THE 11TH INTERNATIONAL CONFERENCE ON MODELING AND APPLIED SIMULATION

SEPTEMBER 19-21 2012

VIENNA, AUSTRIA

EDITED BY

MICHAEL AFFENZELLER

AGOSTINO G. BRUZZONE

FABIO DE FELICE

DAVID DEL RIO VILAS

CLAUDIA FRYDMAN

MARINA MASSEI

YURI MERKURYEV

PRINTED IN RENDE (CS), ITALY, SEPTEMBER 2012

ISBN 978-88-97999-02-7 (Paperback)
ISBN 978-88-97999-10-2 (PDF)

© 2012 DIME UNIVERSITÀ DI GENOVA

RESPONSIBILITY FOR THE ACCURACY OF ALL STATEMENTS IN EACH PAPER RESTS SOLELY WITH THE AUTHOR(S). STATEMENTS ARE NOT NECESSARILY REPRESENTATIVE OF NOR ENDORSED BY THE DIME, UNIVERSITY OF GENOA. PERMISSION IS GRANTED TO PHOTOCOPY PORTIONS OF THE PUBLICATION FOR PERSONAL USE AND FOR THE USE OF STUDENTS PROVIDING CREDIT IS GIVEN TO THE CONFERENCES AND PUBLICATION. PERMISSION DOES NOT EXTEND TO OTHER TYPES OF REPRODUCTION NOR TO COPYING FOR INCORPORATION INTO COMMERCIAL ADVERTISING NOR FOR ANY OTHER PROFIT – MAKING PURPOSE. OTHER PUBLICATIONS ARE ENCOURAGED TO INCLUDE 300 TO 500 WORD ABSTRACTS OR EXCERPTS FROM ANY PAPER CONTAINED IN THIS BOOK, PROVIDED CREDITS ARE GIVEN TO THE AUTHOR(S) AND THE CONFERENCE.

FOR PERMISSION TO PUBLISH A COMPLETE PAPER WRITE TO: DIME UNIVERSITY OF GENOA, PROF. AGOSTINO BRUZZONE, VIA OPERA PIA 15, 16145 GENOVA, ITALY. ADDITIONAL COPIES OF THE PROCEEDINGS OF THE MAS ARE AVAILABLE FROM DIME UNIVERSITY OF GENOA, PROF. AGOSTINO BRUZZONE, VIA OPERA PIA 15, 16145 GENOVA, ITALY.

ISBN 978-88-97999-02-7 (Paperback)
ISBN 978-88-97999-10-2 (PDF)

THE 11TH INTERNATIONAL CONFERENCE ON MODELING AND APPLIED SIMULATION

September 19-21 2012, Vienna, Austria

ORGANIZED BY

DIME – UNIVERSITY OF GENOA

LIOPHANT SIMULATION

SIMULATION TEAM

IMCS – INTERNATIONAL MEDITERRANEAN & LATIN AMERICAN COUNCIL OF SIMULATION

DIMEG, UNIVERSITY OF CALABRIA

MSC-LES, MODELING & SIMULATION CENTER, LABORATORY OF ENTERPRISE SOLUTIONS

MODELING AND SIMULATION CENTER OF EXCELLENCE (MSCOE)

LATVIAN SIMULATION CENTER – RIGA TECHNICAL UNIVERSITY

LOGISIM

LSIS – LABORATOIRE DES SCIENCES DE L'INFORMATION ET DES SYSTEMES

MIMOS – MOVIMENTO ITALIANO MODELLAZIONE E SIMULAZIONE

MITIM PERUGIA CENTER – UNIVERSITY OF PERUGIA

BRASILIAN SIMULATION CENTER, LAMCE-COPPE-UFRJ

MITIM - MCLEOD INSTITUTE OF TECHNOLOGY AND INTEROPERABLE MODELING AND SIMULATION – GENOA CENTER

M&SNET - MCLEOD MODELING AND SIMULATION NETWORK

LATVIAN SIMULATION SOCIETY

ECOLE SUPERIEURE D'INGENIERIE EN SCIENCES APPLIQUEES

FACULTAD DE CIENCIAS EXACTAS. INGEGNERIA Y AGRIMENSURA

UNIVERSITY OF LA LAGUNA

CIFASIS: CONICET-UNR-UPCAM

INSTICC - INSTITUTE FOR SYSTEMS AND TECHNOLOGIES OF INFORMATION, CONTROL AND COMMUNICATION

NATIONAL RUSSIAN SIMULATION SOCIETY

CEA - IFAC

TECHNICALLY CO-SPONSORED

IEEE – CENTRAL AND SOUTH ITALY SECTION CHAPTER

I3M 2012 INDUSTRIAL SPONSORS

CAL-TEK SRL

LIOTECH LTD

MAST SRL

I3M 2012 MEDIA PARTNERS

INDERSCIENCE PUBLISHERS – INTERNATIONAL JOURNAL OF SIMULATION AND PROCESS MODELING

INDERSCIENCE PUBLISHERS – INTERNATIONAL JOURNAL OF CRITICAL INFRASTRUCTURES

IGI GLOBAL – INTERNATIONAL JOURNAL OF PRIVACY AND HEALTH INFORMATION MANAGEMENT

HALLDALE MEDIA GROUP: MILITARY SIMULATION AND TRAINING MAGAZINE

HALLDALE MEDIA GROUP: THE JOURNAL FOR HEALTHCARE EDUCATION, SIMULATION AND TRAINING

EUROMERCI

EDITORS

MICHAEL AFFENZELLER
UNIVERSITY OF APPLIED SCIENCES UPPER AUSTRIA, AUSTRIA
Michael.Affenzeller@fh-hagenberg.at

AGOSTINO BRUZZONE
MITIM-DIME, UNIVERSITY OF GENOA, ITALY
agostino@itim.unige.it

FABIO DE FELICE
UNIVERSITY OF CASSINO, ITALY
defelice@unicas.it

DAVID DEL RIO VILAS
UNIVERSITY OF A CORUNA, SPAIN
daviddelrio@udc.es

CLAUDIA FRYDMAN
LABORATOIRE DES SCIENCES DE L'INFORMATION ET DES SYSTEMES, FRANCE
Claudia.frydman@lsis.org

MARINA MASSEI
LIOPHANT SIMULATION, ITALY
massei@itim.unige.it

YURI MERKURYEV
RIGA TECHNICAL UNIVERSITY, LATVIA
merkur@itl.rtu.lv

THE INTERNATIONAL MULTIDISCIPLINARY MODELING AND SIMULATION MULTICONFERENCE, I3M 2012

GENERAL CO-CHAIRS

AGOSTINO BRUZZONE, *MITIM DIME, UNIVERSITY OF GENOA, ITALY*
YURI MERKURYEV, *RIGA TECHNICAL UNIVERSITY, LATVIA*

PROGRAM CHAIR

FRANCESCO LONGO, *MSC-LES, MECHANICAL DEPARTMENT, UNIVERSITY OF CALABRIA, ITALY*

THE 11TH INTERNATIONAL CONFERENCE ON MODELING AND APPLIED SIMULATION

GENERAL CO-CHAIRS

MICHAEL AFFENZELLER, *UNIVERSITY OF APPLIED SCIENCES UPPER AUSTRIA, AUSTRIA*
MARINA MASSEI, *LIOPHANT SIMULATION, ITALY*

PROGRAM CO-CHAIRS

FABIO DE FELICE, *UNIVERSITY OF CASSINO, ITALY*
DAVID DEL RIO VILAS, *UNIVERSITY OF A CORUNA, SPAIN*

MAS 2012 International Program Committee

Michael Affenzeller, *Upper Austrian Univ. of AS, Austria*
Thècle Alix *IMS Université Bordeaux 1, France*
Gabriel Aprigliano Fernandes, *LAMCE COPPE UFRJ, Brazil*
Naamane Aziz, *LSIS, France*
Enrique Gabriel Baquela, *Universidad Tecnológica Nacional, Argentina*
Imre Barna, *MTA KFKI Atomic Energy, Hungary*
Eleonora Bottani, *University of Parma, Italy*
Agostino Bruzzone, *University of Genoa, Italy*
José M. Cecilia, *Universidad Católica San Antonio, Spain*
Mohamed Chadli, *Universite de Picardie Jules Verne, france*
Diego Crespo Pereira, *University of A Coruna, Spain*
Gerson Cunha, *LAMCE COPPE UFRJ, Brazil*
Fabio De Felice, *University of Cassino, Italy*
David Del Rio Vilas, *University of La Coruna, Spain*
Istvan Farkas, *MTA KFKI Atomic Energy, Hungary*
Claudia Frydman, *LSIS, France*
Luca Gambardella, *IDSIA, Switzerland*
Andrea Grassi, *University of Modena and Reggio Emilia, Italy*
Ginés Guerrero, *University of Murcia, Spain*
Mayer Gusztav, *MTA KFKI Atomic Energy, Hungary*
Rafael Gutierrez, University of Texas, USA
Yilin Huang, *Delft University of Technology, Netherlands*
Alessio Ishizaka, *University of Portsmouth, UK*
Janos Sebestyen Janosy, *MTA KFKI Atomic Energy, Hungary*
Isabella Lami, *Politecnico di Torino, italy*
Pasquale Legato, *University of Calabria, Italy*
Giorgio Locatelli, *University of Lincoln, UK*
Francesco Longo, *MSC-LES, University of Calabria, Italy*
Juan Luis, Instituto Tecnológico de Costa Rica, Costa Rica
Marina Massei, *Liophant Simulation, Italy*
Rina Mary Mazza, *University of Calabria, Italy*
Aldo McLean, *University of Tennessee at Chattanooga, USA*
Paolo Melillo, University of Naples, Italy
Roberto Montanari, *University of Parma, Italy*
Somnath Mukhopadhyay, *University of Texas at El Paso, USA*
Rong Pan, *Arizona State University, USA*
Federica Pascucci, *University of Roma 3, Italy*
Guillerme Pereira, *Universidade do Minho, Portugal*
Horacio Emilio Pérez Sánchez, *Universidad de Murcia, Spain*
Antonella Petrillo, *University of Cassino, Italy*
Gabriele Oliva, Campus Bio-Medico di Roma, Italy
Mustapha Ouladsine, *LSIS, France*
Mamadou Seck, *Delft University of Technology, Netherland*
Roberto Setola, *University of Roma 3, Italy*
Adriano Solis, *York University, Canada*
Alberto Tremori, *University of Genoa, Italy*
Alexander Verbraeck, *Delft University of Technology, Netherland*
Giuseppe Vignali, *University of Parma, Italy*
Gregory Zacharewicz, *IMS Université Bordeaux 1, France*

Tracks and Workshop Chairs

Inventory Management Simulation
Chair: Adriano Solis, York University, Canada

Production Systems Design
Chair: David del Rio Vilas, University of La Coruna, Spain

Decision Support Systems Applications
Chairs: Fabio De Felice, Univerisity of Cassino, Italy; Antonella Petrillo,Univerisity of Cassino, Italy

Simulation in Energy Generation and Distribution
Chair: Janos Sebestyen Janosy, MTA KFKI Atomic Energy Research Institute, Hungary

Promoting Enterprise Interoperability by Service Modeling & Simulation
Chairs: Thecle Alix, IMS Universite Bordeaux 1, France; Gregory Zacharewicz, IMS Universite Bordeaux 1, France

Workshop on Modeling and Simulation of Food Processing and Operations
Chairs: Andrea Grassi, University of Modena and Reggio Emilia, Italy; Giuseppe Vignali, University of Parma, Italy, Eleonora Bottani, University of Parma, Italy; Roberto Montanari, University of Parma, Italy

Simulation Based Design
Chairs: Yilin Huang , Delft University of Technology, Netherlands; Alexander Verbraeck, Delft University of Technology, Netherlands

Automation
Chair: Naamane Aziz, Laboratoire des Sciences de l'Information et des Systemes, France

Workshop on virtual and augmented reality
Chairs: Gerson Cunha, LAMCE/COPPE/UFRJ - Brasil; Gabriel Aprigliano Fernandes, LAMCE/COPPE/UFRJ - Brasil

Simulation and Human Factors Engineering
Chair: Diego Crespo Pereira, University of A Coruna

Simulation based Optimization
Chairs: Pasquale Legato, University of Calabria, Italy; Rina Mary Mazza, University of Calabria, Italy.

High Performance Simulation of biological systems
Chairs: Horacio Pérez-Sánchez, Universidad de Murcia, (Murcia) Spain; José M. Cecilia Universidad Católica San Antonio, (Murcia) Spain

GENERAL CO-CHAIRS' MESSAGE

WELCOME TO MAS 2012!

On behalf of the Program Committee it is our pleasure to present to you the proceedings of the 11th International Conference on Modeling and Applied Simulation – MAS 2012.

MAS was established in 2002 as a joint event to HMS 2002 Bergeggi, Italy in order to provide an opportunity for Project Meetings. Several editions of MAS events were held during the past years in different European Countries (and not only). In particular, during the last three years MAS was held respectively in Tenerife (Spain), Fes (Morocco) and Rome (Italy) all the times collocated as a joint event with the International Multidisciplinary Modeling and Simulation Multiconference (I3M). The organizational structure of MAS 2012 is similar to other international scientific conferences. The backbone of the conference is the scientific program, which is complemented by workshops and open debates dedicated to special topics from the field of applied Modeling & simulation (M&S) and computer technologies; application fields include logistics, supply chain management, production control, business and industrial organization. In 1996 Prof. Agostino Bruzzone's Research Team (University of Genoa, Italy) was driving the European Simulation Symposium (today the EMSS part of I3M) into the Simulation in Industry Event for demonstrating the potential of M&S applied in real world; in similar way MAS is an International Conference in M&S Applications.

We are delighted to report that among many submissions more than 50 papers from more than 20 different countries were selected by the program committee for presentation and discussion during the conference and subsequent publication in the conference proceedings. All submissions were reviewed by an international panel of at least 3 expert referees. We acknowledge the invaluable assistance of the program committee and the international referees, who opted to provide detailed and insightful comments to the authors.

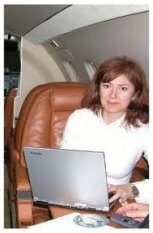
Bridging the gap between theory and practice and incorporating theoretical results into useful products is still one of the main key issues for industrialized countries. Especially in the context of modeling and applied simulation it seems essential that researchers accept the challenge of solving real-world problems, making the science and technology based on mathematics, computer science and computational intelligence contribute to the progress of our developing communities.

MAS 2012 brings together fundamental and applied research contributions that reflect current trends and developments in modeling and applied simulation. We sincerely hope that the conference will encourage cross-fertilization between the various communities, bridging not only the gap between different fields but also, equally important, between different continents and cultures.

In closing, we would like to thank all authors for submitting their works, and all members of the Program Committee who contributed greatly to making the conference a success.

We wish you a fruitful and inspiring conference and a pleasant stay in Vienna.

Michael Affenzeller
University of Applied Sciences, Upper Austria, Austria



Marina Massei
Liophant Simulation, Italy

Fabio De Felice
University of Cassino, Italy

David Del Rio Vilas
University of a Coruna, Spain

ACKNOWLEDGEMENTS

The MAS 2012 International Program Committee (IPC) has selected the papers for the Conference among many submissions; therefore, based on this effort, a very successful event is expected. The MAS 2012 IPC would like to thank all the authors as well as the reviewers for their invaluable work.
A special thank goes to all the organizations, institutions and societies that have supported and technically sponsored the event.

LOCAL ORGANIZATION COMMITTEE

AGOSTINO G. BRUZZONE, *MISS-DIPTEM, UNIVERSITY OF GENOA, ITALY*
ENRICO BOCCA, *SIMULATION TEAM, ITALY*
ALESSANDRO CHIURCO, *MSC-LES, UNIVERSITY OF CALABRIA, ITALY*
FRANCESCO LONGO, *MSC-LES, UNIVERSITY OF CALABRIA, ITALY*
FRANCESCA MADEO, *UNIVERSITY OF GENOA, ITALY*
MARINA MASSEI, *LIOPHANT SIMULATION, ITALY*
LETIZIA NICOLETTI, *CAL-TEK SRL, ITALY*
ALBERTO TREMORI, *SIMULATION TEAM, ITALY*

This International Workshop is part of the I3M Multiconference: the Congress leading **Simulation around the World and Along the Years**

Index

DEVELOPING A SIMULATION TRAINING TOOL FOR ULTRASONOGRAPHY

John A. Sokolowski, PhD[(a)] Catherine M. Banks, PhD[(b)] William Richards, MS[(c)]

[(a)]Virginia Modeling, Analysis and Simulation Center of Old Dominion University, Suffolk, Virginia - USA
[(b)]Virginia Modeling, Analysis and Simulation Center of Old Dominion University, Suffolk, Virginia - USA
[(c)]Virginia Modeling, Analysis and Simulation Center of Old Dominion University, Suffolk, Virginia - USA

[(a)]jsokolow@odu.edu [(b)]cmbanks@odu.edu [(c)]wxrichar@odu.edu

ABSTRACT

This paper discusses the methodology and development of a simulation training tool for ultrasonography. It is designed for training medical practitioners whose practice involves comprehensive and problem-specific physical examination of the patient with the use of ultrasound. Inherent in this user-dependent technology is the need to ensure user capability and appropriate usage of the ultrasound. Simulation is integral to ensuring that capability. Simulation training is able to execute training in a real-world, temporal mode; it can house large digital libraries of ultrasound images for a breadth of experiences; and it can accommodate a repetition of exercises to reinforce learning. Credibly, this simulation training tool for ultrasonography will engage instructional materials in the form of actual patient ultrasonographic images.

Keywords: ultrasonography, pathology, transducer, probe

1. INTRODUCTION

Ultrasonography is an ultrasound-based diagnostic imaging technique used for visualizing subcutaneous body structures to include tendons, muscles, joints, vessels and internal organs for possible pathology or lesions (pathology is defined as the scientific study of the nature of disease and its causes, processes, development, and consequences). It is capable of performing both diagnostic and therapeutic procedures. As such, portable bedside ultrasound devices have revolutionized the practice of medicine, and they are utilized across many sub-fields of medicine (anesthesiology, cardiology, emergency medicine, gastroenterology, gynecology and obstetrics, neurology, ophthalmology, urology, musculoskeletal). Experts in the study and use of ultrasonography recognize this point-of-care medicine is defining the future of patient – physician interaction with pathologies assessed upon examination as bedside visualization. This technology, however, is user-dependent, so ensuring user capability and appropriate usage is necessary. To fully exploit the capability of ultrasonography, medical practitioners using ultrasound must have increased and pathology-specific training to facilitate cognitive and mechanical proficiency. This paper discusses the methodology and development of a simulation training tool to meet that need.

The approach taken is outlined in the following four parts. Part 2 – *Why a Simulation Training Tool is Needed* answers why additional training is needed for point-of-care medicine, who needs this training, and how simulation is integral for this training. Part 3 – *Meeting the Needs of the Training Tool* discusses the necessary focus of the training from substance to interface. Part 4 – *Developing the Tool* discusses the development of the tool such as interface, hardware, and simulator design. Part 5 – *Results and Conclusion* provides an assessment of a prototype tool and discussion for further development of the tool.

2. WHY A SIMULATION TRAINING TOOL IS NEEDED

Medical simulation is able to execute training in a multiplicity of modes, house large digital libraries for a breadth of experiences, and accommodate a repetition of exercises to reinforce learning. Ultrasound practitioners need to understand what they are looking at – impossible to assess with present training modality.

2.1. Why this Tool

Medical students today do not always recognize pathology when they see it, and/or they understand pathology with cognitive skills, but not with the ultrasound image they have extracted. There are also cases whereby students are found to lack the dexterity and mechanical skills needed for the extraction of images. Most curriculums require cognitive examinations wherein the image is provided and the student must simply associate the name of the pathology with the supplied image. Technically, students are given instruction and hands-on training for image extraction, but they do not always recognize the pathology they have retrieve via the extracted ultrasound image due to the fact that he does not have the cognitive skill to do so or because the image is not well-retrieved or a combination of both these handicaps (Levitov 2009). In the US the standard of 20-30 most common pathologies must be obtained and recognized by the student. A gold standard of 150 procedures / images exists, but even then the student may be incompetent (Levitov 2011).

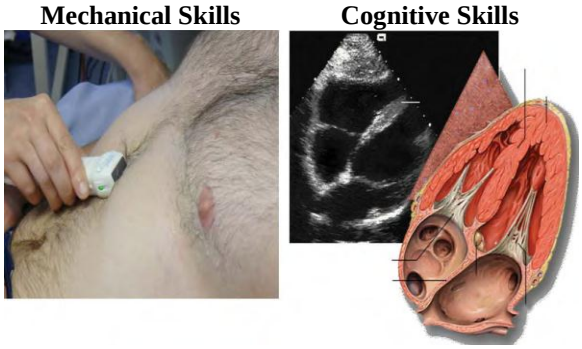


Figure 1: Bedside Ultrasonography Skills

2.2. Who to Train

Thus a simulation training tool for ultrasonography is needed to provide additional and incisive exercises medical students, residents, fellows, and practicing clinicians whose practice involves comprehensive and problem-specific physical examination of the patient. This training can benefit students and clinical practitioners by teaching them to: 1) extract images from a comprehensive library of pathologies, 2) develop of a care plan, and 3) debrief to communicate clinical skill.

2.3. Simulation – Integral for this Training

As such, a tool is need to augment current standards of ultrasonography education through immersive simulation training that includes a comprehensive library of pathological images which the students must extract, recognize, and discuss as part integrating advanced ultrasound principles for patient care.

Arguably, simulation is the ***only way to conduct this training*** because bedside ultrasonography is point-of-care medicine and as such the physician should be able to perform both diagnosis and therapeutic procedures at the bedside, using ultrasound to guide interventional procedures. Training with patient *in vivo* could prove difficult for both patient and clinician: it might be disconcerting to a patient if the physician cannot extract an image, or recognize pathology at the bedside. The immediacy of the technology could prove embarrassing for a physician if he cannot diagnosis or prescribe care through the use of ultrasound in the presence of the patient. Thus, simulation can avoid both these uncomfortable scenarios.

Simulation is the ***best way to provide this training***. Primarily because it is the most ***effective*** means to train and assess the student's knowledge and capability via a body of pre-determined pathologies that must be extracted and discussed via simulated exercises. This training experience will develop, enhance, and make expert their cognitive and technical skills. Second, simulation training is the most ***efficient*** means to exploit the capability of ultrasonography by facilitating repeat-ability of exercises, independent practice sessions. It also is an ***expedient*** means to ensure the future of bedside medicine.

The current users of ultrasonography include physicians practicing internal, critical care, obstetrics-gynecology, and cardiovascular medicine. A broad future audience of clinicians exists because ultrasonography can prove to be a defining skill for practicing physicians as it is superior over the traditional craft (Levitov 2011). Dr. Abraham Verghese of Stanford University (California) is recognized as providing the best physical examination in the US. Verghese is defining (perhaps exposing) the limitation of skills of physicians in physical examination (New York Times 2012). Ultrasonography can mitigate that limitation of skills.

3. INCORPORATING TRAINING REQUIREMENTS INTO THE TRAINING TOOL

To ensure the needs of the training are met, consultation with medical experts is necessary. Their expertise facilitates an understanding of three important concepts before tool development can proceed: the concept of *point-of-care-medicine*, the current type of training provided, and preferences for

the simulation training. These first two concepts are discussed in Section Two above. The following articulates the needs of concept three: preferences for the simulation training.

3.1. Substance of the Tool Design and Interface

Experts in the use and training of ultrasonography would like to see a medical curriculum re-work that requires students to incorporating their knowledge of image acquisition with their cognitive skills of pathology recognition. In short, it means the training should consist of a means to present students with images of multiple pathologies. Essentially, this calls for incorporating images of pathologies into an environment which forces students to interrogate the patient, which can come in the form of a cadaver, mannequin, or simulated-actor. How does one do this without re-inventing the wheel? Modify a traditional instrument into a training tool with specified education and training requirements.

Current ultrasound technology engages a general-purpose sonographic machine for most imaging purposes (see Figure 2).

Figure 2: Medical Ultrasound Scanner

Most ultrasound procedures are done using a transducer on the surface of the body (see Figure 3). The probe contains multiple acoustic transducers to send pulses of sound into a material. Whenever a sound wave encounters a material with a different density (acoustical impedance), part of the sound wave is reflected back to the probe and is detected as an echo. The time it takes for the echo to travel back to the probe is measured and used to calculate the depth of the tissue interface causing the echo. The greater the difference between acoustic impedances, the larger the echo is. If the pulse hits gases or solids, the density difference is so great that most of the acoustic energy is reflected and it becomes impossible to see deeper.

Figure 3 Ultrasound Transducers

This makes clear to needs of the training tool: 1) the images to need be in real-time and, 2) the tool must be a HANDLING tool so that the students can manipulate the transducer to obtain images. As such the training tool must have a reasonable degree of technical difficulty to obtain images. This can be done in two ways: re-animate cadavers or use a simulation platform.

For purposes of this tool development, we determined to use a platform that can be dislodged, yet not completely virtual. The platform must be able to recognize how the transducer is positioned. For added training benefit, the tool must allow transducer to gradually imperfect or distort an image as a teaching experience. When developing aspects of position accuracy with the transducer, the simulation must be able to note patient size and take into account that probes can measure within centimeters for linear or angle-wise images and within 10-15 degrees for triangular or trapezoid images.

The real-world nature of this training tool is premised on using actual patient images – a library of pathologies – images that occur in reality of practice and then tying the right image in real-time.

3.2. Intended Results of Tool Development

Developers of the tool must make every effort to hold to the requirements and preferences of the physicians who have provided their expertise in the development of the tool. As such, the simulation training tool should:

- be highly interactive, diagnostic training tool
- engage user-friendly simulation technology
- developed as ultrasonography information-based
- consist of (real) patient image-based, with images occurring in reality of practice
- facilitate image extraction engaging different transducers and their varying capabilities
- require a reasonable degree of technical difficulty to obtain images for *handling* experience
- present the trainee with images in real-time via a simulated platform
- include inconsistent and inaccurate image extraction as a teaching experience
- provide retrieved and actual patient outcome as realized in the case study image
- serve as a training competency verification tool

4. DEVELOPING THE TOOL

A simulation training tool for ultrasonography must be grounded in electrical engineering and computer science software design to replicate ultrasonography training experiences that are immersive and temporal; *i.e.*, as the ultrasound probe, the transducer, is manipulated images are extracted in real-time. The simulations are drawn from actual patient images serving as the basis for a variety of training scenarios. The tool combines electronic sensing components that detect the position and orientation of the simulated transducer with software that controls proper image selection so that the user is presented with the appropriate image based on his or her skill of probe manipulation. Essentially, this tool requires developing the relationship between transducer and image to create simulator capability so the student can practice manipulating the probe to retrieve pre-set or randomly selected pathologies. In short, we are deconstructing the traditional instrument to create a simulated, real-world training experience.

It is important to note that the simulator will be built with three distinct parts that will operate in cooperation with each other: the graphical user interface (GUI), the hardware interfaces that the user will physically manipulate (such as probes and dials), and the imagery data itself, which will be displayed based on the pathology of the simulated patient and the position of the probe.

4.1. Interface Design

The GUI will be a fairly straight-forward windows based program allowing the trainee to start up the program, select a patient case-study and then initialize the probe. The GUI will read the data coming off the hardware dials, buttons, and the probes providing the user with the appropriate responses to their manipulation of the probe and controls.

4.2. Hardware Design

Once the trainee has selected a patient, he or she will need to select the correct probe to use based on the patient's pathology. The trainee will have to manipulate the probe into the correct position on the mannequin or standardized patient to see the recorded images. If an incorrect probe is used, the simulator may not show the correct image. As some of the probes are meant to be inserted into the patient's body, the simulator will have to include a mannequin that can accept a probe in the correct locations. In addition to the probe, the trainee will have to manipulate some dials and buttons to alter the field of view and strength of the transducers as would be the case on an actual Ultrasound Scanner. These buttons and dials can be interfaced with the GUI using a standard off the shelf USB input/output (IO) board, such as an Arduino or other similar devise (Arduino 2012).

4.3. Simulator Design

The images that will be displayed on the simulator's screen will be stored in a large database containing information about the pathology, age, and sex of the patient. In addition, the angle of view and location on the patient will be stored so that the simulator will know where on the mannequin the probe will have to be located in order to display the appropriate images.

4.4. Hardware Probe Design

Developing the probe and tracking its movement is the most critical part of the simulator design, and will require an Inertial Magnetic Unit (IMU) to track the movement of the probe. An IMU uses the earth's magnetic field and an inertia sensor to tell the software in which direction the device is pointed, and the direction it has been moving. Unfortunately, the accuracy of these devices is compromised as they have a tendency to drift over time (Florida Conference 2002). However, while it will be possible to reduce this drift through the use of a Kalman filter algorithm, it may be useful to

implement another method of tracking the probe as well to improve its accuracy (SIGGRAPH 2001). This can be accomplished by complimenting the IMU with the use of a Microsoft Kinnect to provide visual tracking of the users hand or a Wiimote using its IR camera to track IR emitters on the probe.

The use of a Microsoft Kinnect, along with an IMU may provide the best methodology for tracking the probe's position and orientation. This system is relatively cheap, comes with an easy to use API, and does not necessarily require an unblocked line of sight to the probe. In fact, it should be possible to track the trainee's hand as well as the probe, as the Kinnect software is setup to do just that. However, this system must be tested against its intended use, as the mannequin, which looks just like a human, may potentially interfere with the tracking of the trainee's hand by the Kinnect. Optionally, a wiimote can be pointed at the mannequin and IR emitters fitted to the probe so that the wiimote can track the location and orientation of the probe, as long as the line of sight to the probe from the wiimote is kept clear.

Technically the accuracy of the probe's location with respect to the virtual model and the mannequin or standardized patient is required to have a high level of fidelity in order for it to be effective. This means that the more accurately the probe and the location of the mannequin can be monitored the higher fidelity the virtual image can potentially display. However, the accuracy of the probe's relative position is not the only factor in ensuring the validity of the technical aspects of this simulator. The Virtual model must also show deformation of internal organs as the probe is pressed up against the subject's body, and to verify this, a subject matter expert such as a practitioner of Ultrasonography is required to provide feedback (Gerovichev 2004).

5. SUMMARY

The need for developing simulation medical training tools is quite obvious – technology outstrips educational products *i.e.,* the technology and devices are readily available, but the training is lacking. It has been said that the patient-doctor interface has not changed since 1860s. Yet as this is written, bedside ultrasonography is taking hold of that problem and turning it around. Ultrasonography is the first serious attempt to advance bringing medicine, vis-à-vis the patient-doctor interface, into the 21st century, although some medical experts place these advances to about the 1980s (Levitov 2011).

Developing medical simulation training tools requires close cooperation with the users, the clinicians, determine precise educational content. Simulation engineers and those in the modeling and simulation community simply do not have the expertise nor the materials needed to ensure that the training materials are appropriate. As shown in this discussion of ultrasonography education, those training-specific materials are the patient images – a library of pathologies – images that occur in reality of practice to populate the tool. With the requirements and images in hand, actual development of the tool can begin. The engineering expertise comes in the form of tying the right image in real-time – programming, simulation, and interface design.

This methodology paper outlines how to develop a simulation training tool for ultrasonography. And although this is not the only effort in existence, a cursory review of what is available indicates that many training products are too costly and some have been criticized as not well-thought out. If the expert users of ultrasonography are correct, that the use of ultrasound will change the future of bedside medicine, then it is fair to conclude the U.S medical curriculum will adjust to include ultrasonography education and training as an integral part of its curriculum.

REFERENCES

Arduino
http://www.arduino.cc/ (accessed March 2012)

Walchko, KJ. Mason, P., 2002. *Inertial Navigation*. Proceedings of Florida Conference on Recent Advances in Robotics, pp. 91-98. May 22-24, Miami (Florida, USA).

Levitov A, Apostolos D, Slonin A., 2011. *Bedside Ultrasonography in Clinical Medicine*. New York: McGraw-Hill.

Levitov A, Mayo P, Slonin A., 2009. *Critical Care Ultrasonography*. New York: McGraw-Hill.

New York Times. *Scientist at Work: Dr. Abraham Verghese, Physician Revives a Dying Art: The Physical Exam*. Available from: http://www.nytimes.com/2010/10/12/health/12profile.html?pagewanted=all [accessed March 2012]

SIGGRAPH 2001, Course 8, *An Introduction to the Kalman Filter*, Greg Welch, Gary Bishop. Available from:
http://www.cs.unc.edu/~tracker/media/pdf/SIGGRAPH2001_CoursePack_08.pdf
[accessed March 2012]

Gerovichev, O. 2004. The effect of visual and haptic feedback on computer-assisted needle insertion. *Computer Aided Surgery*, 6(9):243–249, 2004.

AUTHORS BIOGRAPHY

John A. Sokolowski PhD, is the Executive Director for VMASC, supervising 50 researchers and staff with an annual funded research budget of $10 million. He supervises research and development in Transportation, Homeland Security, Defense, Medical M&S, Decisions Support, Business & Supply Chain, and Social Science (real-world) M&S applications. He is contributor and co-editor of *Modeling and Simulation in the Medical and Health Sciences*, Wiley Publications 2011.

Catherine M. Banks PhD, is Research Associate Professor at VMASC. Her focus is on qualitative research among the social science disciplines to serve as inputs into various modeling paradigms: game theoretical, agent-based, social network, and system dynamics. Dr. Banks' research includes models representing humans and human behavior to include the translating / mapping of data for quantitative representations, modeling states and their varied histories of revolution and insurgency, political economy and state volatility, and medical simulation. She has authored and edited books and journal articles on these topics and is contributor and co-editor of *Modeling and Simulation in the Medical and Health Sciences*, Wiley Publications 2011.

William T. Richards MS, is a Project Scientist at VMASC. His focus is on software development, especially user interfaces, as well as artificial intelligence, geospatial information systems, and hardware interfaces. Mr. Richards has developed a number of prototypes such as simulators and simulations for use in training individuals how to perform various tasks or conduct research studies.

GENERATION OF ALTERNATIVES FOR MODEL PREDICTIVE CONTROL IN MANUFACTURING

Sören Stelzer[(a)], Sören Bergmann[(b)], Steffen Straßburger[(c)]

Ilmenau University of Technology
Helmholtzplatz 3
D-98684 Ilmenau, GERMANY

[(a)]soeren.stelzer@tu-ilmenau.de, [(b)] soeren.bergmann@tu-ilmenau.de, [(c)] steffen.strassburger@tu-ilmenau.de

ABSTRACT

Manufacturing systems are dynamic systems which are influenced by various disturbances or frequently changing customer requests. A continuous process of decision making is required. Model Predictive Control is a common model-based approach for control but needs adaption to be applicable to discrete-event simulation. In this paper we introduce an approach to model and generate non trivial control options and decisions often made in the operation of manufacturing systems. We also show how complex scenarios can be generated. To support a wide-range of applications our approach is based on the core manufacturing simulation data (CMSD) information model. We implement the design and generation of complex scenarios by processing and combining modeled control options. By using our approach, which also applicable to decision support systems, we can enable model-based closed-loop control based on a symbiotic simulation system and automated model generation and initialization.

Keywords: simulation, CMSD-IM, design of experiments, decision support system, model predictive control

1. INTRODUCTION

Manufacturing systems are dynamic systems and subject to various internal and external disturbances, which often influence the expected behavior in an undesired way. Additionally, they have to deal with growing uncertainties, flexibility, and high cost pressure.

These facts lead to changing circumstances for the decision making process. Decisions have to be made in higher frequency, which directly leads to a shorter time available for finding them. Therefore, a continuous process of decision making and controlling is required to make sure the aimed goals can be achieved. Additionally, the complexity of internal and external processes is rising. For the same reason the amount of gained data in the connected information systems is also increasing.

To keep up with the tightened situation decision makers are forced to use Decision Support Systems (DSS). To also handle the complexity of today's manufacturing systems discrete event simulation (DES) is used in conjunction with DSS. This leads to model-driven DSS (Heilala et al. 2010). Related approaches are Online-Simulation (Davis 1999; Hanisch, Tolujew, and Schulze 2005), Simulation based Early Warning Systems (SEWS) (Hotz 2007), and Symbiotic Simulation Systems (Aydt et al. 2008a).

In our research work we investigate an automated control approach called Model Predictive Control (MPC) to enable a closed-loop control for manufacturing systems, using techniques like Symbiotic Simulation and SEWS. While MPC is well studied in the field of automatic control engineering, there is very limited research and application of MPC using discrete event simulation techniques. To enable MPC to manufacturing control or decision support, we identified three major research tasks: the formal description of alternatives, the generation of complex scenarios based on combinations of these alternatives, and the generation and appropriate initialization of simulation models with the state of the real system.

In this paper we focus on the methodology of how common decision and control alternatives in the operation of manufacturing systems can be formally described and automatically generated. We also show how based on this description complex scenarios can be designed and generated. Our modeling approach is based on an applied information model, provided by the Core Manufacturing Simulation Data Standard (SISO 2010). Initialization, generation and execution of simulation models have been discussed in previous work (Bergmann, Stelzer, and Strassburger 2011).

The remainder of this paper is structured as follows. In section 2, we discuss related work. In section 3, we illustrate requirements for model-predictive control in manufacturing using discrete event simulation. We also introduce the information model used in our approach. In section 4, we discuss typical decisions and control options used to influence manufacturing systems and how they can be described using the CMSD-IM. In section 5, we describe our

methodology of how we intelligently and automatically generate simulation scenarios from the modeled control options. In section 6, we summarize the results of our approach and give an overview about current and future work.

2. DISCUSSION

2.1. Related Work

The goal of decision makers and controllers is to find an optimal solution for a given situation from a set of alternatives. Therefore it is required to determine the current situation as correct as possible and to estimate the impact of all available decisions or control inputs. Due the high complexity of the internal processes and large amount of data, the estimation of the impact of an alternative to the manufacturing system can not be handled anymore by humans (Heilala et al. 2010).

Discrete Event Simulation (Banks et al. 2000) is a well-accepted technique for planning, investigation and operation of manufacturing systems and supply chains (VDI 3633-1). The application of modeling and simulation for manufacturing systems is not a novel approach, but a dominating part of simulation applications are focused on the planning of new systems or their modification. When the actual operating of manufacturing systems is investigated, the discussed applications are often limited to scheduling problems.

To extend the benefits of modeling and simulation to the operation of manufacturing systems a closer integration of simulation techniques and manufacturing systems is needed. In 2004, Fowler and Rose discussed the future of modeling and simulation and defined a couple of challenges for research and applications. Among them, the closer integration and interaction of simulation techniques and information systems was identified as grand challenge (Fowler and Rose 2004).

The closer integration of simulation techniques and manufacturing enables the handling of complex processes and huge amounts of data required for decision support or control. The Online-Simulation approach (Hanisch, Tolujew, and Schulze 2005) focuses on a closer integration of simulation and manufacturing systems by obtaining or keeping simulation models up-to-date with the investigated system. Besides the introduction of different methods for initializing the simulation with data from the real system, Hanisch et al. also discussed requirements and aspects of data and model quality.

In previous work, we have already presented a solution for model generation and initialization as a potential way to obtain up-to-date simulation models (Bergmann and Strassburger 2010; Bergmann, Stelzer, and Strassburger 2011). The chosen approach is based on the Core Manufacturing Simulation Data Information Model (CMSD-IM).

Symbiotic Simulation (Aydt et al. 2008a) and simulation based early warning systems (SEWS) (Hotz 2007) are further approaches which are focusing on the application of a closer integration of manufacturing information systems and simulation environments. While SEWS are an application of a close integration of information systems to monitor manufacturing systems, symbiotic simulation is discussing the possible interactions and benefits of simulation environments and manufacturing systems. Further, the aspects of symbiotic simulation systems enable a wide range of new applications of modeling and simulation.

The goal of both approaches is to enhance the quality of the decision making process. A key benefit of these approaches is the possibility to consider the current state of the system under investigation. This enables situation-based decision support and control, like prediction of the trajectory of crucial processes. It is also possible to predict the behavior of the investigated system after applying a control input or decision. This is also found in literature as “what-if”-analysis (WIA). Aydt et al. (2008b), for example, showed in a semi-conductor manufacturing application how decision support can be enabled by variation of simple model parameters. There is no information given on modeling requirements or how to implement a symbiotic simulation.

The observed applications of simulation-based decision support in manufacturing are limited to the variation of parameters or scheduling. This is primarily caused by the lack of appropriate methods for describing and modeling of complex control options and decisions. Beyond the simple variation of parameters or schedules, there is a wide variety of decisions and control options (discussed in section 4) which can not be described by simple parameters. The definition of complex scenarios often leads to an extensive manual modeling process. This makes it difficult to automatically generate such scenarios and iterate through them. We consider this an important requirement for simulation-based control.

Neglecting these more complex control options, manufacturing control is already applied and discussed in several papers. A closed control loop involving manufacturing execution systems (MES) as controller is suggested by Kletti (2007) (Figure 1). His goal is to use a model of the system to evaluate a set of possible alternatives and choose the optimal one, regarding current objectives. These alternatives should consider the current system state and include possible decisions, strategies or control values. In reality there are no MES applications which implement an automatic control loop using discrete simulation techniques. Instead they used deterministic forward calculations (often wrongfully named “simulation”) and neglect the dynamic and stochastic behavior of the system.

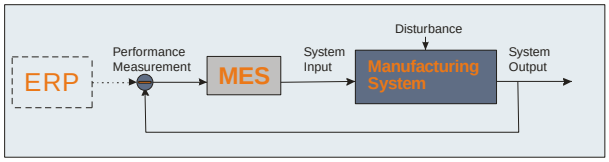


Figure 1: Propagated control loop for manufacturing execution systems (according to Kletti 2007)

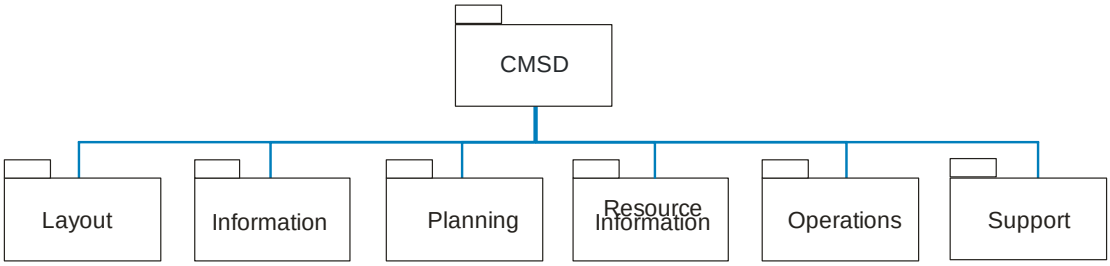


Figure 2: Overview of packages provided by CMSD-IM

There is also related work in the area of automatic control engineering. Finding (semi)-optimal control vectors, analog to alternatives, using a model of the system is called model-predictive control (MPC). Model predictive control was applied to manufacturing systems by Wang et al. They were using analytical models (Wang, Rivera, and Kempf 2007), which are not available and applicable in most cases of manufacturing control and decision support.

To make MPC an accepted approach for manufacturing control it is necessary to use the discrete event simulation technique as a model base. This is caused by the lack of other modeling approaches which can describe the complex dynamic and stochastic properties of systems in this domain.

Regarding the requirements for MPC, symbiotic simulation systems are well suited for establishing a closed-loop control. What is missing is a methodology of how common decision and control alternatives in the operation of manufacturing systems can be formally described and automatically generated.

For this it is necessary to describe and model given alternatives in a way, which allows automated and enumerable combination. With this the controller will be enabled to directly evaluate the search space for finding an optimal scenario.

We also show how based on this description complex scenarios can be designed, generated and afterwards be evaluated through simulation.

Before we describe the modeling of alternatives in CMSD-IM, we have to discuss requirements for model predictive control based on a standard focusing on interoperability of manufacturing information systems and simulation environments.

3. REQUIREMENTS FOR MODEL PREDICTIVE CONTROL

The requirements for enabling a closed-loop control or decision support of manufacturing systems can be separated into four major aspects. Most of them are based on requirements for data exchange, model generation, symbiotic simulation, and online simulation. At first we have to discuss the representation of the investigated manufacturing system and secondly the automated generation and initialization of simulation models. The third aspect is the aggregation of results and the last aspect concerns how to generate complex scenarios to formulate WIAs.

Based on our work on automated model generation and initialization we are using the CMSD-IM as standard for data exchange and modeling manufacturing systems.

3.1. CMSD

The primary objective of the CMSD Information Model is to facilitate interoperability between simulation systems and other manufacturing applications. The CMSD standard provides data structures and an information model (Figure 2) which was designed to firstly support the exchange of modeling information. To cover the complexity of production and logistic systems and a wide range of modeling approaches, the standard allows aspects of the system to be mapped in CMSD in multiple ways.

The capabilities of CMSD were demonstrated in several research projects (Leong et al. 2006; Johansson et al. 2007), which mostly focus on the developing of new and modified production systems. Our own work has focused on using CMSD to support the operational phase of manufacturing systems (Bergmann, Stelzer, and Strassburger 2011).

3.2. Modeling of Manufacturing Systems

The basic idea of model-driven approaches like symbiotic simulation or model predictive control is to use a model of the investigated system to obtain information about its behavior. To assure the correctness of these results, a verified and validated model is needed. In our work we are using the CMSD-IM to store and structure information. This is done by analyzing the real system or collecting information from information systems connected to the real system. Information about processes, resources and materials can, for instance, be imported from Enterprise Resource Planning (ERP) or Manufacturing Execution Systems (MES).

CMSD-IM enables a data-driven modeling approach decoupled from modeling or simulation tools. The information model of CMSD-IM consists of several classes representing common objects found in manufacturing systems like machines and workers. It is also capable of defining process plans as well as representing job and order information.

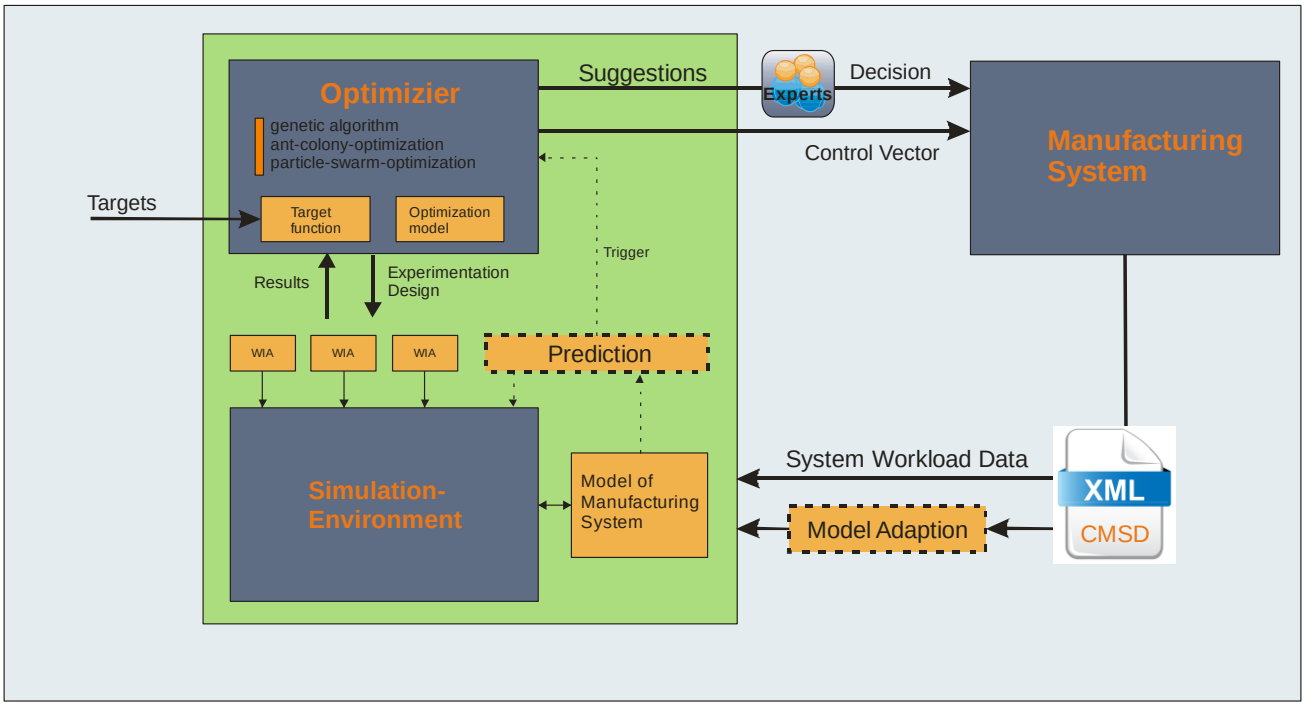


Figure 3: Model Predictive Control / Decision Support System based on automated model generation and initialization

CMSD-IM further allows the checking for logical correctness and therefore facilitates the detection of any missing information extracted from the external information systems.

By using our web-based interface, decision makers can interactively complete necessary data. It is also possible to manipulate the conceptual model via the web-based user-interface.

All objects and attributes of the CMSD-IM are presentable in an XML description using an associated schema. This characteristic enables the automated processing which is required for model generation.

3.3. Description of Alternatives and Generation of complex Scenarios

In contrast to systems which are analytically describable manufacturing systems are modeled using discrete event simulation models. With this, we have no simple way to determine an optimal solution.

In the case of non-trivial manufacturing scenarios, we therefore have multiple dependencies of the system on potentially many parameters. We therefore cannot simply single out a simple variable for optimization.

We therefore require a way to explicitly describe and model the different alternatives, which can afterwards be selected and combined by the controller or the decision support system. To achieve the automated processing of model descriptions we choose CMSD-IM to describe the alternatives (see section 4).

To combine the modeled alternatives it is required to process them and build logically correct and valid scenarios, which afterwards will be automatically generated and executed (section 5).

3.4. Automated Model Generation and Initialization

To support decision makers or enable control to systems using discrete event simulation different opportunities have to be evaluated by simulating them. Therefore an executable simulation model is required. Finding (semi)-optimal solutions to an actual situation typically requires a huge amount of simulation runs, with different model variations.

From previous work we already know that CMSD-IM is well suited for representing manufacturing systems. To transform the CMSD-IM description into executable simulation models we are using an automated model generation approach. The model generation allows the creation of simulation models for different simulation tools, for example Plant Simulation (Siemens 2012) or SLX (Henriksen 1999).

To use simulation as an operational decision support tool we also need to appropriately initialize the simulation model, i.e., a mechanism to keep the simulation model up-to-date with the real system. In previous work we described a methodology based on CMSD-IM to do this. This approach combines the automated model generation and initialization. This is enabled by CMSD-IMs ability to keep system load and state information of modeled entities as well.

3.5. Distribution of Experiments and Result Aggregation

After the scenario is designed and an executable model is built by the automatic model generation, the model has to be executed. In our framework we are able to distribute large amounts of scenarios which have to be evaluated. This is described by Bergmann, Stelzer, and Strassburger (2012).

The results of parallel executed simulation experiments have to be collected and aggregated to performance indicators. This requires the model to have provided information necessary to estimate these values. We are again using CMSD-IM to hold result information. This enables the packaging of model, scenario description, and results in a single data exchange format.

The aggregated results are used by the controller to specifically generate new scenarios and can be presented to decision makers.

Based on the above discussed requirements, model predictive control or model-based decision support systems can be implemented (shown in Figure 3). The established cycle of generation, execution and evaluation of alternative scenarios (WIA) is identical to common simulation-based optimization loop. The advantage of this approach is the ability to consider current state information of the real system and the complex search-space description

4. DESCRIPTION OF ALTERNATIVES

While investigating applications in manufacturing system, we identified typical fields for decision alternatives as shown in Table 1.

Factors		
Job Schedule	**Shift Schedule**	**Machine Utilization**
Job-priorities	Human Resource Flexibility	Alternative Capacities
customer-priorities	Extra shifts	Express Deliveries
date-oriented priorities	Overtimes	Partial Deliveries
		External Capacities/ Out-Sourcing

Table 1: Common alternatives in manufacturing control

The set of available alternatives, of course, highly depends on the specific application scenario. Therefore some of the discussed alternatives may not be available or reasonable in every case.

After defining a set of common control options, we investigated ways for formally describing the possible alternatives. In relation to our previous work, we focused on how the Information Model provided by CMSD can be used for their representation.

4.1. Job Scheduling

The CMSD-IM provides a set of classes for the definition of schedules. A schedule in CMSD is reflected in the class Schedule, which consists of several ScheduleItems. In case of job scheduling, each ScheduleItem contains information when the assigned job has to be processed. Due the fact that more than one instance of the Schedule-Class could exist, CMSD is also capable of holding a set of different alternative schedules reflecting, for instance, different strategies of job-scheduling (e.g. job priorities vs. earliest due dates). A whole schedule contains a list of job references, tagged with time stamps for starting and finishing the associated job. The job reference can also link to an order reference, which is a list of jobs associated with a requesting party. With that it is also possible to prioritize customers.

With these capabilities of CMSD we are capable of defining alternative schedules based on job priorities, customer priorities, date-oriented priorities or combinations of all three.

Another common approach in job scheduling, also describable in CMSD, is to define the order of tasks by resource. In this case, a ScheduleItem in a schedule points to a defined ProcessStep instance and contains information about when it should be started. The ProcessStep-Class itself defines all required Resources like machines, workers or materials. For scheduling it is required to have previously defined all possible combinations of tasks and machines, which can perform these.

With this approach we can even further define the exact order of tasks (ProcessStep) of jobs, giving us full flexibility in representation of alternative job schedules. Beyond the task based scheduling, CMSD-IM allows the combination of ProcessSteps from different ProcessPlans. In this case, there are no limitations for scheduling of parallel machines. However, additional processing is necessary and later discussed in section 4.3.

4.2. Shift Scheduling

Workers in CMSD can have defined times, in which they are available for performing their operations. In a wide-range of applications the management of human resources is a major issue. Besides the task of scheduling times or shifts, it is also crucial to take into account the personal skills when assigning workers to process steps. The CMSD-IM provides several modeling approaches to address this. Every resource, including workers, is tagged with a CalendarReference. A Calendar defines a static list of times associated to Shifts or a reference to a ShiftSchedule. ShiftSchedules are more flexible and similar to the Schedule-Class used for job scheduling. There can be a set of alternative ShiftSchedules, but a resource can only be associated to one specific ShiftSchedule (via a CalendarReference).

A ShiftSchedule consists of a list of time-tagged ShiftReferences. The Shift-Class defines start and end times, breaks, and applicable days. In combination with a ShiftSchedule, it is possible to define the availability of resources on an arbitrarily level of detail. To perform a continuous scheduling of resources, it is necessary to define every time of activity (excluding breaks) as a Shift and merge them via a ShiftSchedule.

```
<ProcessPlan>
      <Identifier> ProcessPlan01_PartType01 </Identifier>
      <PartsProduced>
            <PartType> <PartTypeIdentifier> PartType01 </PartTypeIdentifier> </PartType>
      </PartsProduced>
      <Process>
            <Identifier> PP01_PartType01_Step01 </Identifier>
            <ResourcesRequired>
                  <Resource> <ResourceIdentifier> Ma02 </ResourceIdentifier> </Resource>
                  <AllowableSetup>
                        <SetupDefinitionIdentifier> Setup_Ma02_PartType01 </SetupDefinitionIdentifier>
                  </AllowableSetup>
            </ResourcesRequired>
            <OperationTime> <TimeUnit> minutes </TimeUnit> <Value> 35 </Value> </OperationTime>
      </Process>
      <Process/> ... <Process/>
<ProcessPlan>
```

Figure 4: ProcessPlan Definition (XML)

In some applications there are planned or potential deviations from the common Shift. In these cases we suggest to define a concurrent set of Shifts and group them into an alternative ShiftSchedule. In this approach it is only necessary to replace the schedule used for most cases with the special schedule.

With these capabilities of CMSD we are capable of representing arbitrary alternative shift schedules which can subsequently be used for evolution in the simulation.

4.3. Variation of Processes

Modeling the variation of processes is quite different to job or shift scheduling in CMSD. Processes are described by the ProcessPlan class which typically consists of a list of ProcessStep instances (see figure 4). A typical control option in the operation of a manufacturing system is the definition and selection of process alternatives for job or product types.

In case of machine failures or scheduled maintenance operations, alternative resources can be used to reduce the impact of such temporal bottlenecks. Like parallel machines, in CMSD-IM every possible process has to be pre-defined. If one or more tasks have to be changed in a process, a new ProcessPlan derived from the original one has to be defined. In this case, alternative process descriptions can be linked to (sub)-jobs or orders, changing the assigned EffortDescription. A similar line of action has to be performed if parallel machine setup is used. In principle, it is also possible to describe alternative resources in a ProcessStep instance, but they are not useful for scheduling of processes on parallel machines.

If the manufacturing system is operated by task-based scheduling, managing temporal variations of the material flow is also possible by using the Schedule-Class. As discussed in 4.1, the ScheduleItem-Class is able to describe the effort of tasks via ProcessSteps. Unfortunately CMSD provides no reference to an assigned job or order, which would be necessary to determine the effort for this.

To address this lack, we suggest using the CMSD provided property concept to attach a job or an order-reference to every ScheduleItem. Using this approach it is possible to assign every ScheduleItem to its proper job or order and determine the planned effort.

In summary, we can use CMSD to describe process alternatives in two different ways (direct modeling of alternatives in the ProcessPlan class vs. indirect modeling using the Schedule class). This gives the possibility to describe the usage of alternative capacities (compare table 1).

4.4. Additional Modeling

Partial or complete outsourcing of orders is modeled in the CMSD-IM using the Order or Job classes. Every instance is assigned to an executing party, which can be targeted to an external source. Further it is also possible to model costs of jobs, orders, resources, and tasks. By this, we can easily model express-deliveries of required materials or preconditioned orders and their financial impact on any cost functions.

5. GENERATION OF ALTERNATIVES AND COMPLEX SCENARIOS

While the previous section discussed the explicit description of alternatives to an existing model and how to reflect these alternatives in CMSD-IM, this section focuses on detecting further alternatives (possibly implicitly stored in the CMSD-IM) and composing more complex scenarios based on previously described alternatives and their variations.

The finding of alternatives beyond the explicitly modeled process definitions is based on investigating the predefined ProcessPlan instances. This is done by processing of the contained ProcessSteps and finding matches of output and input behavior, like consumed or produced parts or part types.

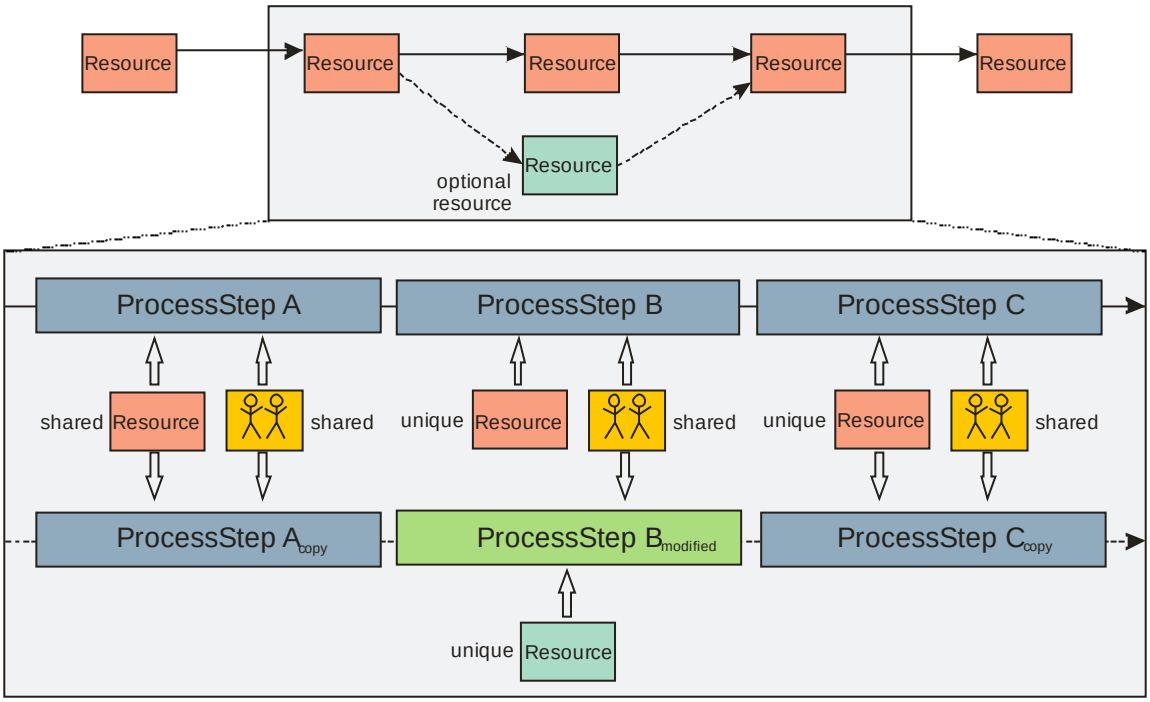


Figure 5: Modification of process definitions by exchange of required resources to create alternatives.

In our approach we assume that each ProcessStep reflects a task, which has to be performed by a defined set of resources. We also assume that tasks are exchangeable by another task exactly matching its input-output mapping. Based on these assumptions resources used to perform equivalent tasks are also exchangeable. Further information about required skills, setup-states and differing processing times can be determined by evaluating the associated resource references.

5.1. Iterating Process Alternatives

While generating schedules of jobs or human-resources are widely covered by commercial scheduling libraries, we here focus on generating alternative processes based on the modeled manufacturing systems. There are two prerequisite for our approach. At first the system and its current state is reflected in a CMSD-IM description. Second the manufacturing allows adaption, i.e., the system has a certain redundancy or unused resource potential.

Beyond the explicit modeling of varied processes, it is possible to determine alternative resources by processing the CMSD-IM model. Equivalent to the CMSD-IM approach, there are two stages for finding alternatives. The first stage is to process available ProcessPlan's to find a matching of the produced parts or rather types of parts. In case of the ProcessPlan-level, they accord with a job. If there were explicitly modeled alternative processes for a job type, they would have matching input and output.

The next step is to use the determined information of equivalent resources to generate alternative process description (see top of figure 5). This requires a second pass, recursively processing ProcessPlans. For every existing ProcessStep there are two possible actions. If there is no alternative resource availed, the ProcessStep is kept untouched. Else, in the first pass an alternative ProcessStep was found. In this case, the processing is branched (forked) and an alternative ProcessPlan will be created, using the previously scanned ProcessSteps (see of figure 5). For the current ProcessStep we are able to create a new ProcessStep based on the alternative ProcessStep. To keep the consistency of the model, it is required to copy the required setup-states and worker skills. This transformation requires the absence of referenced human resources and machine-based skill.

If human resources have to be considered, the non-human resources are handled like described above, but the required human resources have to be compared by skills. If the skills of the worker performing the replaced operation are sufficient to perform the alternative task, we can reuse the references. If not, the human resources from the alternative ProcessStep have to be taken to the newly created alternative.

5.2. Dynamical building of Process Definitions

Another second practicable approach to determine alternative process definitions for exceptional situations or for analyzing possibilities is to use the current system load (for example: job situation) and to schedule these using all available operations. This requires the existence of an external scheduling library or software. The result of the scheduling process is transformed into a Schedule-Class instance of CMSD-IM. In this case it is necessary to use ScheduleItems based on ProcessSteps and tag them with the assigned job/order.

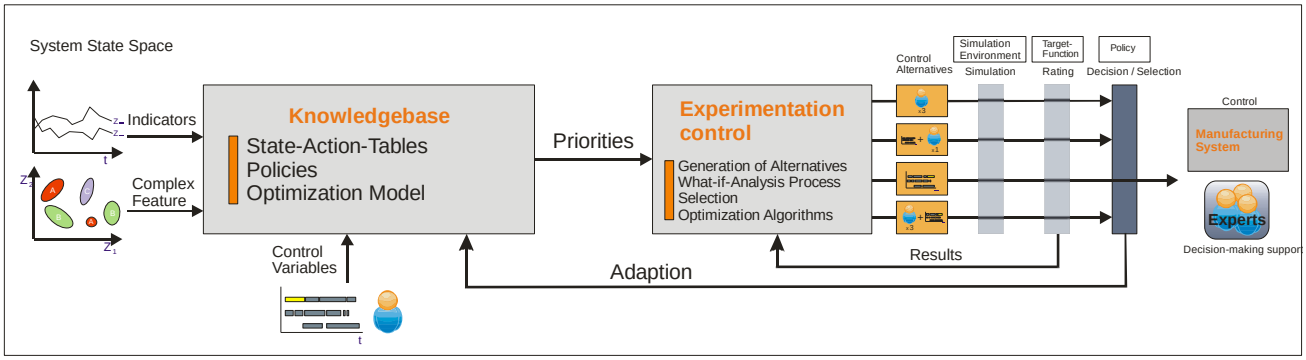


Figure 6: Description of algorithm used for selection of described alternatives and complex scenario (WIA) combination.

By processing a selected Schedule along the job reference it is possible to build an individual ProcessPlan and attach it to the job object by the JobEffortDescription.

5.3. Design and Generating of Complex Scenarios

For truly enabling MPC based on using discrete event simulation to evaluate different control options, a approach for generating complex scenarios is needed. Complex scenarios can contain combinations of the previously discussed alternatives.

In previous work we already discussed the automated model generation and initialization. Further we discussed several strategies to distributed execution of simulation experiments as a prerequisite for an efficient "what-if-analysis" process (WIA).

The missing link, like discussed in section 1, is the generation of complex alternative scenarios based on described decisions or control options. To enable a systematical experimentation, it was necessary to investigate the possibilities how to consistently model control or decision options (section 4).

In most scenarios it is required to apply a subset of alternatives to influence the system behavior. A controller and also decision support systems have to build complex scenarios, based on combination of available alternatives. In case of MPC, the controller has to do this automatically.

In our approach this is done by building complex scenarios by selecting subsets of the described decision and control alternatives. To define the subset, we are using a binary genetic encoding. To solve the problem that some alternatives, like scheduling or parameters are themselves iterable, we also use a genetic encoded description. By concatenation of the different encoding string, complex scenario descriptions are created. This string is used to generate a derivation of the base model of the manufacturing system, represented as a CMSD-IM description.

Using a binary genetic encoding also enables us to clip several dimensions of the search-space. To reduce the required iterations to find (semi)-optimal solutions we use a subset-selection algorithm, which starts with a minimal set of alternatives used for building complex scenarios. This subset is extended by unused alternatives, triggered by defined performance indices (figure 6). The used information model also enables us to determine the validity of the generated model.

5.4. Comparison and Discussion

To evaluate our approach to describe and automatically generate alternative control options or decisions and build complex scenarios, we implemented a CMSD-IM based symbiotic simulation environment.

The core of the test environment is an automated model generator, which is also able to initialize the generated models with the current system state. The model generator transforms a CMSD-IM description of a manufacturing system into an executable simulation model. This was recently discussed in previous work (Bergmann 2010; Bergmann, Stelzer and Strassburger 2011). The information flow is defined as shown in Figure 7.

We also implemented both ways of generating alternative process definitions to compare the impact on complexity and usability. Iterating the available process definitions and generating new ones based on compatible single-operations can generate a large amount of alternative process-definitions. This is amplified by the amount of alternative single-operations, especially induced by parallel machine-configurations or a high degree of redundancy. In this case we suggest focusing on explicit modeling and not using implicit information.

The second way of generating alternative processes (section 5.2) needs the presence of an external scheduling library which has to be provided with a set of process definitions. As previously discussed, alternative ProcessSteps can be regarded as parallel machine problems. This can lead to well-known problems regarding the solution complexity.

Our focus is to use a combinatorial iteration comparable to genetic encoding to build schedules. This only requires a mechanism to filter invalid schedules, which can be easily implemented.

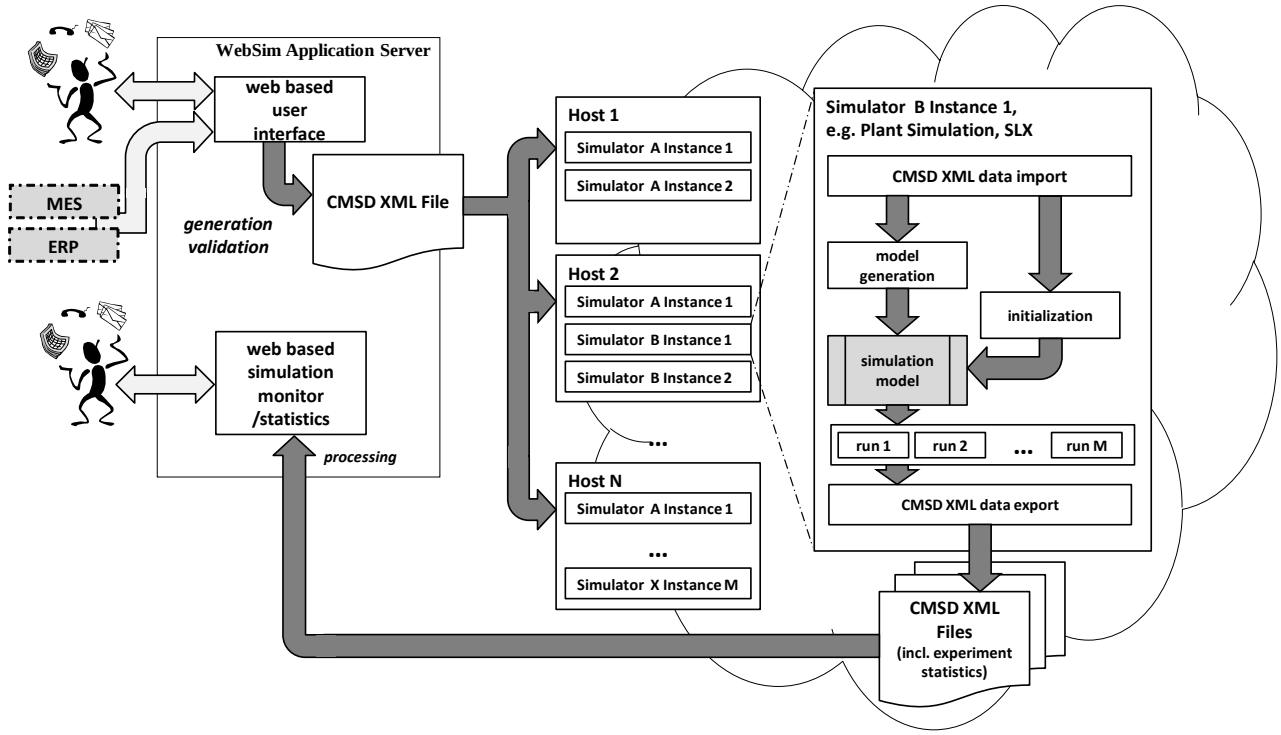


Figure 7: Information flow of the test environment for model generation and distributed execution.

6. CONCLUSION AND FUTURE WORKS

To enable model-based control, like MPC, for manufacturing systems, a way had to be found to resolve the problem of how to automatically iterate the search-space spanned by common control alternatives. In this paper we introduced an approach how to formally describe and model typical decision and control alternatives in the operation of manufacturing systems. We also presented a methodology for using implicitly stored information for detecting further control alternatives, which can be combined with explicitly modeled alternatives to complex scenarios. Using a Symbiotic Simulation System as a base, we enabled situation-based model-predictive control. The Symbiotic Simulation was implemented based on previously discussed work on automated model generation and initialization.

As shown in the discussion, our approach can also be useful to enhance model-driven decision support systems. The results of the work described in this paper enable us to investigate the MPC approach in real and larger scale manufacturing applications, which is a subject of future work.

Future work will also focus on intelligently handling the large amount of computation time needed to evaluate simulation experiments.

We also focus on applications for Exception based Manufacturing Execution Systems and the preventive evaluation of future scenarios in Early Warning Systems.

REFERENCES

Banks, J., Carson, J., Nelson, B. L., Nicol, D., 2000. *Discrete-Event System Simulation (3rd ed).* Upper Saddle River, New Jersey: Prentice-Hall, Inc.

Aydt, H., Turner, S. J., Cai, W., Low, M. Y. H., 2008. "Symbiotic simulation systems: An extended definition motivated by symbiosis in Biology". *Proceedings of the 22nd Workshop on Principles of Advanced and Distributed Simulation*, 109–116. June 3-6, Rome (Italy).

Aydt, H., Turner, S. J., Cai, W., Low, M. Y. H., Lendermann, P., Gan, B. P., 2008. "Symbiotic Simulation Control in Semiconductor Manufacturing". *Proceedings of the International Conference on Computational Science*, 5103:26–35. June 23-25, Krakow (Poland).

Bergmann, S., Strassburger, S., 2010. "Challenges for the Automatic Generation of Simulation Models for Production Systems". *Proceedings of the 2010 Summer Simulation Multiconference*, 545-549. July 11-14, Ottawa (Canada).

Bergmann, S., Stelzer, S., Strassburger, S., 2011. „Initialization of Simualtion Models using CMSD". *Proceedings of the 2011 Winter Simulation Conference*, 2228-2239. December 11-14, Phoenix (Arizona, USA).

Bergmann, S., Stelzer, S., Strassburger, S., 2012. „A new web-based method for distribution of simulation experiments based on the CMSD standard". *Proceedings of the 2012 Winter Simulation Conference.* December 9-12, Berlin (Germany).

Davis, W. J., 1998. “On-Line Simulation: Need and Evolving Research Requirements”. In: Banks, J., ed. *Handbook of Simulation*. John Wiley & Sons Inc, 465-516.

Fowler, J. W., Rose, O., 2004. “Grand Challenges in Modeling and Simulation of Complex Manufacturing Systems”. *SIMULATION* 80(9): 469–476.

Heilala, J., Maantila, M., Montonen, J., Sillanpää, J., Järvinen, P., Jokinen, T., Kivikunnas, S., 2010. “Developing Simulation-based Decision Support Systes for Customer-driven Manufacturing Operation Planning”. *Proceedings of the 2010 Winter Simulation Conference*. December 5-8, Baltimore (Maryland, USA).

Hanisch, A., Tolujew, J., Schulze, T., 2005. “Initialization of Online Simulation Models”. *Proceedings of the 2005 Winter Simulation Conference*, 1795-1803. December 4-7, Orlando, (Florida, USA).

Henriksen, J. O., 1999. “SLX - The X is for eXtensibility.” *Proceedings of the 1999 Winter Simulation Conference*, 167-175. December 5-8, Phoenix (Arizona, USA).

Hotz, I., 2007. “Ein Simulationsbasiertes Frühwarn-system zur Unterstützung der operativen Produktionssteuerung und -planung in der Automobilindustrie” (A simulation based early warning system to support the operational production planning and scheduling in the automotive industry (in German)). Thesis (PhD). Otto-von-Guericke University Magdeburg (Germany)

Johansson, M., Leong, S., Lee, Y. T., Riddick, F., Shao, G., Johansson, B., Skoogh, A., and Klingstam, P., 2007. “A Test Implementation of the Core Manufacturing Simulation Data Specification”. *Proceedings of the 2007 Winter Simulation Conference*. 1673-1681. December 9-12, Washington, DC (USA).

Kletti, J., 2007. *Konzeption und Einführung von MES-Systemen*. Berlin:Springer.

Leong, S., Lee, Y. T., Riddick, F., 2006. “A Core Manufacturing Simulation Data Information Model for Manufacturing Applications”. *Proceedings of the 2006 Fall Simulation Interoperability Workshop*. September 12-15, Orlando (Florida, USA).

Siemens Product Lifecycle Management Software Inc., 2012. “Plant Simulation”. Available from: http://www.plm.automation.siemens.com/en_us/products/tecnomatix/plant_design/plant_simulation.shtml [March 2012].

SISO, 2010. “Standard for: Core Manufacturing Simulation Data – UML Model”. Core Manufacturing Simulation Data Product Development Group. Available from: http://www.sisostds.org/DigitalLibrary.aspx?Command=Core_Download&EntryId=31457 [February 2011].

VDI 3633-1, 2000. “Simulation of systems in materials handling, logistics and production - Fundamentals”. VDI-Society Production and Logistics. Berlin:Beuth Verlag.

Wang, W., Riviera, D. E., Kempf, K. G., 2007. „Model Predictive Control strategies for Supply Chain Management in Semiconductor Manufacturing”. *International Journal of Production Economics* 107:56-77.

AUTHORS BIOGRAPHY

SÖREN STELZER is a PhD student at the Ilmenau University of Technology. He is a member of the scientific staff at the Department for Industrial Information Systems. He received his diploma degree in Computer Science from the Ilmenau University of Technology. During his study he was working in the Neuro-informatics and Cognitive Robotics Lab of the Ilmenau University of Technology. After his study he worked in optimization of power plants in several projects. His research interests are simulation based optimization, model predictive control, artificial learning and discrete event simulation. His email is soeren.stelzer@tu-ilmenau.de.

SÖREN BERGMAN is a PhD student at the Ilmenau University of Technology. He is a member of the scientific staff at the Department for Industrial Information Systems. He received his diploma degree in Information Systems from Ilmenau University of Technology. Previously he worked as corporate consultant in various projects. His research interests include generation of simulation models and automated validation of simulation models within the digital factory context. His email address is soeren.bergmann@tu-ilmenau.de.

STEFFEN STRASSBURGER is a professor at the Ilmenau University of Technology in the School of Economic Sciences. Previously he was head of the “Virtual Development” department at the Fraunhofer Institute in Magdeburg, Germany and a researcher at the DaimlerChrysler Research Center in Ulm, Germany. He holds a Ph.D. and a Diploma degree in Computer Science from the University of Magdeburg, Germany. He is a member of the editorial board of the Journal of Simulation. His research interests include distributed simulation as well as general interoperability topics within the digital factory context. He is also the Vice Chair of SISO’s COTS Simulation Package Interoperability Product Development Group. His web page can be found via www.tu-ilmenau.de/wi1. His email is steffen.strassburger@tu-ilmenau.de.

WRINKLE EFFECT IN CLOTH SIMULATION USING FLUIDITY CONTROL

Jaruwan Mesit

Grambling State University
403 Main Street, Grambling, Louisiana 71245

mesitj@gram.edu

ABSTRACT

In complex environment, the wrinkle effect is essential in cloth simulation for realism in animated soft objects. Animating wrinkle effect requires more computation which leads the unacceptably slow simulation. In this paper, we use fluidity control in general soft body with variation of parameter values to present the wrinkle effect in cloth simulation. The fluidity control in our simulation is fast since only nearby surface points are used in this computation. For the general soft body model in this paper, the structure of the model is provided the structure control while the surface deformation is presented by the fluidity control. For the realism of the model, the gravitational control is exerted for free fall motion of the model. The result presented in the paper shows that the various parameter values in the control can generate wrinkle effects in the cloth simulation.

1. INTRODUCTION

Nowadays, the computer simulations has become more realistic. The techniques for soft body are becoming increasingly general. The general soft body model enables simulation of a variety of soft body materials by adjusting the parameter values for specific soft body behaviors. This research focuses on how to apply general method of soft body model into specific property of the material such as cloth. The property can be specified by the parameters of body control, surface deformation, volume control, and gravitation. To control these parameters we use body mesh structure to maintain configuration among surface points, fluid modeling to deform the details of the surface, internal pressure to approximated the simulated molecules within the soft body. Finally, free fall motion of soft body is generated by gravitational field.

Remainder of this paper is structured as follows. Section 2 describes previous works. Section 3 discusses a general soft body model. Sections 4 presents wrinkle effect using fluidity control. Section 5 demonstrates the results of fluidity control in cloth simulation. Section 6 concludes our work.

2. PREVIOUS WORKS

In cloth simulation, several techniques have been proposed by many researchers to present the cloth behaviors of drapes, folds, and wrinkles, for different properties of textiles. In some cloth-specific properties, the cloth behaviors deals with large deformations for cloth in high flexibility. In recent work, underwater cloth simulation has been described in the term of internal and external dynamics of cloth underwater.

In complex environments, such as air, water, or oil, Eberhardt *et al*. describe the fast, flexible, particle-based model to animate the drapes of different types of cloth, which require the calculation of exact trajectories of moving particles (Eberhardt, Weber, & Strasser, 1996). They have investigated a suitable description of internal forces for each particle in the form of force plots of tension, shearing, and bending. These later are used to calculate trajectory of a particle via an integration of the Lagrange differential equation. the particle locations, inside or at the border of cloth provide different levels of air resistance to particular particles. The visual results of different textures present draping and vibrating effects in their simulations.

Bhat *et al.* investigate cloth simulation from video data of real fabrics in (Breen et al., 2003). An algorithm to estimate the parameters of cloth simulation from video data of real fabric is proposed in this paper for the needs of parameter adjustment to achieve the appearance of a particular fabric. Two consecutive video frames are used to provide the cloth parameters from the folds of fabrics. The dynamic and static tests on small fabric swatches give appropriate simulation parameters which are later used to simulate the fabrics that are worn by a human actor. Four different types of fabrics (linen, fleece, satin, and knit) are simulated to demonstrate the performance of this approach.

Bridson *et al.* present the simulation of clothing with folds and wrinkles in (R Bridson, Marino, & Fedkiw, 2003). A mixed explicit/implicit time integration in numerical analysis is used for the cloth appearance. A detail in contact regions of the cloth is achieved by a physically correct bending model combining with an interface forecasting technique. In this proposed method, a post-processing method preserves folds and wrinkles on cloth-

character collision. Additionally, a dynamic constraint mechanism supports the control in large scale folding. Later, the simulation realism is achieved by using these techniques to control folds and wrinkles on cloth simulation.

Bridson *et al.* propose the technique for robust treatment of collision, contact and friction for cloth animation for actual modeling of cloth thickness in (Robert Bridson, Fedkiw, & Anderson, 2005). For the smooth and interference free data in sharp folds and wrinkles in a cloth mesh, the post-processed scheme with subdivisions has been used in this paper. Since the sharp folds can generate the intersections between elements of subdivisions, the repulsion and collision impulses are used to adjust the cloth positions with no intersection at the end of adjustment. In addition, the static friction model provides the stable folds and wrinkles as shown in the simulation of a curtain that is draped over a ball.

Decaudin *et al.* present the folds generated by the collision with the virtual mannequin (Decaudin et al., 2006). The different buckling patterns for a cylinder of fabric describes the patterns of diamond buckling, twist buckling, and axis aligned folds. Diamond buckling is presented by the compression of a cylinder of cloth maintaining a zero Gaussian curvature which has a diamond shape that appears when an elbow or a knee of a character is bent or the sleeves of a sweater are pulled up. The parallel folds are generated when the body twists and when a loose skirt hangs under gravity. The proposed method to produce these folds works on an fabric cylinder. The visual result of this simulation shows the realistic 3D mannequin dressed in the designed garment.

Müller *et al.* introduce a method to avoid the velocity layer required to provide the new position of the model (Müller, Heidelberger, Hennix, & Ratcliff, 2007). Since there is instability problem in the explicit Euler method, point based dynamics provide the new position immediately after the model constraints, including constraints of collision. The new locations of points are calculated by the solver approach that tries to satisfy all constraints of the model. Using this method, all points in the models can be generated immediately during the simulation. To show the effectiveness of the method, cloth simulation with animated game characters is presented in the paper.

Ozgen *et al.* simulate underwater cloth with fractional derivatives in (Ozgen, Kallmann, Ramirez, & Coimbra, 2010) by using a particle-based cloth model that includes half-derivative viscoelastic elements to describe the internal and external dynamics of the cloth. The cloth responds to fluid stresses and the behavior of particles in a viscous fluid by this method. The fluid viscosity component in fractional cloth model produces bump propagations on the surface of the model. The equation of motion in this paper is Fractional Differential Equation (FDE) to where both explicit and implicit numerical solution techniques can be extended. To present the realism of the simulation, the underwater cloth deformation has been demonstrated and compared with real clothes.

3. A GENERAL FORMAL MODEL OF SOFT BODIES IN MOTION

This section, the general formal model of soft body model can be described as a structure of the model that is composed of a list of surface points, a list of edges connecting between the points, and a list of triangles connecting between the edges of the model. The structure of the model is defined as

$\boldsymbol{p_i}$ is the list of the surface points.

F_i is a set of forces for each surface point of the soft body ,

v_i is a set of velocities for each surface point of the soft body,

Each force at point i is defined as $F_i = \langle F_{bi}, F_{ai}, F_{vi}, F_{gi} \rangle$, where $\boldsymbol{F}_{bi}^t$ is the force of the body structure control, F_{ai}^t is the force of the fluidity control, $\boldsymbol{F}_{vi}^t$ is the force of the volume control, and $\boldsymbol{F}_{gi}^t$ is the force of gravity. All of these controls are presented at point i.

To satisfy a level of softness of the model, we present a parametric model of different control. As indicated previously, the composite force, $\boldsymbol{F}_i^t$, of the surface point i at time t is based on the controls of components. With this assumption, $\boldsymbol{F}_i^t$ is defined as:

$$\boldsymbol{F}_i^t = \alpha \boldsymbol{F}_{bi}^t + \beta \boldsymbol{F}_{fi}^t + \gamma \boldsymbol{F}_{vi}^t + \delta \boldsymbol{F}_{gi}^t, \qquad (1)$$

where parameters, $\alpha, \beta, \gamma,$ and δ are parameters of body structure control, fluidity control, volume control, and gravitational field, respectively, where $\alpha > 0$ and $\beta, \gamma, \delta \geq 0$. Constraints may exist for relations among these components.

In this manner, the various force parameters exerted on each soft body surface point can be adjusted to obtain specific types of soft body behaviors.

4. WRINKLE EFFECT USING FLUIDITY CONTROL

Fluidity is the study how the fluid flows and sometimes is known as hydrodynamics. The properties of the fluid are normally described in the terms of density(ρ) , pressure ($\boldsymbol{F}_{\boldsymbol{fpi}}^{\boldsymbol{t}}$), and viscosity ($\boldsymbol{F}_{\boldsymbol{fvi}}^{\boldsymbol{t}}$).

In this paper, we present the wrinkle effect using fluidity control in the general soft body model. For the behavior of fluidity control, the free moving particles, which are the surface points in our model, interact with nearby surface points within a radius. We use the model technique called the smoothed particle hydrodynamics (SPH) developed originally for astrophysics problems and later used in interactive applications of particles based on fluid simulation (Müller, Charypar, and Gross 2003; Müller et al. 2005). For the implementation of the model, SPH is simpler than other fluid modeling such as FEM or FVM. An interpolation method is used in SPH to distribute quantities in a local neighborhood of each particle using radial symmetrical smoothing kernels. Poly6, spiky, and viscosity smoothing kernels in Müller, Charypar, and Gross (2003) and Müller et al. (2005) describe fluid density, fluid pressure, and viscosity forces in our model. Thus, during fluid force calculation, fluid density and fluid pressure are computed to generate fluid pressure force and fluid viscosity force as presented as follows.

Related to the model developed in Müller, Charypar, and Gross (2003) and Müller et al. (2005), fluid density is given by:

$$\rho_i^t = \sum_j m_j W_{poly6}(l_{ij}^t, h_w), \ \forall j \text{ such that } l_{ij}^t \le h_w \quad , \tag{2}$$

where ρ_i^t is the density at surface point i at time t, m_j is the mass at surface point j, and h_w is the core radius of SPH.

Next, fluid pressure is generated from fluid density as:

$$L_i^t = k(\rho_i^t - \rho^0), \tag{3}$$

where L_i^t is the fluid or liquid pressure at surface point i at time t, k is the gas constant, ρ_i^t is the density at surface point i at time t, and ρ^0 is the initial density.

Fluid pressure force at the soft body surface point i, $\boldsymbol{F}_{fpi}^t$, is computed as:

$$\boldsymbol{F}_{fpi}^t = -\sum_j m_j \frac{L_i^t - L_j^t}{2\rho_j^t} \nabla W_{spiky}(l_{ij}^t, h_w), \ \forall j \text{ such that } l_{ij}^t \le h_w, \tag{4}$$

where m_j is the mass at surface point j, L_i^t and L_j^t are fluid or liquid pressure values at surface points i and j respectively at time t, ρ_j^t is the density at surface point j at time t, and h_w is the core radius of SPH.

Finally the fluid viscosity force at the soft body surface point i, $\boldsymbol{F}_{fvi}^t$, is generated by:

$$\boldsymbol{F}_{fvi}^t = \mu \sum_j m_j \frac{\boldsymbol{v}_j^t - \boldsymbol{v}_i^t}{\rho_j^t} \nabla^2 W_{viscosity}(l_{ij}^t, h_w), \ \forall j \text{ such that } l_{ij}^t \le h_w, \tag{5}$$

where μ is the viscosity of fluid, m_j is the mass at surface point j, $\boldsymbol{v}_i^t$ and $\boldsymbol{v}_j^t$ are the velocities at surface points i and j respectively at time t, ρ_j^t is the density at surface point j at time t, and h_w is the core radius of SPH.

The fluidity control generated by fluid force, $\boldsymbol{F}_{\boldsymbol{ai}}^{\boldsymbol{t}}$, is the combination of two different forces: (1) fluid pressure force, $\boldsymbol{F}_{fpi}^t$, and (2) fluid viscosity force, $\boldsymbol{F}_{fvi}^t$. Hence, we define the fluid force as follows:

$$\boldsymbol{F}_{\boldsymbol{ai}}^{\boldsymbol{t}} = \boldsymbol{F}_{\boldsymbol{fpi}}^{\boldsymbol{t}} + \boldsymbol{F}_{\boldsymbol{fvi}}^{\boldsymbol{t}} \quad , \tag{6}$$

where $\boldsymbol{F}_{fpi}^t$ is fluid pressure force at surface point i and $\boldsymbol{F}_{fvi}^t$ is fluid viscosity force at surface point i.

In order to model volume of the soft body an internal pressure force must push the surface points outward. This volume is created by pressure force generated by the molecules within the soft body. Without volume, the soft body may become flat, much like fabric or cloth after colliding with the environment.

5. RESULTS

Three demonstrations with different parameter values of fluidity control have been presented in this section. The environment of the scene composes of a hanger with 2,151 vertices and a cloth with 12,290 vertices hanging on the hanger where the collision detection occurs between the

hanger and the cloth to hold the cloth. The simulation has three different sets of parameters $\alpha, \beta, \gamma,$ and δ which are body structure control, fluidity control, volume control, and gravitational field, respectively, where $\alpha > 0$, and $\beta, \gamma, \delta \geq 0$.

The animated cloth illustrates the main effects of fluidity control from fluid parameter to implement the winkle effect of the cloth. The parameter sets are presented in table 1 and results of parameter set A, B, and C are demonstrated in figures 1 - 3, figures 4 - 6, and figures 7 - 9, respectively.

Table 1: Experiment parameters for winkle effect in cloth simulation

Parameter sets	α	β	γ	δ
A	1	0	1	1
B	1	1	1	1
C	1	2	1	1

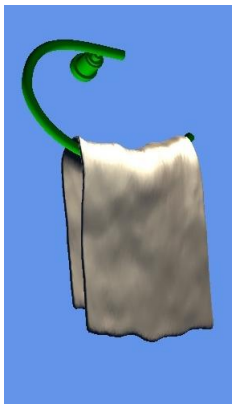


Figure 1. The screenshot is captured at 1st frame of the simulation for experiment set A.

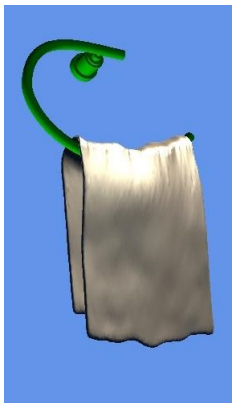


Figure 2. The screenshot is captured at 200th frame of the simulation for experiment set A.

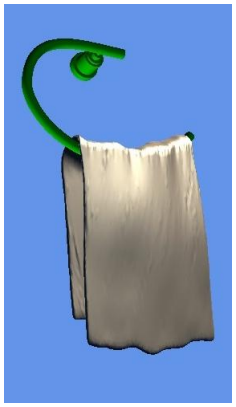


Figure 3. The screenshot is captured at 400th frame of the simulation for experiment set A.

Figure 4. The screenshot is captured at 1st frame of the simulation for experiment set B.

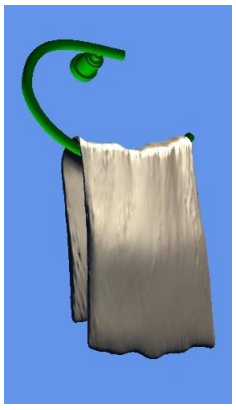


Figure 5. The screenshot is captured at 200th frame of the simulation for experiment set B.

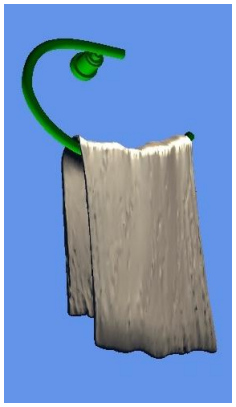


Figure 6. The screenshot is captured at 400th frame of the simulation for experiment set B.

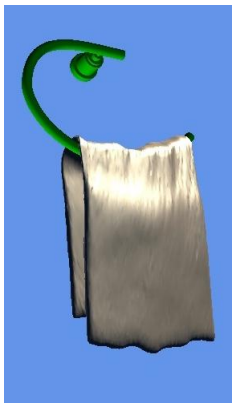


Figure 7. The screenshot is captured at 1st frame of the simulation for experiment set C.

Figure 8. The screenshot is captured at 200th frame of the simulation for experiment set C.

Figure 9. The screenshot is captured at 400th frame of the simulation for experiment set C.

The result of the parameter set A which is presented in figures 1 to 3 shows that the simulation of cloth becomes smooth when the fluidity control is set to 0. When the parameter value of fluidity is increased to 1 as set in parameter set B, the result of the cloth presets more wrinkle on the surface as demonstrated in figures 4 - 6. When the parameter value of fluidity is increased to 2 as set in parameter set C, the result of the cloth presets even more wrinkle on the surface as demonstrated in figures 7 - 9. The results of these parameter sets shows that the fluidity control can create the wrinkle effect on the cloth simulation.

6. CONCLUSION

As we have presented in the simulation set, the our general soft body control that composes of body control, surface deformation, and gravitation, can be adjusted to simulate different types of soft bodies. Focusing of the fluidity control the winkle effect that can be used in cloth simulation in different levels. If the cloth needs more wrinkle effect, then it can be achieved by increasing the fluidity parameter value. In the future work, the wrinkle effect can be applied in the cloth on animated characters.

REFERENCES

Breen, D., (editors, M. L., Bhat, K. S., Twigg, C. D., Hodgins, J. K., M., S., Khosla, P. K., et al. (2003). Estimating Cloth Simulation Parameters from Video. *Eurographics/SIGGRAPH symposium on Computer Animation 2003* (pp. 37-51).

Bridson, R, Marino, S., & Fedkiw, R. (2003). Simulation of clothing with folds and wrinkles. *SCA '03: Proceedings of the 2003 ACM*

SIGGRAPH/Eurographics symposium on Computer animation (pp. 28-36). Aire-la-Ville, Switzerland, Switzerland: Eurographics Association.

Bridson, Robert, Fedkiw, R., & Anderson, J. (2005). Robust treatment of collisions, contact and friction for cloth animation. *ACM SIGGRAPH 2005 Courses on - SIGGRAPH '05* (p. 2). New York, New York, USA: ACM Press. doi:10.1145/1198555.1198572

Decaudin, P., Julius, D., Wither, J., Boissieux, L., Sheffer, A., & Cani, M.-P. (2006). Virtual Garments: A Fully Geometric Approach for Clothing Design. *Computer Graphics Forum*, *25*(3), 625-634. doi:10.1111/j.1467-8659.2006.00982.x

Eberhardt, B., Weber, A., & Strasser, W. (1996). A fast, flexible, particle-system model for cloth draping. *IEEE Computer Graphics and Applications*, *16*(5), 52-59. doi:10.1109/38.536275

Müller, M., Charypar, D., & Gross, M. (2003). Particle-based fluid simulation for interactive applications. *SCA '03: Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation* (pp. 154-159). Aire-la-Ville, Switzerland, Switzerland: Eurographics Association.

Müller, M., Heidelberger, B., Hennix, M., & Ratcliff, J. (2007). Position based dynamics. *Journal of Visual Communication and Image Representation*, *18*(2), 109-118. doi:10.1016/j.jvcir.2007.01.005

Müller, M., Solenthaler, B., Keiser, R., & Gross, M. (2005). Particle-based fluid-fluid interaction. *SCA '05: Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation* (pp. 237-244). New York, NY, USA: ACM. doi:http://doi.acm.org/10.1145/1073368.1073402

Ozgen, O., Kallmann, M., Ramirez, L. E., & Coimbra, C. F. (2010). Underwater cloth simulation with fractional derivatives. *ACM Transactions on Graphics*, *29*(3), 1-9. doi:10.1145/1805964.1805967

OPEN BENCHMARK DATABASE FOR MULTIDISCIPLINARY OPTIMIZATION PROBLEMS

Tuomo Varis[(a)], Tero Tuovinen [(b)]

[(a)]Department of Mathematical Information Technology, University of Jyväskylä
[(b)] Department of Mathematical Information Technology, University of Jyväskylä

[(a)]tuomo.j.varis@jyu.fi, [(b)]tero.tuovinen@jyu.fi

ABSTRACT
Solving new increasingly complex problems requires development of new methods and tools but verification of their correctness and efficiency in absence of actual experimental data is difficult. In this paper we propose an open database of benchmark cases for multidisciplinary optimization validation that will serve as a reference point for discovery and validation of optimization methods and facilitate adoption of such methods in the industry. The paper describes the goals of the database, the process of acquiring content to the database, its initial content and technical implementation.

Keywords: multidisciplinary, optimization, benchmarking, validation

1. INTRODUCTION

It is well known, that the mathematical and numerical analysis and optimization is going forward to multidisciplinary and multi-scale (e.g. more complex) problems. The natural reason for this trend is the tremendous increase in computational power in late decades. We are now in a situation, where everyone has the possibility to buy practically speaking unlimited resources of computing time through the Internet at a very decent price. A complex model with numerically estimated results will give us some insight in many phenomena that have never been modeled before. And, at the same time, verification and validation of multidisciplinary optimization methods and tools becomes a more and more crucial step for the process. The benchmarks defined before will not be enough for this new generation of problems. The optimization will provide one abstract layer over the regular analysis and complicates the process by having certain implications on for example how meshing is handled.

In this paper, we present an open database for multidisciplinary optimization problems which we have developed in order to tackle this challenge. Previously, similar benchmark databases have been developed for example by Ingenet (INGENET, 2008) and Flownet (Marini et al., 2002) projects. In this article, we will introduce a guide for our open database. The scope of the system comes from complex multidisciplinary optimization problems. We have listed the defined benchmark cases. For developing similar benchmark databases, we will propose one technical solution that has been tested in use. The database and its content are open for everyone on the Internet at the database web site (Design Test Case Database, 2009). Submitting new content requires a free registration and validation from our team before being published.

The aim of this study is to create a database, where scientists can propose and publish definitions of multidisciplinary and multi-objective optimization benchmark cases in study along with example solutions. Later on, other scientists in the field have a way to reconstruct the same benchmark with their tools and compare the results in a decent manner. The best cases will be computed multiple times with different methods and the pool of results available will grow. The openness of the system will give everyone a possibility to contribute and get feedback from their simulations. As a result, the most efficient and reliable methods and tools for solving each type of problem can be found.

2. GOALS OF THE DATABASE

Goals of the database three-fold:

1. to serve as a reference point for discovery and validation of optimization methods for different types of problems
2. to give scientists working on the field an opportunity to compare and demonstrate their methods, tools and expertise, and
3. to promote usage of design optimization techniques to industry.

Design optimization is an effective tool for enhancing properties of existing products by improving their designs through advanced algorithms and computational simulations instead of time and money consuming experiments on physical prototypes. The types of optimization problems vary greatly depending on the application domain and correct methods need to be chosen for the optimization to be efficient and successful. One major goal of our database is to work as a reference point for scientists and engineers working with design optimization in their search for the method best suitable for their problem at hand. This is done by providing benchmark cases from various fields, from

electromagnetics to acoustics and aerodynamics. In order to be useful as a reference point, the benchmark cases are defined in a generic yet rigorous manner and the example solutions provided by contributors include detailed information on the methods used to reach the solution along with analysis on both the progression of the optimization and the optimized design. When working in optimization it is also important to have a reference for validation of the methods used. When no actual experimental data is available, applying new techniques to an existing well defined problem and comparing the results to examples in the database can be done to gain insight on the performance and reliability of the methods. Due to the multi-step nature of simulation-driven design optimization, where a mistake in any phase (importing and remodeling geometry, meshing and simulation) can have drastic effects and lead to either inferior or altogether incorrect results, it is necessary to build on a pool of existing experiments to be confident about the methods used.

The database also gives the scientists and engineers working on optimization a new forum for interacting with other experts on the field. While papers are the preferred method for publicizing research in the scientific community, the database gives everyone a lucrative opportunity to prove their algorithms and codes on a variety of problems and publish the results on-line. We have also organized Database Workshop events that revolve around the benchmark cases in the database. These events are an excellent opportunity for networking but also have a competitive nature by allowing the scientists themselves present their results and compare them with those of others. The database also gives the scientists and engineers working on optimization a new forum for interacting with other experts on the field. While papers are the preferred method for publicizing research in the scientific community, the database gives everyone a lucrative opportunity to prove their algorithms and codes on a variety of problems and publish the results on-line. We have also organized Database Workshop events that revolve around the benchmark cases in the database. These events are an excellent opportunity for networking but also have a competitive nature by allowing the scientists themselves present their results and compare them with those of others.

While optimization methods have proven to be useful in industrial application, the adoption of such techniques has not yet reached companies outside very engineering-heavy industries such as aerospace. One purpose of our database is to show that optimization has applications beyond the narrow scope sometimes perceived by the industry. Some of the benchmark cases already available on-line have been created by engineers from large companies and despite generality are directly applicable to problems often encountered in product design and manufacturing. One of our goals is to help company engineers to realize the benefits of optimization and simulation-based prototyping and encourage co-operation with academic experts. Increased collaboration would help to bridge the gap between the academic and industrial worlds.

3. THE PROCESS OF ACQUIRING DEFINITIONS OF THE BENCHMARK CASES

The first target for this study was to find the benchmark definitions from the industry. We noted that in industry, there are lots of open questions concerning multidisciplinary problems and optimization and the need for this kind of system is urgent. However, defining the benchmark cases was difficult, the line between confidential and public knowledge was thin and the expertise of the engineers was targeted in a very narrow scope on the field and thus were usually unable to describe the benchmark in general scientific context. As the result, we received benchmark test case definitions that were not easily reproduced and therefore, the potential contributors coming from different engineering fields or academia did not see the benefits for collaborating with the system. The feedback concerning the system was poor.

The second target was to define the cases on an academic basis. We did noticed, that the definitions of the problems and the generality was much easier to achieve and the benefits were not restricted to a single branch of the industry. Working in the scientific community was much more natural. As the result, we got some very well defined benchmark cases and several solutions for comparison purposes. The system showed its possibilities. Still, we were lacking the audience. The website of the system received relatively few hits and in seminars related to the benchmarks we had to challenge each scientist at a time to participate and contribute to our database. The amount of knowledge in the system increased slowly in time and but realizing the benefits of the system invited more scientists to participate and contribute. Scientists have to understand the value of the forum to justify the time spent in order to participate and convert their solvers and other tools to support the formats requested by the system.

Our third target will be the full openness of the system. The system is highly dependent on the quality and amount of computation results provided. The scientist that has done research on a problem and provided the related benchmark definition has to take the initiative to provide the first results on that case. The benefit for the scientist in this stage is the number of possible citations produced by the knowledge and comparison. When the first results are available along with a description of the methods required to compute the case it is much easier for other participants to follow up and either improve the existing solution or take their own approach to the problem. But when the benchmark case definition is clear and without ambiguity and all needed data is available, the system begins to live its own life.

4. THE DEFINITIONS OF THE BENCMARK CASES

In order for the stored computation results to fulfill their purpose as a reference for validation, it is critical to ensure all participants have solved the same problem. While the methods and tools used may vary, the system under optimization along with the modeled physical phenomena must be the same. Thus the definitions of the benchmark cases need to be rigorous, contain all relevant information required to re-compute the case and leave no ambiguity for the interpretation. Our database contains an on-line template based on the problem definitions by Désidéri et al.(1991) for the definitions. The published benchmark case definitions are generated directly from the information entered in the template. Each benchmark case description has the following structure:

- Introduction
    - The introduction describes the main difficulties and challenges of the benchmark case along with a short description of the application area. The introduction should justify the importance and usefulness of computing the benchmark.
- Objectives
    - This section describes the goal of the optimization in context with the application area.
- Requirements
    - Requirements of the benchmark list the types of tools required to successfully compute the case such as FEA software, meshers and optimizers.
- Computational domain
    - This section describes the computational domain of the benchmark case along with the geometry of the object(s) under optimization. Usage of illustrations along with parameters and measurements is recommended.
- Modelling: physical properties
    - Each benchmark case needs to provide a description of the exact physical conditions and properties for the simulations.
- Boundary and/or initial conditions for computations
    - This section defines the boundary conditions and the initial state of the system to be simulated.
- Material parameters
    - This describes whether the benchmark is dealing with solid and/or fluid materials.
- Optimization
    - The quantity to either minimize or maximize.
- Design parameters
    - This section describes the parameters of the geometry that can be altered in order to alter the properties of the object.
- Objective function definition
    - Objective function is the mathematical representation of the fitness of the object under optimization.
- Results
    - This section describes the format and content of the results, such as quantities and plots of interest, requested to be updated in the database after successful computation of the case.

After the submission of a new benchmark case on-line, the definition is evaluated for correctness and completeness by our team and improvements are requested if deemed necessary. When the definition is considered finished it is published on-line in the database. The creator of the benchmark case is named as the chairman of the case. The benchmark case chairmen have the opportunity to participate in database workshop events to present the benchmark case and chair the benchmark case session. It is also preferred that the chairmen solve the case themselves and store the initial example solution to the database.

So far the database contains 12 benchmark case definitions of which two have been received from the industry and rest from academic research units. The topics of the cases vary from academic type optimization problems to industrial level problems that represent situations encountered in actual product design.

5. AVAILABLE BENCHMARK CASES

This section summarizes briefly the benchmark cases provided by chairman contributors of the database. Complete and detailed descriptions are available on-line at the Design Test Case Database.

5.1. Academic Benchmark Cases

5.1.1. A Numerical Set-up For Benchmarking And Optimization Of Fluid-Structure Interaction

The main purpose of this benchmark is to describe specific configurations which shall help in future to test and to compare different numerical methods and code implementations for the fluid-structure interaction (FSI) problem which can be additionally coupled with an

additional optimization procedure. This FSI benchmark is based on an older successful 'flow around cylinder' benchmark for incompressible laminar fluid flow (Schäfer and Turek, 1996). Similar to this older configuration we consider the flow to be incompressible and in the laminar regime. The structure is allowed to be compressible, and the deformations of the structure should be significant. The overall setup of the interaction problem is such that the solid object with elastic part is submerged in a channel flow (Figure 1).

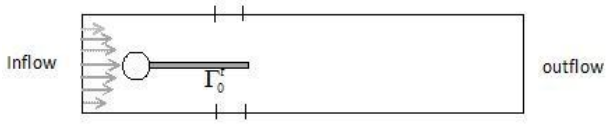


Figure 1: Illustration of the computational domain (Turek, 2009)

5.1.2. Inverse or Optimization Problems for Multiple (Ellipse) Ellipsoid Configurations

This academic test case was developed in order to study algorithmic convergence by splitting the inverse problem (recovery of target pressure on the surface) into smaller sub-problems. It also provides a way to study the behaviour of algorithms with meshes of different quality. Finally, it can be expanded into a simple test platform for multiphysics optimization (computational fluid dynamics, computational electromagnetism, and aeroacoustics), both in 2D and 3D (Leskinen, 2009). This benchmark has been successfully used by Leskinen and Hecht (2011), and Leskinen and Périaux (2011). This benchmark includes three different reconstruction problems where the goal is to recover the original positions of two ellipses or ellipsoids (Figure 2) under varying conditions.

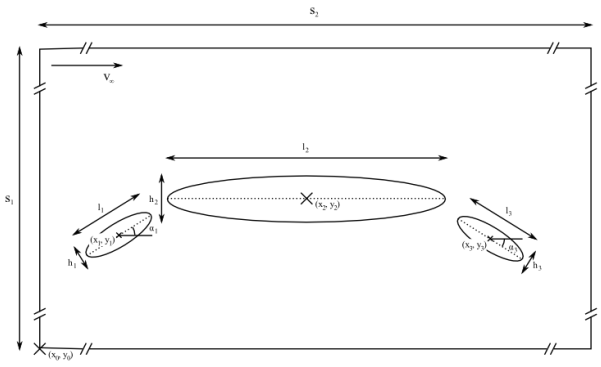


Figure 2: Illustration of the computational domain (Leskinen, 2009)

5.1.3. Optimization of Beam Profile in Fluid-Structure Interaction

The test case combines fluid-structure interaction with optimization in a simple but effective way. The cost function is well defined, has a definite global minimum and its evaluation requires the solution of a strongly coupled fluid-structure interaction problem. The individual problems are easily solved while the coupled problem sets requirements to the efficient coupling of the different sub-problems. The aim is to optimize the geometry of an elastic beam so that it bends as little as possible under the pressure and traction forces resulting from viscous incompressible flow. The profile of the beam has an effect both on the flow and the structural stiffness of the beam, respectively. (Råback, 2009b) (Figure 3)

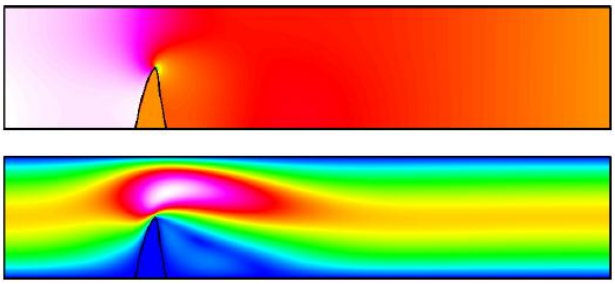


Figure 3: Pressure and velocity fields in an example solution (Råback, 2009a)

5.1.4. Shock Control Bump Optimization on a Transonic Laminar Flow Airfoil

Shock control bumps were found to be effective in reducing the wave drag and the total drag if installed on transonic airfoils or wings. However, their effectiveness relies on the position, height, and size of the bumps. This benchmark case looks into the optimal design parameters for a given laminar flow airfoil, i.e. RAE5243 airfoil (Figure 4), at the design Mach number and Reynolds number. It is divided into two cases: (1) fully turbulent flow; (2) fixed transition at 45%c. The optimization is constrained by a given lift condition. (Qin, 2009)

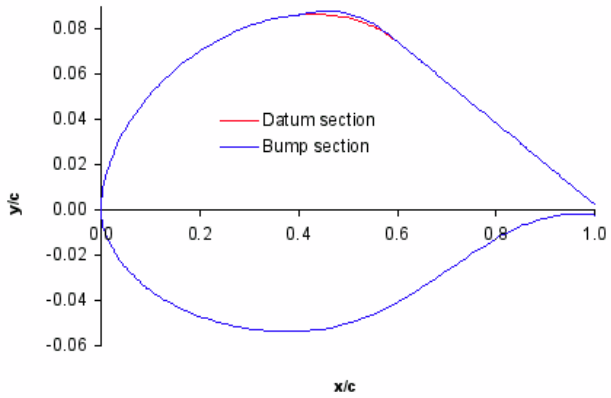


Figure 4: RAE5243 airfoil with a shock control bump (Qin, 2009)

5.1.5. 3D Shock Control Bump Optimisation

This test case extends the optimization of a shock control bump on a RAE5243. The computations are conducted under different flight conditions and in a three-dimensional domain (Figure 5).

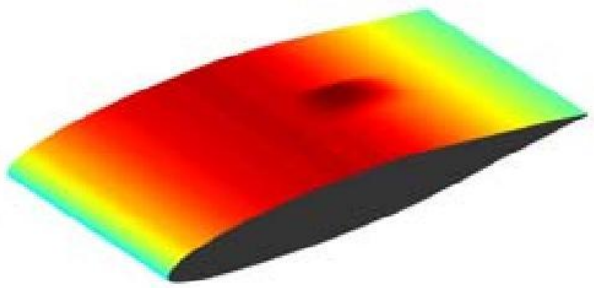


Figure 5: RAE5243 wing with 3D shock control bump (McIntosh & Qin, 2010)

5.1.6. Maximizing the Performance of SHM Systems by Robust Sensor Network Optimization

Recent advanced design tools and material offers complex structures with composite materials. However, the impact on structures causes delamination between composite layers or crack on fiber-reinforced area which the current visual inspection is impossible to check. In addition, current visual inspection will take high time cost on large structures in engineering. This is why Structural Health Monitoring (SHM) system is introduced as a promising technology to maintain healthy structure in increasing engineering applications.

The main goal of this test case is to maximize the Probability of Detection (POD) by selecting an optimal number of sensors and also their locations with efficient optimization methods like Evolutionary Algorithms which will be part of a SHM system to handle complex models. (Chang *et al.*, 2010)

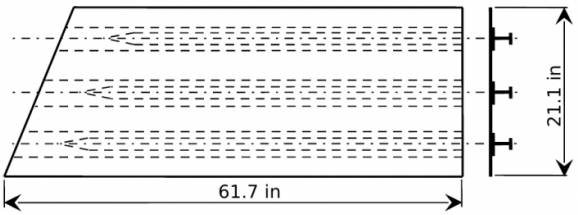


Figure 6: Test model used in the benchmark case (Chang *et al.*, 2010)

5.1.7. Reconstruction of BINACA0012 Geometry Using Discrete and Continuous Optimization

This benchmark case presents an inverse problem consisting of recovery of positions of two BINAC0012 airfoils (Figure 7) in either discrete or continuous search space.

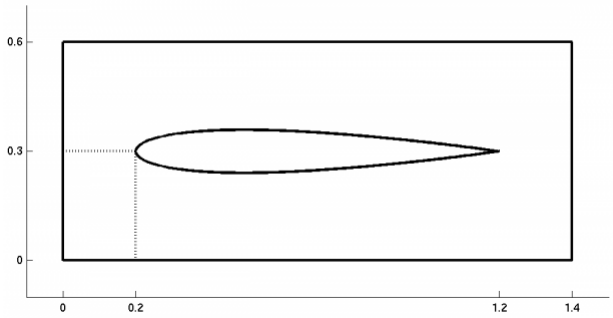


Figure 7: BINAC0012 airfoil in its bounding box (Leskinen and Wang, 2010)

5.2. Industrial Benchmark Cases

5.2.1. MDO of Mobile Phone: Antenna, SAR, HAC and Temperature

This benchmark case draws from common design challenges in mobile phone industry, especially in antenna design (Figure 8). The problem is divided in three different levels with increasing difficulty.

The first option is to optimize the antenna geometry according to the defined objective function. A reference model is provided along with the benchmark case definition that can be used as a reference. The second option is to combine antenna performance and temperature on keyboard and display area. The third, and most challenging, is to optimize the design according the all objectives given: antenna efficiency, temperature, specific absorption rate and hearing aid compatibility. (Jekkonen, 2009)

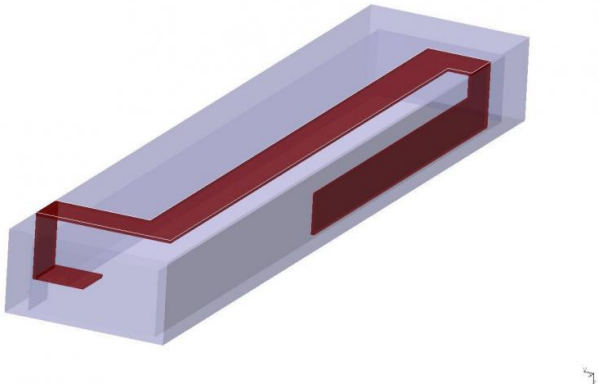


Figure 8: Example geometry for an antenna radiator

5.2.2. Optimization of a Generic Air Control Surface

This benchmark case involves the minimization of the mass of a generic air control surface by shape optimization of the internal spar structure of the air control surface. Figure 9 shows the structure under optimization.

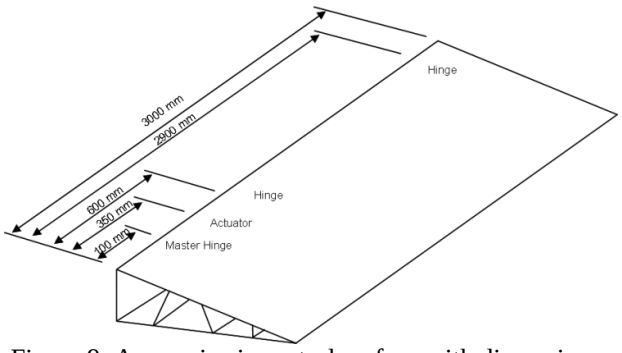


Figure 9: A generic air control surface with dimensions (Hepola, 2009)

5.2.3. Numerical Investigation of 3D Flow Over Horizontal Axis Wind Turbine NREL Phase VI

The goal in this benchmark case is to maximize the generated power of a horizontal axis wind turbine with constant or slight increase of thrust. The wind turbine under examination is a NREL Phase VI, a two-bladed 10.1-meter diameter upwind wind turbine. It is a stall regulated wind turbine, with twisted and tapered blades whose sectional geometry is the S809 airfoil. (Hirsch, 2010) Figure 10 shows the airfoil along with its loading components.

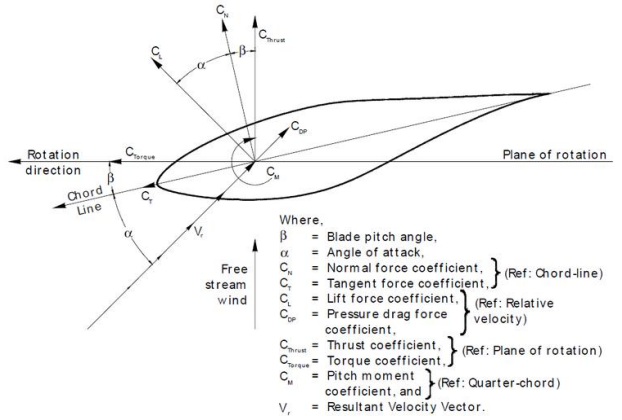


Figure 10: S809 airfoil with the definition of loading components (Hirsch, 2010)

5.2.4. Optimal Flow Divider

The first component in the headbox of a paper machine is a flow divider (Figure 11), which is to be designed to give an equal flow rate over the width of a paper machine. In this benchmark case the goal is to optimize the piece-wise linear back-wall of the flow divider in such a way that the outflow is as even as possible. The back wall is parameterized by equally distributed 5 design variables with one meter spacing. (Hämäläinen, 2010)

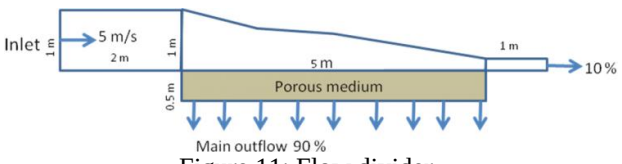


Figure 11: Flow divider

6. TECHNICAL IMPLEMENTATION OF THE DATABASE

In contrast to for example Ingenet database where the content was distributed as a static collection of HTML documents and also on a physical CD medium we want to provide the scientific community with a system that grows in time. We need to be able to add new benchmark case descriptions on the fly and also provide the contributors with a simple method of delivering their results in a uniform manner. We expect the amount of benchmark cases and contributed results to grow significantly as the database gains publicity. To prepare for growth a scalable software tool for efficient publishing and management of the benchmark case descriptions and computation results is needed. We also want to keep the presentation level and the user interface (i.e. the pages the users sees when browsing to the database web site) separate from the actual data. A Web Content Management System (CMS) called Drupal (Drupal.org, 2011) was chosen for this task.

Drupal is a popular open source CMS / web application framework written in PHP. Because it runs on a typical LAMP software stack (Linux, Apache, MySQL, PHP, see Figure 12) our entire database is based completely on open source technologies. HTTP services are provided by the Apache web server running on Linux operating system. Apache uses a PHP interpreter to run Drupal that uses the MySQL database for storing most of the content. Large data files are stored directly on the server file system. Drupal is highly extensible and besides managing, creating and publishing provides all the functionality required by our system: user management, role based and granular access management and dynamic user interfaces. Due to its extensibility and large amount of plug-ins available, many features such as TeX formatted mathematic formulas, file uploads could be implemented without writing custom code.

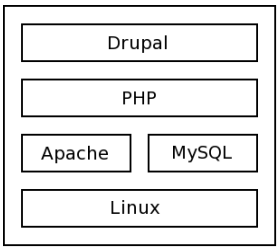


Figure 12: Software stack of the Design Test Case Database

In order to make the database as useful as possible we want to give users the opportunity of examining the actual post-processing data of the example solutions instead of static plots. For this reason we request all example solutions to use a common format for storing meshes and post-processing data. The VTK file format, supported by the Visualization Toolkit libraries and for example ParaView visualization software was chosen as the common format. The reason for this was the fact that it is well documented (Kitware, 2010), supports a

wide array of different types of data, is based on XML and thus human readable and relatively easy to convert to from other formats. By using the open source visualization tool ParaView, all users of our database are able to examine and compare the contributed data in high detail instead of resorting to static low resolution pre-prepared plots (Figure 13). ParaView was also successfully used in our Database Workshop events during presentation of the results.

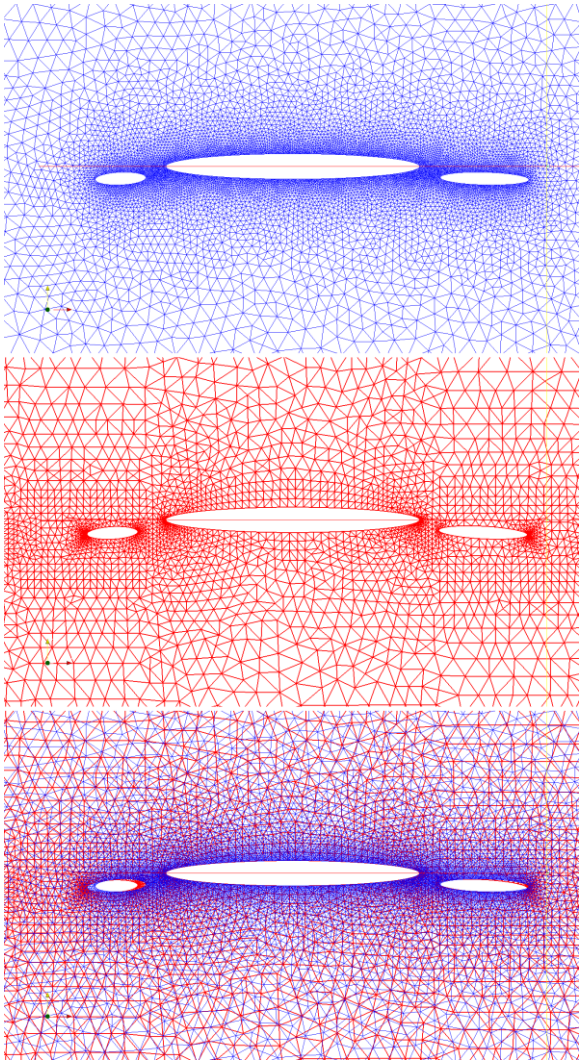


Figure 13: Example of a comparison of two contributed solutions using ParaView

7. CONCLUSIONS

Solving new increasingly complex problems requires development of new methods and tools but verification of their correctness and efficiency in absence of actual experimental data is difficult. In this paper we presented an open database of benchmark cases for multidisciplinary optimization validation that can serve as a reference point for discovery and validation of optimization methods and facilitate adoption of such methods in the industry.

By making this type of benchmark cases available to all interested parties, the long term goal is to improve the quality of future multidisciplinary optimization studies. One goal of this paper was to share the experiences obtained during the process of building the database and to make the environment better known. We have shown how to contribute to the system and described several benchmark cases, defined by chairmen, found in the database. Also we have introduced one way to implement a scalable on-line database for this kind of purpose. The definition and prescribed format for solution comparison have been noted.

ACKNOWLEDGMENTS

We would like to acknowledge professors Jacques Periaux and Pekka Neittaanmäki for discussions and ideas concerning the development of the Design Test Case Database and organizing the Database Workshop events.

This work has been made possible by MASI and DTP programs funded by the Technology Development center of Finland (TEKES) and FiDiPro program funded by the Academy of Finland and TEKES.

Special thanks to all the people who have contributed to the database, chairmen for creating the benchmark case definitions and workshop participants for example solutions.

REFERENCES

Chang, F.K., C. Lee, & J. Periaux, 2010. TA7: Maximizing the Performance of SHM Systems by Robust Sensor Network Optimization. Available from Design Test Case Database: http://jucri.jyu.fi/?q=testcase/35

Désidéri, J.A., R. Glowinski and J. Périaux, editors, 1991. Hypersonic flows for reentry problems, vol. I: Survey lectures and test cases for analysis. Springer: Berlin.

Design Test Case Database, 2009. Design Test Case Database. Available from http://jucri.jyu.fi/

Drupal.org, 2011. Drupal.org - Open Source CMS. Available from http://drupal.org/

Hepola, P., 2009. TI2: Patria AST Test Case. Available from Design Test Case Database: http://jucri.jyu.fi/?q=testcase/6

Hirsch, C., 2010. TI4: Numerical investigation of 3D flow over Horizontal Axis Wind Turbine NREL Phase VI. Available from Design Test Case Database: http://jucri.jyu.fi/?q=testcase/48

Hämäläinen, J., 2010. TI5: Optimal flow divider. Available from Design Test Case Database: http://jucri.jyu.fi/?q=testcase/50

INGENET, 2008. Evolutionary Computing for Industrial Design: The Ingenet Experience. Available from European INGEnet Network Project: http://ceani.ulpgc.es/ingenetcd/database/database.htm

Jekkonen, J., 2009. TI1: MDO of Mobile Phone: Antenna, SAR, HAC and Temperature. Available from Design Test Case Database: http://jucri.jyu.fi/?q=testcase/6

Kitware, 2010. VTK User's Guide Version 11. Kitware Inc.

Leskinen, J., 2009. TA2: Inverse or optimization problems for multiple (ellipse) ellipsoid configurations. Available from Design Test Case Database: http://jucri.jyu.fi/?q=testcase/5

Leskinen, J. and F. Hecht, 2011. Nash Games and Adaptive Meshing in a Steady-State Navier-Stokes Shape Reconstruction Problem. Evolutionary Methods for Design, Optimization and Control, Proceedings of the EUROGEN09 conference, June 15-17. Cracow, Poland.

Leskinen, J. and J. Périaux, 2011. Increasing Paralellism of Evolutionary Algorithms by Nash Games in Design Inverse Flow Problems. Evolutionary Methods for Design, Optimization and Control, CIRA. Proceedings of the EUROGEN11 conference, September 14-16, 2011. Capua, Italy.

Leskinen, J. and H. Wang, 2010. Reconstruction of BINACA0012 geometry using discrete and continuous optimization. Available from Design Test Case Database: http://jucri.jyu.fi/?q=testcase/49

Marini, M., R. Paoli, F. Grasso, J. Periaux, and J.A. Desideri, 2002. Verification and Validation in Computational Fluid Dynamics: the FLOWNET Database Experience. JSME International Journal, Series B , 45 (1).

McIntosh S., and N. Qin, 2010. TA6: 3D Shock Control Bump Optimisation. Available from Design Test Case Database: http://jucri.jyu.fi/?q=testcase/34

Qin, N., 2009. TA5: Shock control bump optimization on a transonic laminar flow airfoil. Available from Design Test Case Database: http://jucri.jyu.fi/?q=testcase/4

Råback, P., 2009a. Optimization of beam profile with Elmer . Available from Design Test Case Database: http://jucri.jyu.fi/?q=node/27

Råback, P., 2009b. TA4: Optimization of beam profile in fluid-structure interaction. Available from Design Test Case Database: http://jucri.jyu.fi/?q=testcase/8

Schäfer, M. and S. Turek, 1996. Benchmark computations of laminar flow around cylinder. Flow Simulation with High-Performance Computers II, volume 52 of Notes on Numerical Fluid Mechanics .

Turek, S., 2009. TA1: A numerical set-up for benchmarking and optimization of fluid-structure interaction . Available from Design Test Case Database: http://jucri.jyu.fi/?q=testcase/14

Turek, S. and J. Hron, 2006. Proposal for numerical benchmarking of fluid-structure interaction between an elastic object and laminar incompressible flow. Lecture Notes in Computational Science and Engineering , Volume 53, 371-385.

REMARKS ON QUBIT NEURON-BASED QUANTUM NEURAL SERVO CONTROLLER

Kazuhiko Takahashi

Doshisha University, 1-3 Miyakodani Tatara Kyotanabe Kyoto, Japan
katakaha@mail.doshisha.ac.jp

ABSTRACT

This paper presents a servo-level controller using a quantum neural network and investigates its characteristics for control systems. A multi-layer quantum neural network that uses qubit neurons as an information processing unit is used to design three types of neural-network-based servo controllers: a direct controller, parallel controller and self-tuning controller. Computational experiments to control a nonlinear discrete time plant are conducted in order to evaluate the learning performance and capability of the quantum neural-network-based servo controllers. The results of the computational experiments confirm both the feasibility and effectiveness of the quantum neural-network-based servo controllers.

Keywords: Quantum neural network, Qubit neuron, Servo controller, Real-coded genetic algorithm

1. INTRODUCTION

After Feynman (1982) introduced the possibility of using quantum mechanical systems for reasonable computing, Deutsch (1989) proposed the first quantum computing model. Subsequently, several quantum computing algorithms (Shor 1994, Grover 1996) have been proposed. Since Kak (1995) originally presented the concept of quantum neural computing, interest in artificial neural networks based on quantum theoretical concepts and techniques (hereafter called quantum neural networks) increased because of the belief that quantum neural networks may provide a new understanding of certain brain functions and also help solving classically intractable problems (Ezhov and Ventura 2000). In quantum computing, 'qubits' (an abbreviation for quantum bits) are the counterparts of the 'bits' of classical computers, and they are used to store the states of circuits in quantum computations. The quantum neural network that utilizes qubit neurons as an information processing unit was proposed by Matsui, Takai, and Nishimura (2000), where the qubit neuron model is the one in which neuron states are connected to quantum states, and the transitions between neuron states are based on operations derived from quantum logic gates. The high learning capability of the quantum neural network with qubit neurons was demonstrated in several basic benchmark tests and applications (Mitrpanont and Srisuphab 2002; Kouda, Matsui, Nishimura, and Peper 2005; Araujo, Oliveira, and Soares 2010; Takahashi, Kurokawa, and Hashimoto 2011). During the past quarter of the century, many studies about the application of both flexibility and learning capability of neural networks to control systems have been conducted worldwide (Narendra and Parthictsarathy 1990). Although many types of neural network-based control systems have been proposed (Hagan, Demuth, and De Jesus 2002, Balderud and Giovanini 2008), the possibility of applying quantum neural networks to servo-level controller applications has not been adequately investigated.

This paper proposes a servo controller design using a quantum neural network and investigates its characteristics for control systems. In Section 2, a multi-layer quantum neural network with qubit neurons is described and designs of three types of quantum neural-network-based servo controllers are presented (hereafter called

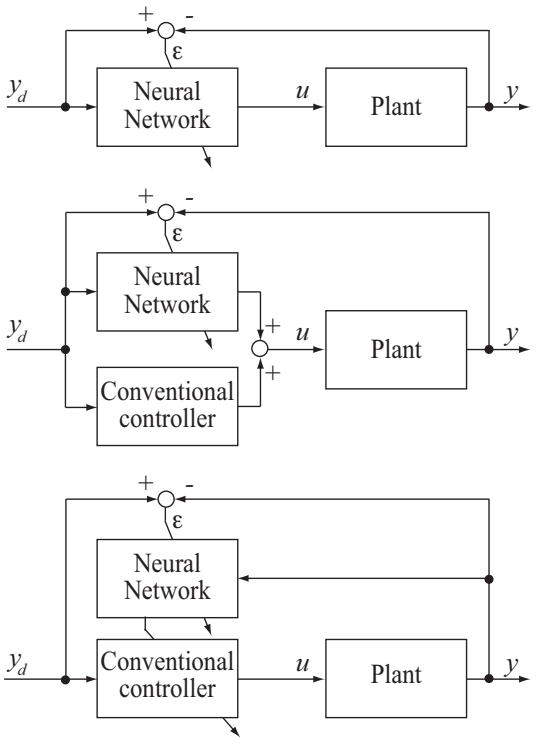


Figure 1: Schematic figure of a neural-network-based servo controller, where y_d is the desired output, y is the plant output, u is the plant input and ϵ is the output error between the desired output and the plant output (top: direct controller; middle: parallel controller; bottom: self-tuning controller.)

quantum neural servo controllers.) In this study, a direct controller, parallel controller and self-tuning controller, as shown in Fig. 1 are presented. In Section 3, computational experiments for controlling a discrete time nonlinear plant are conducted to evaluate the feasibility of the quantum neural network for servo-level controller applications.

2. DESIGN OF A QUANTUM NEURAL SERVO CONTROLLER

2.1 Neural Servo Controller

In this section, the basic idea of designing neural servo controllers is described. To design neural servo controllers, the following single-input single-output discrete time plant is considered as a controlling target plant:

$$y(k+d) = F_p[y(k), \cdots, y(k-n+1), u(k), \cdots, u(k-m-d+1)], \quad (1)$$

where y is the plant output, u is the plant input, n and m are the plant orders, k is the sampling number, d is the dead time of the plant and $F_p(\cdot)$ is the function that expresses plant dynamics. This design uses the following assumptions: the upper limit orders of the plant are known and the dead time of the plant are known. The plant output $y(k)$ depends on the past plant input and output. The plant orders determine the period in which the plant output depends on them. This period is usually shorter than the trial period. By considering the desired plant output $y_d(k)$, the output error $\epsilon(k)$ can be defined: $\epsilon(k) = y_d(k) - y(k)$.

Direct controller: In the direct controller, the output from the neural network u_n is input to the plant directly as shown at the top of Fig. 1: $u(k) = u_n(k)$. Substituting Eq. (1) into the output error and then setting it to zero, an input vector $\boldsymbol{x}_D(k)$ of the neural network can be defined as follows:

$$\boldsymbol{x}_D(k) = [\ y_d(k+d) \quad y(k) \quad \cdots \quad y(k-n+1) \quad u(k-1) \quad \cdots \quad u(k-m-d+1)\ ]^{\mathrm{T}}. \quad (2)$$

Parallel controller: In the parallel controller, the plant input is composed of the outputs obtained from the neural network u_n and a conventional controller u_c as shown in the middle of Fig. 1: $u(k) = u_n(k) + u_c(k)$. Substituting Eq. (1) into the output error and then setting it to zero, an input vector $\boldsymbol{x}_P(k)$ of the neural network can be defined as follows:

$$\boldsymbol{x}_P(k) = [\ y_d(k+d) \quad y(k) \quad \cdots \quad y(k-n+1) \quad u_n(k-1) \quad \cdots \quad u_n(k-m-d+1) \quad u_c(k) \quad \cdots \quad u_c(k-m-d+1)\ ]^{\mathrm{T}}. \quad (3)$$

Self-tuning controller: In the self-tuning controller, the plant input is calculated by a conventional feedback and/or feedforward controller and its parameters are adjusted by the neural network as shown at the bottom of Fig. 1. When a digital PID control law is utilized as the conventional controller, the plant input is defined as follows:

$$\begin{aligned} u(k) = u(k-1) &+ K_P(k)\{\epsilon(k) - \epsilon(k-1)\} \\ &+ K_I(k)\epsilon(k) \\ &+ K_D(k)\{y(k) - 2y(k-1) + y(k-2)\}, \end{aligned} \quad (4)$$

where K_P(k), $K_I(k)$ and $K_D(k)$ are the proportional gain, integral gain and differential gain, respectively. In this controller, the gain parameters are given by the neural network as follows: $K_P(k) = u_{n_1}(k)$, $K_I(k) = u_{n_2}(k)$ and $K_D(k) = u_{n_3}(k)$. Substituting Eq. (1) into the output error and then setting it to zero, an input vector $\boldsymbol{x}_S(k)$ of the neural network can be defined as follows:

$$\boldsymbol{x}_S(k) = [\ y_d(k+d) \quad y(k) \quad y(k-1) \quad y(k-2) \quad \cdots \quad y(k-n+1) \quad u(k-1) \quad \cdots \quad u(k-m-d+1) \quad \epsilon(k) \quad \epsilon(k-1)\ ]^{\mathrm{T}}. \quad (5)$$

2.2 Multi-layer Quantum Neural Network with Qubit Neurons

In quantum computing, the two quantum states express one bit of information: $|0\rangle \in \mathbf{C}$ corresponds to the classical computer's bit 0 and $|1\rangle \in \mathbf{C}$ corresponds to bit 1. Here the symbol $|\ \rangle$ is a part of the Dirac notation. The qubit state $|\psi\rangle$ maintains a coherent superposition of states: $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$, where a and b are complex numbers called probability amplitudes that satisfy $|a|^2 + |b|^2 = 1$. The operations of the rotation gate and controlled NOT gate in quantum computations can be expressed by the quantum state rewritten using the phase ϕ. The rotation gate, which is a phase-shift gate that transforms the phase of quantum states, can be expressed as $f(\phi_1 + \phi_2) = f(\phi_1)f(\phi_2)$, where $f(\phi) = \mathrm{e}^{\mathrm{i}\phi}$ (i is an imaginary unit). The controlled NOT gate, which is the phase reverse operation defined with respect to the controlled input parameter γ, can be given by $f(\frac{\pi}{2}\gamma - \phi)$, where $\gamma = 1$ and $\gamma = 0$ correspond to reversal rotation and non-rotation, respectively. Although the phase of the probability amplitude of the quantum state $|1\rangle$ is reversed when $\gamma = 0$, this case can be treated as non-rotation because its observed probability is invariant. By considering these gates, the state of the j-th qubit neuron model in the r-th set z_j^r is defined as follows:

$$\begin{cases} z_j^r = f[\dfrac{\pi}{2}\delta_j^r - \arg(v_j^r)] \\ v_j^r = \displaystyle\sum_l f(\theta_{l,j}^r) f(z_l^{r-1}) - f(\lambda_j^r) \end{cases}. \quad (6)$$

Here, δ_j^r is the reversal parameter corresponding to the controlled NOT gate, $\theta_{l,j}^r$ is the phase parameter corresponding to the phase of the rotation gate and λ_j^r is the threshold parameter having a range $[0, 1]$.

The multi-layer quantum neural network is designed by combining qubit neurons in layers. In the input layer (indicated by superscript r of I), the network input x_l in the range $[0, 1]$ is first converted into quantum states with a phase in the range $[0, \frac{\pi}{2}]$, and then the output, given

by $z_l^I = f(\frac{\pi}{2}x_l)$, is fed into the neurons present in the hidden layer (indicated by superscript r of H). In the hidden and output layers, the outputs from neurons are given by Eq. (6). By considering the probability of the state in which $|1\rangle$ is observed from the j-th neuron in the output layer (indicated by superscript r of O), the output from the network u_{QN_j} is defined as follows:

$$u_{QN_j}(\omega, \boldsymbol{x}) = |\mathrm{Im}(z_j^O)|^2. \tag{7}$$

Here, the vector ω is composed of parameters $\theta_{l,j}^r$, δ_j^r and λ_j^r ($r = I, H, O$), and the vector $\boldsymbol{x}$ is composed of input x_l. In practical applications, the outputs from the quantum neural network are converted from the range $[0, 1]$ into the range $[u_{n_{min}}, u_{n_{max}}]$ with a gain and shift factors: $u_{n_j} = c_0(u_{QN_j} - u_{QN_0})$ where c_0 is the gain factor, and u_{QN_0} is the shift factor.

2.3 Training by Real-coded Genetic Algorithm

The training of the quantum neural servo controller is performed by searching the optimal parameters $\theta_{l,j}^r$, δ_j^r and λ_j^r of the quantum neural network so as to minimize the cost function, $J(\omega) = \frac{1}{2}\sum_k \epsilon^2(k)$. A back-propagation algorithm, which is based on the steepest descent method to minimize the cost function: $\omega_{p+1} = \omega_p - \eta\frac{\partial J(\omega_p)}{\partial \omega_p}$ where η is the learning factor and p is the iteration number, can be applied for training the quantum neural servo controller. However, its learning occasionally falls into a local minimum, and the information of a plant Jacobian $\frac{\partial y}{\partial u}$ is required in order to calculate the derivative of the cost function. In this study, a real-coded genetic algorithm (Akimoto, Sakuma, Ono, and Kobayashi 2009) is utilized in training the quantum neural servo controller. Thus, the parameter values of the vector $\boldsymbol{w}$ are used directly as gene parameters of an individual. The real-coded genetic algorithm is composed of a multi-parental crossover and a generation alternation model. A real-coded ensemble crossover is used as the multi-parental crossover. The real-coded ensemble crossover is a generalization of the enhanced unimodal normal distribution crossover, and has some probability distribution in the multi-parental crossover operation in order to avoid the asymmetry and bias of children distribution. In the real-coded ensemble crossover, the new individuals (children), $\boldsymbol{g}_c$, are generated using multi-parental individuals, $\boldsymbol{g}_j$ ($j = 1, 2, \cdots, N + K$: N is the dimension of the problem; in this study it is the dimension of the vector $\boldsymbol{w}$):$\boldsymbol{g}_c = \boldsymbol{g}_0 + \sum_{j=1}^{N+K}\nu_j(\boldsymbol{g}_j - \boldsymbol{g}_0)$, where $\boldsymbol{g}_0$ indicates the center of gravity of the parents and ν_j is the stochastic variable that follows the probability distribution $\varphi(0, \frac{1}{N+K})$. In a generation alternation model, the just generation gap which replaces parents with children completely in every generation, is utilized. In the just generation gap, the numbers of population, parents and children are recommended to be $(15 \sim 50)N$, $N + K$ and $10N$, respectively. To evaluate the individuals, a fitness function of the q-th individual at the p-th generation is defined by the reciprocal of the cost function $J(\omega_p^q)$.

3. COMPUTATIONAL EXPERIMENTS

The quantum neural servo controllers are numerically investigated using a discrete time nonlinear plant in the computational experiments. The equation of the plant in which the second-order system is dominant is as follows:

$$y(k+1) = F_s[-\sum_{i=1}^{3} a_i y(k-i+1) + \sum_{i=1}^{2} b_i u(k-i+1) + c_{non}y^2(k)], \tag{8}$$

where a_3 and c_{non} are the coefficients of the parasitic term and the nonlinear term and the function $F_s(\cdot)$ has the nonlinear characteristic of saturation:

$$F_s(x) = \begin{cases} 1 & (x \geq 1) \\ x & (-1 < x < 1) \\ -1 & (x \leq -1) \end{cases}$$

In the experiments, the values of plant parameters were set to $a_1 = -1.3$, $a_2 = 0.3$, $b_1 = 1$, $b_2 = 0.7$, $a_3 = 0.03$ and $c_{non} = 0.2$ (Yamada 2010).

In designing the quantum neural servo controller, the plant was assumed to be a linear second order plant: $d = 1$, $n = 2$ and $m = 1$. Thus, the number of qubit neurons in the input layer was 4, 6 and 7 for the direct controller, parallel controller and self-tuning controller, respectively. In each controller, the number of qubit neurons in the hidden layer was set to 4. In the parallel controller, P-control law was utilized as a conventional controller: $u_c = K_P\epsilon(k)$, where $K_P = 0.5$ in the computational experiment. The training conditions were as follows: the dimension of the problem N were 30, 38 and 54 for the direct controller, parallel controller and self-tuning controller, respectively, The value of K was 1, total number of generations was 2000 and probability distribution φ was uniform in the range [-1, 1]. In the training process, the desired plant output $y_d(k)$ was a rectangular wave in order to take account of frequency richness. The number of samples within one cycle of the rectangular wave was 50 and the amplitude of the wave was changed randomly in the range [-0.5, 0.5]. In the open test, the desired plant output, the rectangular wave, had an amplitude of ± 0.4 and the number of samples within one cycle varied from 50 to 60, 30 and 40 in order.

As a reference for comparing the results of the quantum neural servo controller, the following conventional multi-layer neural network that uses sigmoid functions as an information processing unit was utilized to design the neural servo controller (hereafter called sigmoid neural servo controller).

$$u_{MN_j}(\boldsymbol{w}, \boldsymbol{x}) = s[\sum_i w_{2_{ji}} s(\sum_l w_{1_{il}} x_l)], \tag{9}$$

where x_l is the input to the lth neuron in the input layer, u_{MN_j} is the output of the jth neuron in the output layer, $w_{i_{jl}}$ ($i = 1, 2$) is the weight that includes the threshold, $s(\cdot)$ is the sigmoid function, and the vectors $\boldsymbol{w}$ and

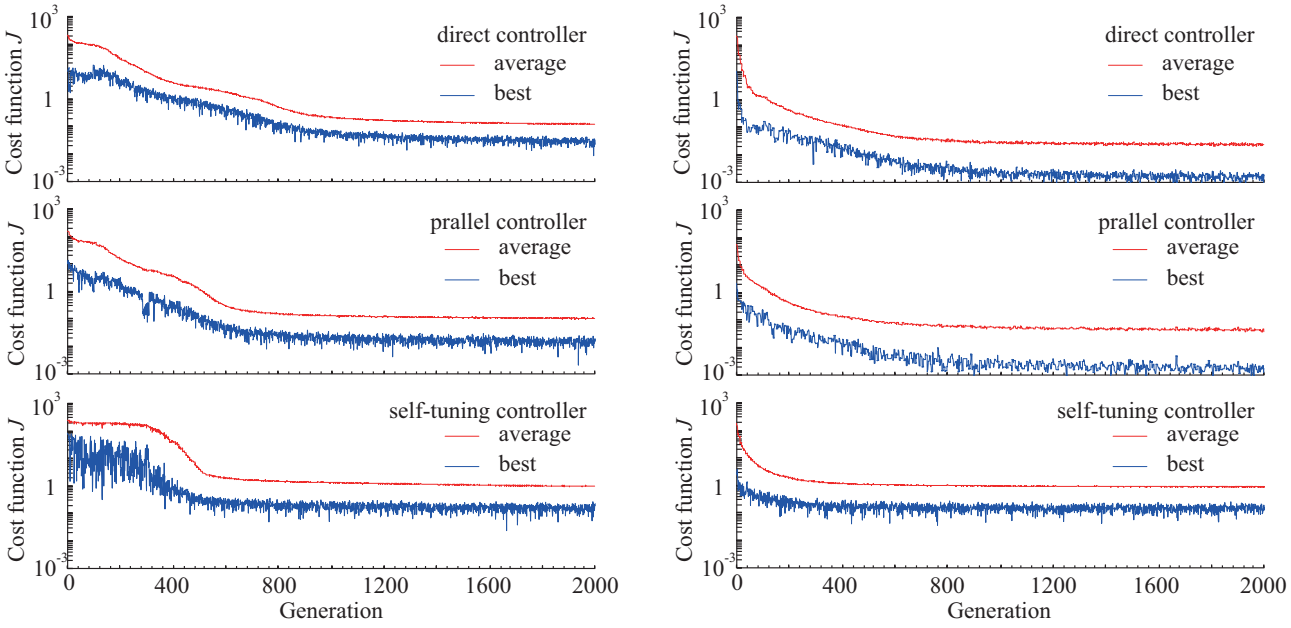


Figure 2: Example of neural servo controllers' training process (left: quantum neural servo controller; right: sigmoid neural servo controller; top: direct controller; middle: parallel controller; bottom: self-tuning controller.)

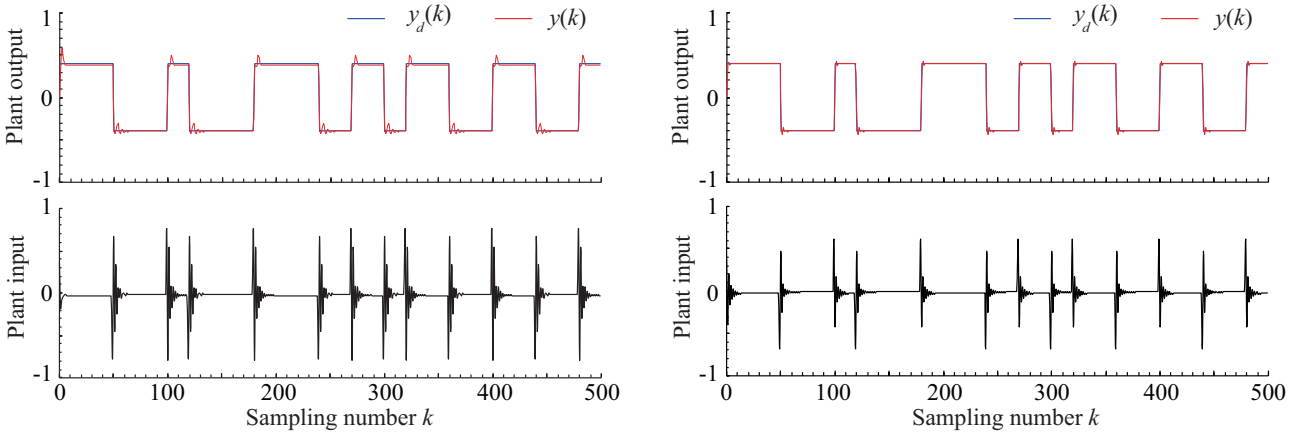


Figure 3: System response controlled by the direct neural servo controller (left: quantum neural servo controller [J = 0.224]; right: sigmoid neural servo controller [J = 0.243]; top: desired and plant outputs; bottom: plant input.)

$\boldsymbol{x}$ are composed of the weight $w_{i_{jl}}$ $(i = 1, 2)$ and input x_l respectively. The learning of the sigmoid neural servo controller is performed by the real-coded genetic algorithm in order to minimize the cost function, $J(\boldsymbol{w}) = \frac{1}{2}\sum_k \epsilon^2(k)$. The training conditions used here were the same as that in the quantum neural servo controller. The number of neurons in the hidden layer was set to 6 in each controller so that the dimension of the problem in the sigmoid neural controller was almost the same as that in the quantum neural controller. Thus N was 31, 41 and 58.

Figure 2 shows an example of the training process. Here, the horizontal axis represents the generation and the vertical axis represents the cost function. As shown in Fig. 2, the cost function decreased as the generation progressed. Although the averaged cost functions are almost same in the quantum neural servo controller and the sigmoid neural servo controller, the cost functions of the sigmoid neural servo controller with the best individual are smaller than those of the quantum neural servo controllers in the direct and parallel controllers.

Figures 3, 4 and 5 are examples of responses controlled by the direct controller, parallel controller and self-tuning controller, respectively, after the neural servo controller's training converged. In each controller, the individual that has the best fitness function was used. Because the plant output tracks the desired output, as shown in each figure, the quantum neural servo controllers can achieve the control task of making the nonlinear plant follow the desired output. By comparing the three types

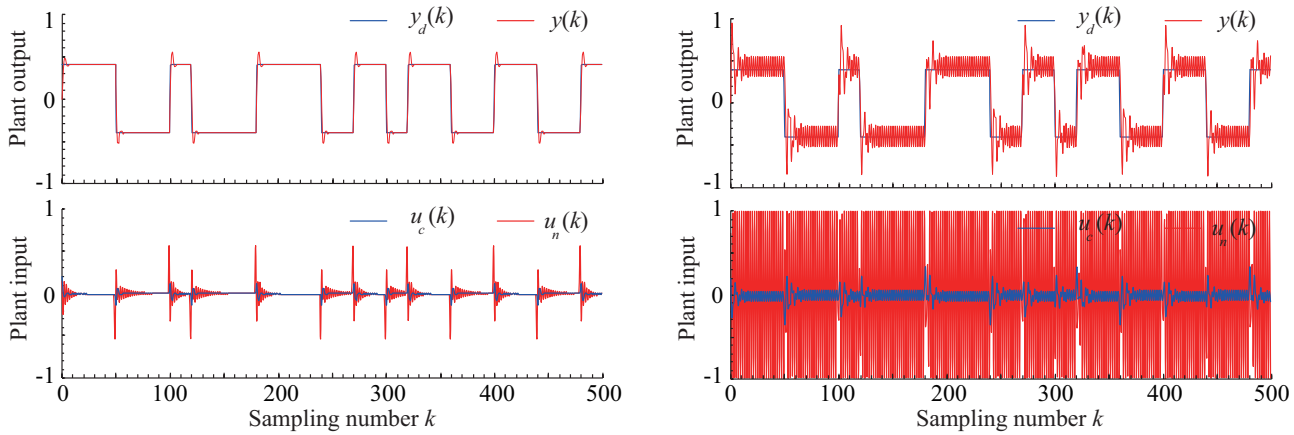


Figure 4: System response controlled by the parallel neural servo controller (left: quantum neural servo controller [J = 0.695]; right: sigmoid neural servo controller [J = 7.631]; top: desired and plant outputs; bottom: plant input.)

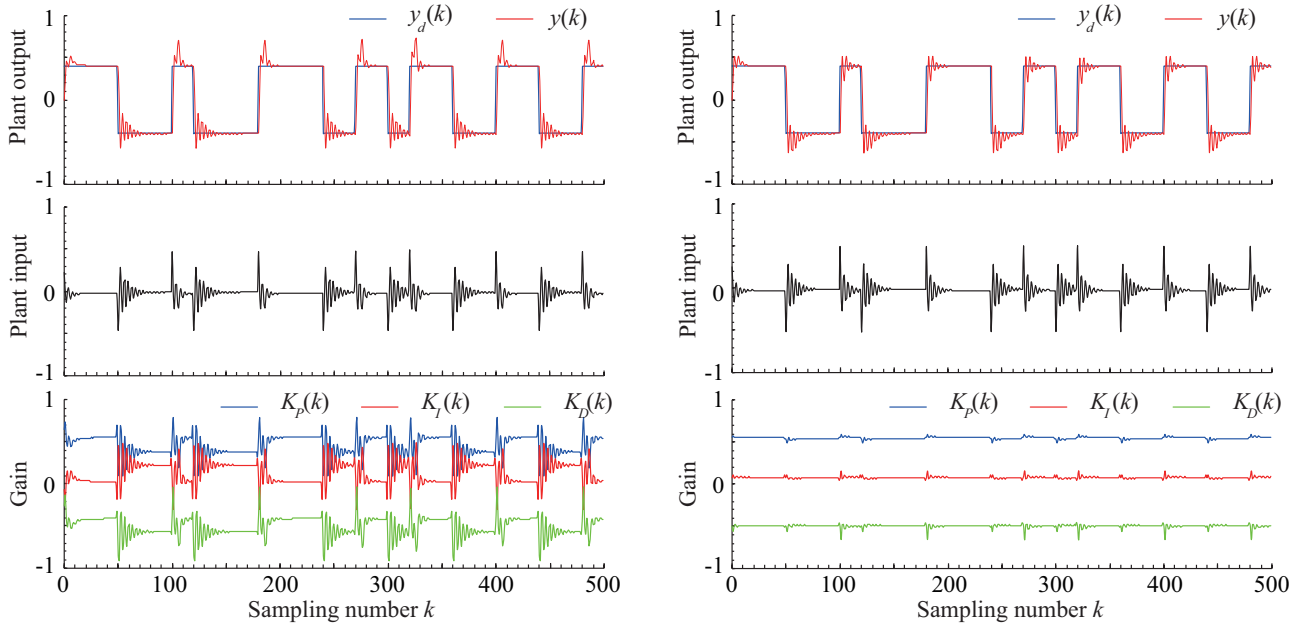


Figure 5: System response controlled by the self-tuning neural servo controller (left: quantum neural servo controller [J = 5.779]; right: sigmoid neural servo controller [J = 5.336]; top: desired and plant outputs; middle: control input; bottom: PID gain parameters.)

of quantum neural servo controllers, we observe that the direct controller has the lowest cost function in the open test. In the direct controller, the control performance of the quantum neural servo controller is almost the same as that of the sigmoid neural servo controller. In the parallel controller, shown in Fig. 4, the output from the quantum neural network is dominant in the plant input (the output from the conventional controller is almost zero). In the self-tuning controller, although the gain parameters are tuned by the output from the quantum neural network, as shown in the bottom of Fig. 5, the output error increases rather than the direct and parallel controllers because the control performance depends on the PID law. These results indicate the feasibility of the quantum neural servo controllers and show that the real-coded genetic algorithm has an advantage in the learning of the quantum neural servo controller because it does not require the Jacobian information of the plant in the training process.

4. CONCLUSION

This paper presented a servo controller designs using a quantum neural network and investigated their characteristics for control systems. The multi-layer quantum neural network that uses qubit neurons as an information processing unit was used to design three types of neural-network-based servo controller: the direct controller, parallel controller and self-tuning controller. Computational

experiments to control the nonlinear discrete time plant were conducted in order to evaluate the learning performance and capability of the quantum neural servo controllers. The results of the computational experiments confirmed both the feasibility and effectiveness of the quantum neural servo controllers.

REFERENCES

Akimoto, Y., Sakuma, J., Ono, I. and Kobayashi, S., 2009. Adaptation of Expansion Rate for Real-coded Crossovers. *Proceedings of 11th Annual Conference on Genetic and Evolutionary Computation*, 739–746.

Araujo, R. A., Oliveira, A. L. I. and Soares, S. C. B., 2010. A Quantum-Inspired Hybrid Methodology for Financial Time Series Prediction. *Proceedings of the 2010 International Joint Conference on Neural Networks*, 1–8.

Balderud, J. and Giovanini, L., 2008. Adaptive Control and Signal Processing Literature Survey. *International Journal of Adaptive Control and Signal Processing*, 22(3), 318–321.

Deutsch, D., 1989. Quantum Computational Networks. *Proceedings of the Royal Society of London*, Series A, 425, 73–90.

Ezhov, A. A. and Ventura, D., 2000. Quantum Neural Networks. *Future Directions for Intelligent Systems and Information Sciences*, 213–234.

Feynman, R., 1982. Simulating Physics with Computers. *International Journal of Theoretical Physics*, 21, 467–488.

Grover, L. K., 1996. A Fast Quantum Mechanical Algorithm for Database Search. *Proceedings of the 28th Annual ACM Symposium on the Theory of Computing*, 212–219.

Hagan, M. T., Demuth, H. B. and De Jesus, O., 2002. An Introduction to the Use of Neural Networks in Control Systems. *International Journal of Robust and Nonlinear Control*, 12(11), 959–985.

Kak, S. C., 1995. On Quantum Neural Computing. *Information Science*, 83, 143–163.

Kouda, N., Matsui, F., Nishimura, H. and Peper, F., 2005. Qubit Neural Network and Its Learning Efficiency. *Neural Computing and Application*, 14(2), 114–121.

Matsui, N., Takai, M. and Nishimura, H., 2000. A Network Model Based on Qubit-like Neuron Corresponding to Quantum Circuit. *Electronics and Communications in Japan*, 83(10), 67–73.

Mitrpanont, J. L. and Srisuphab, A., 2002. The Realization of Quantum Complex-Valued Backpropagation Neural Network in Pattern Recognition Problem. *Proceedings of the 9th International Conference on Neural Information Processing*, 1, 462–466.

Narendra, K.S. and Parthictsarathy, K., 1990. Identification and Control of Dynamics System Using Neural Networks. *IEEE Transactions on Neural Networks*, 1(1), 4–27.

Shor, P. W., 1994. Algorithms for Quantum Computation: Discrete log and factoring. *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science*, 124–134.

Takahashi, K., Kurokawa, M. and Hashimoto, M., 2011. Controller Application of a Multi-Layer Quantum Neural Network Trained by a Conjugate Gradient Algorithm. *Proceedings of the 37th Annual Conference of the IEEE Industrial Electronics Society*, 2278–2283.

Yamada, T., 2010. Transformation of Neural Network Weight Trajectories on 2D Plane for A Learning-type Neural Network Direct Controller. *Artificial Life and Robotics*, 15, 413–416.

START-UP BUSINESS SUCCESS PREDICTION BY MEANS OF ARTIFICIAL NEURAL NETWORKS

Francisco Garcia Fernandez [(a)], Ignacio Soret Los Santos [(b)], Santiago Izquierdo Izquierdo [(c)], Francisco Llamazares Redondo [(d)], Francisco José Blanco Jiménez [(e)]

[(a)] Universidad Politecnica de Madrid. Spain.
[(b)] ESIC Business & Marketing School. Spain.
[(c)] MTP Metodos y Tecnologia. Spain.
[(d)] ESIC Business & Marketing School. Spain.
[(e)] Universidad Rey Juan Carlos. Spain

[(a)]francisco.garcia@upm.es , [(b)]ignacio.soret@esic.e, [(c)]s_izdo@hotmail.com, [(d)]francisco.llamazares@esic.es, [(e)]francisco.blanco@urjc.es

ABSTRACT

There is a great interest to know if a new company will be able to survive or not. Investors use different tools to evaluate the survival capabilities of middle-aged companies but there is not any tool for start-up ones. Most of the tools are based on regression models and in quantitative variables. Nevertheless, qualitative variables which measure the company way of work and the manager skills can be considered as important as quantitative ones.

Develop a global regression model that includes quantitative and qualitative variables can be very complicated. In this study, a more powerful modeling tool is used to predict the company early years success. Artificial neural networks can be a very useful tool to model the company survival capabilities. They have been large specially used in engineering processes modeling, but also in economy and business modeling.

Keywords: Artificial neural networks, Enterprise solvency, Entrepreneurship success, SME

1. INTRODUCTION

Small and medium-sized enterprises are actually a country driving force, both in economic and employment creation. It is, therefore, important to promote the creation of new enterprises and ensure their survival, especially in the first critical three years of life. To do so, arise different policies for supporting of entrepreneurs, with the objectives of boosting entrepreneurship and support the birth of business initiatives by helping them in their beginnings.

In this aspect, from the point of view of the sponsoring body, is important to know which the degree of survival that can have the new company. Obtaining a model that attempts to explain the success or failure in a company from quantitative and qualitative variables is a task that currently involved numerous researchers and great efforts. Such a model would not only help emerging companies to predict the outcome, but would facilitate a great tool for detecting of possible weaknesses causing of a business failure. However, until now, these studies have focused more on set up companies than in new established firms.

These models, mainly regression ones, are based on quantitative variables as the Altman ratios and qualitative variables such as training manager, product quality, or training , innovation, price and quality control policies established in the company. However, to date, these models have not got an acceptable results to apply them to real life, mainly because of limitations in regression fits to model the relationships between the different variables involved in each case.

In the case of start-up businesses to predict survival studies are few and with very poor results.

Our study will use a new methodology in its application to start-up businesses. An artificial neural network will be developed for business success/failure of start-ups modeling based on both financial data and qualitative information relating to the various managerial policies. To obtain these data a series of personal surveys shall be carried out on behalf of the promoter team in each company, out of a total of 125. In order to unify the qualitative variables quantification is important to perform the surveys by the same team.

Artificial neural network are a type of mathematical structures that imitates the functioning of the brain. Their uses have been increased in the last years due to the high computers development. Artificial neural networks are capable of extracting knowledge from a series of sampling data and applying it later to unknown data. As a universal function aproximator has become a powerful tool in modeling. The major beneficiary has been the industrial scope with the production processes modeling, but also in the economic field have been a great development, ranging from modeling of Forex market to different products market value (Jalil and Misas, 2007) and the study of the enterprise solvency

prediction (Lacher et al, 1995; Jo et al, 1997; Yang et al, 1999; Hsiao et al, 2009).

The power of artificial neural networks for modeling complex relationships between variables and the success achieved in other fields, where have surpassed traditional regression models, makes us to be optimistic about its use in this study.

This study aims to predict the degree of survival of a start-up company from their first two years of activity data and detect possible competitive weaknesses.

2. MATERIALS AND METHODS

2.1. Entrepreneurship in Spain

The Spanish entrepreneurship's basic indexes through 2009 have been affected by the economic crisis. After a moderate drop (8%) in 2008, the Total Entrepreneurial Activity index (TEA) experienced a great drop (27.1%) in 2009, returning to 2004 levels (de la Vega García 2010) (Fig 1).

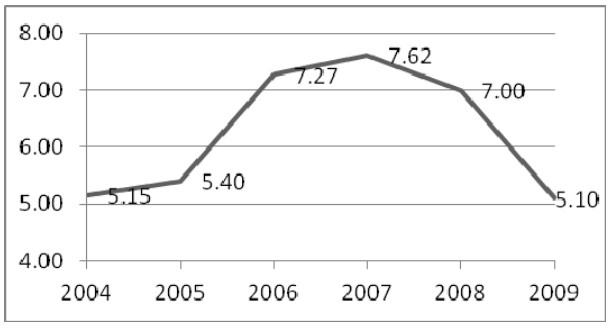


Fig.1. 2009 Executive Report. Global Entrepreneurship Monitor- Spain

According to that data, there are 1,534,964 nascent businesses (between 0 and 3 months old). The owner-managers of a new business (more than 3 months but not more than 3.5 years) have also declined in 2009, returning to 2007 levels.

As in other comparable, innovation-driven, countries, the typical early stage entrepreneur in Spain is male (62.5% of all entrepreneurs), with a mean age of 36.6, and well educated (55.4% with a university degree). The female entrepreneurial initiatives have been declined in 2009 and the difference between female and male Total Entrepreneurial Activity index (TEA) rates is now bigger than in 2008. The gender difference in the TEA index has increased from two to almost three points. Now the female TEA index is 3.33% and the male TEA index is 6.29%.

Although most individuals are pulled into entrepreneurial activity because of opportunity recognition (80.1%), others are pushed into entrepreneurship because they have no other means of making a living, or because they fear becoming unemployed in the near future. These necessity entrepreneurs are 15.8% of the entrepreneurs in Spain.

In Spain, the distribution of early-stage entrepreneurial activity and established business owner/managers by industry sector is similar to that in other innovation-driven countries, where business services (i.e., tertiary activities that target other firms as main customers, such as finance, data analysis, insurance, real estate, etc.) prevail. In Spain, they accounted for 56.5% of early-stage activities and 46.1% of established businesses. Transforming businesses (manufacturing and construction), which are typical of efficiency-driven countries, were the second largest sector, accounted for 25.9% and 24.2% respectively. Consumer services (i.e., retail, restaurants, tourism) accounted for 12.8% and 17.3%, respectively. Extraction businesses (farming, forestry, fishing, mining), which are typical of factor-driven economies, accounted for 6.0% and 8.6%, respectively. The real estate activity in Spain was of great importance, and its decline explains the reduction in the business services sector in 2009.

The median amount invested by entrepreneurs in 2009 was around 30,000 Euros (less than the median amount of 50,000 Euros in 2008). Therefore the entrepreneurial initiative is less ambitious in general.

The factors that mostly constrain entrepreneurial activity are: first, financial support (e.g., availability of debt and equity), which was cited as a constraining factor by 62% of respondents. Second, government policies supporting entrepreneurship, which was cited as a constraining factor by 40% of respondents. Third, social and cultural norms, which was cited as a constraining factor by 32% of respondents.

More than one fifth of the entrepreneurial activity (21.5%) was developed in a familiar model. Therefore, the entrepreneurial initiatives, often driven by family members, received financial support or management assistance from some family members. Nevertheless, the influence of some knowledge, technology or research result developed in the University was bigger than expected. People decided to start businesses because they used some knowledge, technology or research result developed in the University (14.3% of the nascent businesses, and 10.3% of the owner-managers of a new business).

2.2. Questionnaire

The company survival is greatly influenced by its financial capabilities, however, this numerical information is not always easy to obtain, and even when obtained, it is not always reliable.

But there are some other non numerical factors that determine company survival, such as its educational level of its employees, its customer service policies or its technical capabilities.

For this study, both numerical and qualitative data are used to model the company survival.

1.- Financial data.

Altman develop a regression model based in some financial ratios to predict the company success (Lacher et al. 1995): The most used are:

- Working Capital/Total Assets. Working Capital is defined as the difference between current

assets and current liabilities. Current assets include cash, inventories, receivables and marketable securities. Current liabilities include accounts payable, short-terms provision and accrued expenses.

- Retained Earnings/Total Assets. This ratio is specially important because bankruptcy is higher for start-ups and young companies.
- Earnings Before Interest and Taxes/Total Assets. Since a company's existence is dependent on the earning power of its assets, this ratio is appropriate in failure prediction.
- Market Capitalization/Total Debts. This ratio weighs up the dimension of a company's competitive market place value.
- Sales/Total Assets. This ratio measures the firm´s assets utilization.

2.- Qualitative data.

It is very difficult to evaluate qualitative characteristics as quality policies or technical capabilities. In this study the works of Rubio Bañón and Aragón Sánchez (2002) and Aragón Sánchez and Rubio Bañón (2005) are used. They modeled the company positioning and its survival capabilities and the influence of other factors as manager personality or the educational level of its employees:

- Manager academic level, ranged from 1 to 4.
  - PhD or Master (4).
  - University degree (3).
  - High school (2).
  - Basic studies (1).
- Company technological resources, ranged from 1 to 4.
  - The company uses self-made software programs (4).
  - The company uses specific programs but it buys them (3).
  - The company uses the same software than competitor (2).
  - The company uses older software than competitors (1).
- Quality policies, ranged from 1 to 5.
  - The company has quality policies based on ISO 9000 (5).
  - The company controls either, production and client satisfaction (4).
  - A production control is the only quality policy (2).
  - Supplies control is the only quality control in the company (1).
  - The company has not any quality policy.
- Trademark, ranged from 1 to 3.
  - The company trademark is better known than competitors' (3).
  - The company trademark is as known than competitors' (2)
  - The company trademark is less known than competitors' (3).
- Employees education policy, ranged from 1 to 2.
  - The company is involved in its employees education (2).
  - The company is not involved in its employees education (1).
- Number of innovations areas in the company, ranged from 1 to 5.
- Marketing experience, ranged from 1 to 3.
  - The company has a great marketing experience in the field of its products and in others (3).
  - The company has only marketing experience in his field of duty (2).
  - The company has no marketing experience (1).
- Knowledge of the business area, ranged from 1 to 3.
  - The manager knows perfectly the business area and has been working on several companies related whit it (3).
  - The manager knows lightly the business area (2).
  - The manager has no idea on the business area (1).
- Openness to experience, ranged from 1 to 2.
  - The manager is a practical person who is not interested in abstract ideas, prefers works that is routine and has few artistic interest (2).
  - The manager spends time reflecting on things, has an active imagination and likes to think up new ways of doing things, but may lack pragmatism (1).

Table 1. Variables

Variable	Type	Range
Working Capital/Total Assets	Quantitative	R^+
Retained Earnings/Total Assets	Quantitative	R^+
Earnings Before Interest and Taxes/Total Assets	Quantitative	R^+
Market Capitalization/Total Debts	Quantitative	R^+
Sales/Total Assets	Quantitative	R^+
Manager academic level	Qualitative	1-4
Company technological resources	Qualitative	1-4
Quality policies	Qualitative	1-5
Trademark	Qualitative	1-3
Employees education policy	Qualitative	1-2
Number of innovations areas	Qualitative	1-5

Marketing experience	Qualitative	1-3
Knowledge of the business area	Qualitative	1-3
Openness to experience	Qualitative	1-2

2.3. Artificial neural networks

ANN have been widely used in many engineering fields, especially when the relation between the variables involved in the process is not so important as find a suitable solution to our problem. From that point of view, ANNs have become a very important tool to model specially industrial processes, but also others kind of complex processes such as environmental changes, climate, tress growing, financial markets or traffic flow.

ANNs are a complex mathematical structures that try to imitate a biological brain and its way of thinking. They are able to learn from a series of examples and apply that knowledge to unknown situations.

These structures have a series of interconnected elements, (Fig 3) known as artificial neurons and those connections are the responsible of the knowledge storing.

One of the type of ANN most used is the perceptron (Fig. 3). It is made up of three layers, known as input layer, hidden layer and output layer. The input layer receives the initial values of the variables, the output layer shows the final results of the network for the input, and the hidden layer makes all the operations to get the final results.

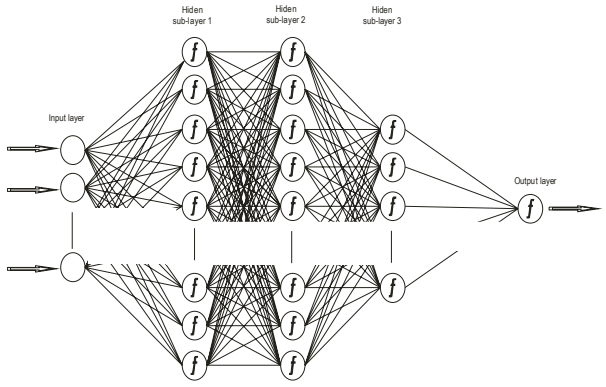


Fig. 3. Artificial neuron network architecture.

The number of neurons in the input layer is the same as independent variables, and the number of neurons in the output layer is the same as dependent variables, but, there is not a procedure to define the number of neurons and the number of layers the hidden layer should have, that means it is very difficult to choose a model, even for an experienced user. In general, the model is obtained by a trial and error process. There are some general recommendations about the final structure as it must be a pyramidal type or about the maximum number of neurons as a function of the number of examples, but they are only recommendations not really rules.

Their main advantages are they are universal non-lineal function aproximators (Hornik 1989), and perceptrons are able to filter the data noise. Data are not need to be in a normal distribution as in other statistical models. Their principal disadvantages are that they need a large amount of data for the training process, and the final model is a non interpretable model, is such a black box, where you do not know how variables are connected one to each other.

The hyperbolic tangent sigmoid function (Eq. 1) was used as the transfer function. This is equivalent to the hyperbolic tangent function and also improves network performance by producing an output more quickly (Demuth et al 2002). To improve the ANN results all the data were normalized according to Eq. 2. The transfer function produces an output in the interval (-1,+1), which means that data normalization is highly appropriate for improving network performance.

$$f(\theta) = \frac{2}{1+e^{(-2\theta)}} - 1 \qquad (1)$$

f(θ): Output value of the neuron.
θ: Input value of the neuron.

$$\theta' = \frac{\theta - \theta_{\min}}{\theta_{\max} - \theta_{\min}} \qquad (2)$$

θ': Value after normalization of vector X.
θmax y θmin: Maximum and minimum values of vector X.

The training method chosen was supervised learning. In order to this the whole initial data set has been divided into three subsets at random without repetition. The training set (60% of the data), test set (20% of the data) and validation set (20% of the data).

To avoid the problem of overfitting during the training phase, the early-stopping technique has been used. Overfitting occurs when the error in the validation set stats to increase while decreases in the training set, it is a clear indication of a generalizing loose capacity. To prevent this situation and design the ANN structure, a specific program has been develop using the Neural Network Toolbox® ver. 4.0.2, from the MATLAB® Program Ver. 6.5.0. Release 13. This program generates different perceptrons with different neurons in their inner layers, compares the training error and validation error every 100 training epochs and also compares all the preceptrons generated between them.

3. RESULTS AND CONCLUSIONS

This work is the initial steps of an ambitious project that pretend to evaluate the survival of start-up companies. Actually the work is on his second stage which is the data recompilation through different individual surveys. We hope to present in the next congress the first results to discuss them.

ACKNOWLEDGMENTS
This Researching Project is supported by ESIC Business and Marketing School, Madrid , Spain.

LIST OF REFERENCES

Aragón Sánchez A., Rubio Bañón, A., 2005.Factores explicativos del éxito competitivo: el caso de las PYMES del estado de Veracruz. *Contaduría y Administración* 216, 35-69.

de la Vega Garcia Pastor, J. 2010. GEM Informe Ejecutivo 2009 España. Madrid: Instituto de Empresa Madrid Emprende. 2009. *Memoria de Viveros de Empresa de la Comunidad de Madrid. Madrid,* Madrid Emprende.

Demuth H., Beale M., Hagan M. 2002. *Neural Network Toolbox User's guide, version 4.* Natick: The Mathworks Inc., MA 01760, USA. 808pp.

Hornik, K., 1989. Multilayer Feedforward Networks are Universal Approximators. *Neural Networks* 2, 359-366.

Hsiao, S.H. and Whang, T.J,. 2009. A study of financial insolvency prediction model for life insurers. *Expert systems with applications*., 36, 6100-6107.

Jalil M.A., Misas, M., 2007. Evaluación de pronósticos del tipo de cambio utilizando redes neuronales y funciones de pérdida asimétricas. *Revista Colombiana de Estadística*, 30, 143-161.

Jo H., Han I., Lee H., 1997. Bankruptcy prediction using case-based reasoning, neural Networks and discriminant analysis. *Expert Systems With Applications,* 13(2), 97-108.

Lacher R.C., Coats P.K., Sharma S.C., Fant, L.F., 1995. A neural network for classifying the financial health of a firm. *European Journal of Operational Research*, 85, 53-65.

Rubio Bañón A., Aragón Sánchez A., 2002. Factores explicativos del éxito competitivo. Un estudio empírico en la PYME. *Cuadernos de Gestión*, 2(1), 49-63.

Yang Z.R., Platt M.B., Platt, H.D., 1999. Probabilistic neural networks in bankruptcy prediction. *Journal of Business Research*, 44, 67-74.

A NEW DESIGN OF FMS WITH MULTIPLE OBJECTIVES USING GOAL PROGRAMMING

Berna Dengiz[(a)], Yusuf Tansel İç[(b)], Selin Coşkun[(c)], Nil Dağsalı[(d)], Damla Aksoy[(e)], Gözde Çizmeci[(f)]

[(a)] Department of Industrial Engineering, Faculty of Engineering, Baskent University, 06810, Baglica, Etimesgut, Ankara, Turkey.
[(a)]bdengiz@baskent.edu.tr, [(b)]ytansel@baskent.edu.tr

ABSTRACT
The operation of the Flexible Manufacturing System (FMS) includes complex and conflicted issues that result in the system performance. In the operational improvement studies in a FMS usually make in determination of a single-response measure. This paper presents an application for multi-response simulation optimization of a FMS via DOE (design of experiment), regression meta-model and goal programming (GP) together. A real FMS with four work stations is modeled by ARENA simulation software to optimize system performance measures considering five design and control parameters.

Keywords: Flexible Manufacturing System, Design of Experiment, Regression Meta-Model, Goal Programming.

1. INTRODUCTION

A Flexible Manufacturing System (FMS) is an automated group technology machine cell, consisting of a group of processing stations (usually CNC machine tools), interconnected by an automated material handling and storage system which is controlled by an integrated computer system (Groover, 2008). The operation and design of the FMS includes complex and conflicted factors that result in the productivity of the system (Park et al., 2001; Groover, 2008).

The operational and design factors are considered separately due to the complexity of system. However, most analytical and simulation modeling study finished so far has focused on mainly one or two decision problems among system loading, machine loading, part selection, machine grouping, tool allocation, and scheduling parts (Park et al., 2001). The following criteria have been most likely used in the FMS modeling studies: system utilization, job tardiness, due dates, production rate, work-in-process inventory, set-up time and tool changes, balance of machine usage, flow time, (Park et al., 2001; Savsar, 2005; Um et al., 2009). Most past study on the operation or design of the FMS considered only a single performance response as their objective function to optimize (Guo et al., 2003; Chan et al., 2007; Ozmutlu et al., 2004; Savsar, 2005, Kumar et al., 2011; Mahdavi et al., 2010). During last decade, however, a few researchers have used multi-objective decision-making approaches to solving FMS design problems with more than one response considering only hypothetical systems (Park et al., 2001; Kumar and Sridharan, 2009; Um et al., 2009; Joseph and Sridharan, 2011; Javadian et al., 2011). In this study, we focus on to improve of a real FMS achieve a global optimization in the improvement of a FMS performance with four work stations in a company in Ankara/Turkey considering design and operation related decision variables.

In the literature, Park et al. (2001) and Um et al. (2009) used hypothetical systems with hypothetical assumptions in their studies. Park et al. (2001) proposed a method for simultaneously optimization operational and design factors of a FMS with the multiple objectives via DOE, regression analysis and compromise programming. Eight operational and design factors were simultaneously optimized by compromising four performance measures that are obtained using regression analysis (Park et al., 2001). Also, Um et al. (2009) presented the combined study for the analysis of a FMS with an Automated Guided Vehicle system (AGVs). In their study to maximize the operating performance of FMS with AGVs, some factors were investigated, including the velocity, number, and dispatching rule of AGV, scheduling, work-piece types, and buffer sizes. They also considered the three performance measure namely minimizing the vehicle utilization, minimizing the congestion, and maximizing the throughput. Um et al (2009) used simulation-based optimization methods that are Multi-Objective Non-Linear Programming (MONLP) and Evolution Strategy (ES). MONLP obtained the design factors of the FMS through factorial design and regression analyses. However ES used to verify each factor for simulation-based optimization (Um et al., 2009).

Hypothetical models are useful to investigate system behaviors. On the other hand real systems generally need evaluating their own operating characteristics. In other words, hypothetical models seem to be inadequate to estimate a FMS performance in desired detail.

The main objective of this study is to present an operational improvement approach for existing real FMS design considering design and operational

variables using simulation optimization integrated DOE, regression meta-model and GP

2. A DESIGN AND PERFORMANCE EVALUATION MODEL FOR AN FMS WITH MULTIPLE OBJECTIVES

Simulation models have been widely used to mitigate the restrictions of the analytical models for designing and analyzing the FMS (Park et al., 2001, Um et al., 2009).

The FMS performance is determined by running the simulation model via a DOE. Firstly a 2^k full-factorial design is applied in this study to DOE scenario for the multiple-objective problem. Secondly, performance responses of the FMS are determined using a simulation tool and DOE method, statistical analysis are applied by ANOVA to obtain main and interaction effects of the design factors. Thirdly, the FMS performance responses are then transferred in a mathematical form with the identified significant main and interaction effects through a regression. Finally, a goal programming (GP) model is used by setting the response functions as objective functions and including FMS constraints. After that, the most suitable levels of design parameters in the GP model are determined. The GP is a relatively popular methodology in the literature and it has been used by many authors in different areas (Badri, 1999; 2001; Yurdakul, 2004; Lee et al., 2010; Liang, 2009; Ic et al., 2012). The proposed FMS design and optimization process can be described as in Figure 1.

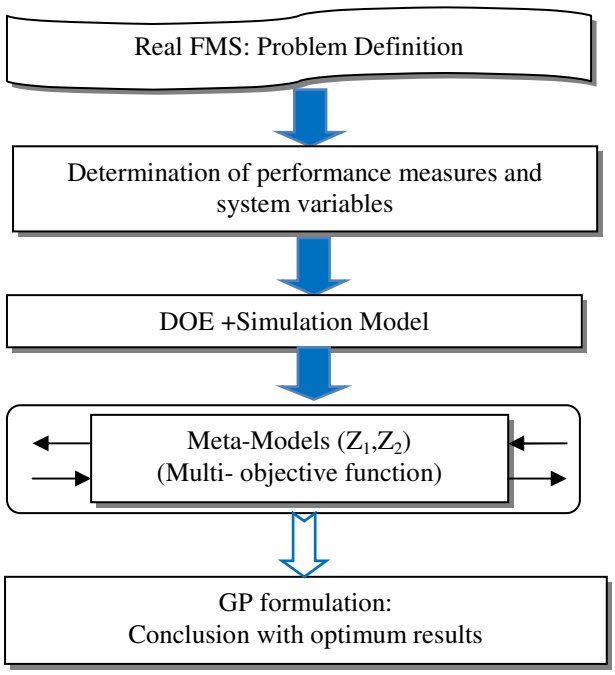


Figure 1. FMS design and optimization process

3. REAL CASE APPLICATION

This section presents an application for multi-response simulation optimization of a FMS via DOE, regression meta-model and GP. 'X Manufacturing (Turkey) Co.' is produces an extensive range of over 100 product varieties of automotive units, cams, cranks, shafts, motor blocks, pistons and transmission elements for world leading automotive manufacturers.

The FMS considered in this research, which is producing brake cylinder casing, gear box and flywheel housing for automobile manufacturers. FMS studied in this research is a dedicated type FMS which allows a dedicated process routing of parts to machining centers. There are four CNC machining centers (MAZAK FH6800) with one separate local buffer storages (20 pallet capacity) for work pieces. Work pieces in FMS are moved via transporting robot on bidirectional paths, and processed at one of the appropriate CNC machining centers. FMS layout is given in Figure 2. If there is no work pieces for the load-unload station the transporting robot is completed their process, then it stays idle at the current CNC. One of the attribute of the dedicated FMS system is its no-routing flexibility, which no allows work piece to be processed on more than one alternative CNC per process (Park et al., 2001; Groover, 2008). The alternative CNC machining centers for a specific process of a part are pre-determined via prior analysis to balance the workloads among CNC machining centers. Hence, when a work piece enters the FMS, the dedicated CNC machining centers for each process are already known based on the work piece type of the process (Park et al., 2001). The buffers feeding machining centers have same sizes.

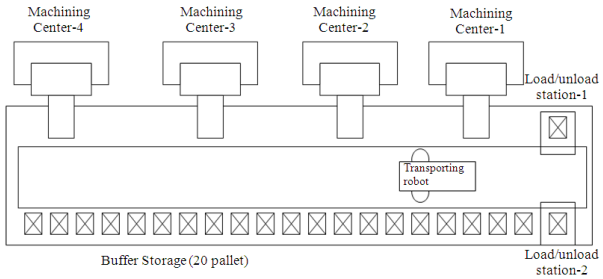


Figure 2. FMS layout

The company has applied employee training activities, line-balancing techniques for improvements of the FMS. However, the FMS still faced obstacles due to delay of product delivery to the customer for different causes. The main cause behind this delay is the long Cycle Time (CT).

Five design and operation variables determined based on system behaviors and expert opinions: number of operator in load-unload station, velocity of material handling robot, number of pallets, number of cutting tool, pallet routing scheduling rules (Table 1). Two performance measures such as maximizing throughput (units/h) and minimizing cycle-time (h) are considered.

Table 1. FMS design and operation factor's levels

Factors		Levels	
Symbol	Content	Low (-1)	High (+1)
x_1	number of operator in load-unload station	2	4
x_2	velocity of material handling robot (m/min)	1	2
x_3	number of cutting tool	120	200
x_4	pallet routing scheduling rules	Random	Dedicated to first non-busy pallet
x_5	number of pallet	20	40

3.1. Design of Experiments

Five factors are determined in the FMS design problem to optimize two performance measures of cycle time (h) and throughput (units/h). They include five design parameters so the DOE for the FMS problem involves five factors and two levels in each factor as given in Table 1.

3.2. System Modeling of the FMS via Simulation

The FMS presented in Figure 2 is modeled using the ARENA® Simulation Software by Rockwell Automation. Work pieces enter the system based upon exponential distribution. The number of process required for each job type is in the range of three to five operations, and then when a work piece comes into the FMS, the processes are assigned depending upon the job type.

The processing times in FMS vary from seconds to minutes depending upon the nature of required operations. The processing times of operations are pre-determined and fixed for each part type. Work piece handling time is computed by dividing the rectangular distance by transporter speed. Each simulation experiment was carried out for the operation of the FMS over a period of 2 month or 1152 h (48working days and 24h per day).

For validation and verification of the simulation model, all the necessary data are collected via DOE from the simulation experiments. To test its validity, the TRACE command, one of the ARENA output commands, is used to verify the model. This permits the user to watch step by step, generated and running the model on the time basis to see how well it represents the FMS under specific assumptions considering real FMS (Ayag, 2007).

3.3. Design of Experiment

Using the full-factorial DOE, 320 ($2^5 \times 10$ replication) simulation runs were conducted on MINITAB®. The ANOVA is applied to determine the significance of main and interaction effects of the five design factors (Park et al., 2001). It should be noted that since the significance level is set at 5%, effects with a $p \leq 0.05$ significantly contribute to the corresponding FMS performance measure. Functions of two FMS performance measures (cycle time and throughput) are obtained using regression analysis with identified significant main and interaction effects.

3.4. Validation of the Meta-Model

Simulation validation illustrates how well the model represents the real world system. Validation of the meta-model is obtained with respect to the underlying simulation (Dengiz et al., 2006). Therefore, the validation of a meta-model is obtained by making many comparisons between the outputs of the meta-model and the simulation model (Dengiz et al., 2006; Dengiz and Akbay, 2000; Kleijnen and Sargent, 2000; Dengiz and Belgin, 2007). To decide whether to accept a meta-model the Absolute Relative Error-ARE (see Kleijnen and Sargent, 2000 and Dengiz et al., 2006) is used. To assure the validation of the meta-model built in this study, the-meta model was tested against simulation runs at fifteen randomly selected design points within their permissible ranges (Dengiz et al., 2006). Then, the results obtained from this simulation runs were compared with the values obtained from meta-model using the same combination of parameters (Dengiz et al., 2006).

3.5 Multi-Objective Simulation Optimization

Regression models of two FMS performance measures (throughput and cycle time) are obtained using MINITAB. The FMS design problem can be formulated via a multi-objective programming technique as follows:

$$\text{MIN } Z = P_1*d_1^- + P_2*d_2^+ \quad (1)$$

Subject to

$$651.42 + 18.64*x_1 + 7.38*x_2 - 0.17*x_3 + 20.08*x_4 - 1.00*x_5 + 42.26*x_1*x_2 - 17.70*x_1*x_4 + 42.96*x_2*x_4 + 1.43*x_4*x_5 - 43.78*x_1*x_2*x_4 + d_1^- - d_1^+ = 800; \quad (2)$$

$$1.09787 - 0.03075*x_1 + 0.09345*x_2 + 0.00963*x_3 + 0.01978*x_4 - 0.00140*x_5 - 0.07315*x_1*x_2 + 0.05018*x_1*x_4 + 0.01586*x_2*x_3 + 0.03012*x_2*x_4 - 0.01318*x_2*x_5 + 0.01220*x_3*x_4 + 0.08693*x_1*x_2*x_4 + 0.01830*x_1*x_4*x_5 - 0.01417*x_2*x_3*x_5 - 0.01690*x_2*x_4*x_5 + 0.01355*x_1*x_2*x_3*x_4 + d_2^- - d_2^+ = 1; \quad (3)$$

$$2 \leq x_1 \leq 4 \quad (4)$$

$$1 \leq x_2 \leq 2 \quad (5)$$

$$120 \leq x_3 \leq 200 \quad (6)$$

$$x_4 \in \{0,1\} \quad (7)$$

$$20 \leq x_5 \leq 40 \quad (8)$$

$$d_i^-, d_i^+ \geq 0, \quad i=1,2 \quad (9)$$

$$x_i \in Z, \quad i=1,2,3,4,5 \quad (10)$$

$$P1 >> P2 \text{ (pre-emptive priority levels)} \quad (11)$$

The Goal Programming (GP) model has been solved using MS Excel® Solver tool. The model is admitted as a successful result, and the goal programming process terminates with the solution given below:

Number of operator	: 3
Velocity of transporter (m/min)	: 1.073792
Number of cutting tool	: 120
Pallet selection procedure	: Random
Number of pallet	: 31

4. CONCLUSIONS

This is the first kind of a model to solve a real case multi-response FMS performance improvement problem by using integrated DOE, regression analysis, and goal programming. It permits FMS planners to agreement interactively among conflicting goals while obtaining system performance parameters. The multi response approach with an integration of statistical tools and optimization theory can be used to other multi-response or multi-objective optimization problems that are too sophisticated to determine an objective function in a mathematical form. For future study, the proposed approach can be extended to incorporate the cost and economical issues and machine replacement analysis considering dynamic manufacturing environments such as machine reliability, technological level of CNC machines, mix of part types, utilization of facilities and logistical consequences of design changes. As a result of this study proposed approach can be able to improve FMS designs for determining the operation working conditions of the system.

REFERENCES

Ayag, Z., 2007. A hybrid approach to machine-tool selection through AHP and simulation. International Journal of Production Research,45(9), 2029-2050.

Badri, MA., 1999. Combining the analytic hierarchy process and goal programming for global facility location-allocation problem. *International Journal of Production Economics*, 62, 237-248.

Badri, MA., 2001. A combined AHP-GP model for quality control systems. *International Journal of Production Economics*, 72, 27-40.

Chan, F.T.S., Bhagwat, R., Wadhwa, S., 2007. Flexibility performance: Taguchi's method study of physical system and operating control parameters of FMS. *Robotics and Computer-Integrated Manufacturing*, 23, 25–37.

Dengiz, B., Bektas, T., Ultanir, A.E.,2006. Simulation optimization based DSS application: A diamond tool production line in industry. *Simulation Modelling Practice and Theory*, 14,3, 296-312.

Dengiz, B., Belgin, O., 2007. Paint shop Production Line Optimization Using Response Surface Methodology. *Proceeding of the 2007 Winter Simulation Conference*, pp.1667-1672.

Dengiz, B., Akbay, K.S., 2000. Computer simulation of a PCB production line: meta-modeling approach. *International Journal of Production Economics*, 63, 195–205.

Groover, M.P., 2008. *Automation, Production Systems, and Computer Integrated Manufacturing.* Third Edition, Pearson International Edition, USA.

Guo, Y., Chen, L., Wang, S., Zhou, J., 2003. A new simulation optimisation system for the parameters of a machine cell simulation model. *International Journal of Advanced Manufacturing Technology*, 21, 620–626.

Ic, Y.T., Elaldi, F., Pakdil, F., Ipek, N.E., 2012. Design of Experiment and Goal Programming Application for GMAW Process. *Welding Journal*, 91(4), 106-112.

Javadian, N., Aghajani, A., Rezaeian, J., Sebdani, M.J.G., 2011. A multi-objective integrated cellular manufacturing systems design with dynamic system reconfiguration. Int J Adv Manuf Technol, 56, 307–317.

Joseph, O.A., Sridharan, R., 2011. Effects of flexibility and scheduling decisions on the performance of an FMS: simulation modeling and analysis. *International Journal of Production Research*,1-21,In Press. DOI:10.1080/00207543.2011.575091.

Kleijnen, J.P.C., Sargent, R.G., 2000. A methodology for fitting and validating meta-models in simulation. *European Journal of Operational Research,* 120, 14–29.

Kumar, N.S., Sridharan, R., 2009. Simulation modelling and analysis of part and tool flow control decisions in a flexible manufacturing system. *Robotics and Computer-Integrated Manufacturing,* 25, 829–838.

Kumar, M.V.S., Janardhana, R., Rao, C.S.P., 2011. Simultaneous scheduling of machines and vehicles in an FMS environment with alternative routing. *International Journal of Advanced Manufacturing Technology*, 53, 339–351.

Lee, J., Kang, S-H., Rosenberger, J., Kim, S.B., 2010. A hybrid approach of goal programming for weapon systems selection, *Computers & Industrial Engineering,* 58, 521–527.

Liang, T-F., 2009. Fuzzy multi-objective project management decisions using two-phase fuzzy goal programming approach. *Computers & Industrial Engineering,* 57,1407–1416.

Mahdavi, I., Shirazi, B., Solimanpur, M., 2010. Development of a simulation-based decision support system for controlling stochastic flexible job shop manufacturing systems. *Simulation Modelling Practice and Theory*, 18, 768–786.

Ozmutlu, S., Harmonoskyz, C. M., 2004. A real-time methodology for minimizing mean flow time in FMSs with machine break downs: threshold-based selective rerouting. *International Journal of Production Research,* 42 (23), 4975–4991.

Park, T., Lee, H., Lee H., 2001. FMS design model with multiple objectives using compromise

programming. *International Journal of Production Research*, 39(15), 3513-3528.

Savsar, M., 2005. Performance analysis of an FMS operating under different failure rates and maintenance policies. *The International Journal of Flexible Manufacturing Systems*,16, 229–249.

Um, I., Cheon, H., Lee, H., 2009. The simulation design and analysis of a Flexible Manufacturing System with Automated Guided Vehicle System. *Journal of Manufacturing Systems*, 28,115-122.

Yurdakul, M., 2004. Selection of computer-integrated manufacturing technologies using a combined analytic hierarchy process and goal programming model. *Robotics and Computer-Integrated Manufacturing*, 20, 329-340.

AUTHORS BIOGRAPHY

Berna Dengiz is the dean of Engineering Faculty at the Baskent University. Her field of study is mainly modeling and optimization of complex large sized systems besides heuristic optimization. She has received research funding for her collaborative studies from NATO-B2 program, TUBITAK (The Scientific and Technical Research Council of Turkey), Government Planning Center of Turkey and National Science Foundation's (NSF) of the USA. She has worked as visiting professor at the University of Pittsburgh and the Auburn Universities.

Yusuf Tansel Ic is an Assistant Professor of Department of Industrial Engineering at the Baskent University. He received a PhD degree in Mechanical Engineering from Gazi University Institute of Science and Technology. He has more than 10 years of experience in banking industry. His research interests include application of expert systems to manufacturing systems, modeling and analysis of production systems, multi-criteria decision making, and financial risk management in commercial banks.

FIELD EXPERIMENTS FOR ENGINEERING AUGMENTED REALITY TOOLS

Gabriel A. Fernandes[(a)], Gerson G. Cunha[(b)], Celia Lopes[(c)], Luiz Landau[(d)], Alvaro Luiz Gayoso de Azeredo Coutinho[(e)]

[(a) (b) (c)] GRVa – LAMCE – COPPE - UFRJ
[(d)] LAMCE – COPPE - UFRJ
[(e)] NACAD – COPPE - UFRJ

[(a)]gabriel.ufrj@gmail.com, [(b)]gerson@lamce.coppe.ufrj.br , [(c)]celia@lamce.coppe.ufrj.br, [(d)]landau@lamce.coppe.ufrj.br, [(e)]alvaro@nacad.ufrj.br

ABSTRACT
This paper presents the field experiment results for a civil engineering augmented reality tool. The tool allows you to mark any number of interest points in large areas and link relevant information to them, such as: assembly, maintenance and inspection data. The information is seen on the screen through Augmented Reality technology. The system also has a specific content manager to help organize and manage information. The field experiment points out activity execution performance improvements when using Augmented Reality tools when compared to more traditional task management and execution methods.

Keywords: Augmented Reality, Mixed Reality, Virtual Reality, Civil Engineering.

1. INTRODUCTION

Augmented Reality (AR) is a major step in visualization technology, as it allows a user to interact with the combined strength of eye visible perception and digital information. Digital, in this case, can range from a simple text to a complex interactive multimedia presentation. This combination can be seamingless in a way that the digital content seems totally integrated in the real world, or not directly related in a visual manner, such as reading a QR Code and navigating to webpage.

When contextualized in any engineering field, this technology can be fitted to improve creative thinking and project design ((Shen, Ong et al. 2010), (Fuge, Yumer et al. 2011), (Huang, Yang et al. 2012), (Ran, Wang et al. 2011)). It can help in real assembly tasks or simple training ((Anastassova and Burkhardt 2009), (Ong and Wang 2011)). Also it can be fitted to support maintenance, supervision and inspection operations ((Shin and Dunston 2010), (Lee and Akin 2011)). As last, but not limited to, it can provide useful visual feedback and interaction between a diverse professional scope (Designer, Engineer, Technician, Manager, Client, User and others), allowing fast idea and knowledge transfer.

As presented in previous works ((Fernandes, Cunha et al. 2011), (Fernandes, Coutinho et al. 2010), (Fernandes and Cunha 2009), (Fernandes and Cunha 2010)) AR has been discussed, studied and experimented in various ways. The Applied Virtual Reality Group has focused a great effort in bringing together theory and practice in the AR field. The main AR system, currently under development, connects a project management system and a visualization tool which can be deployed through a wide range of platforms. The system is entirely based on a project database which delivers controlled XML files depending on tasks and user location. The system is scalable and can be easily adapted to interact with new tracking technologies and visualization techniques.

The objective has been to propose practical visualization tools for activity management and support. AR is mostly known for its marketing approach, usually to give products, ideas and services a technological front to the general public. Usually it faces a barrier when crossing to industrial scenario. The tool presented in this work has a polished user interface and a interaction concept based on at least five years feedback from direct projects with the Brazilian oil industry. This has the purpose of founding the current system with controlled field testing and general user feedback.

2. THE EXPERIMENT

This paper presents an experiment developed to test and analyze the previously created AR support system for engineering. The experiment is executed in a controlled environment where a sample group of ten will be asked to execute a series of activities using de the AR system. These activities focus on series of inspection, assembly and maintenance tasks where the user observer and interact with interest points and fill a report checklist during the process.

The experiment measures, through user feedback multiple tool aspects. It compares general visibility/readability under different lighting conditions and screen resolution limitations on mobile hardware; it measures interface design and ergonomic aspects; as for

functionality the experiment measures task execution time and correct completion. The users also input information through a score questionnaire which covers on several unique aspects of task and tool. The experiment collects subjective user input through multiple choice questions on personal perception and thoughts.

The environment is based on the second floor of the LABCOG building at the Technological Park at the Federal University of Rio de Janeiro. It proposes the ideal testing spot due to its unfinished status. Figure 1 presents a top view of the floor with four suggested enviroments. Each enviroment contains three to four interest points and each point has designated symbol which designates its location.

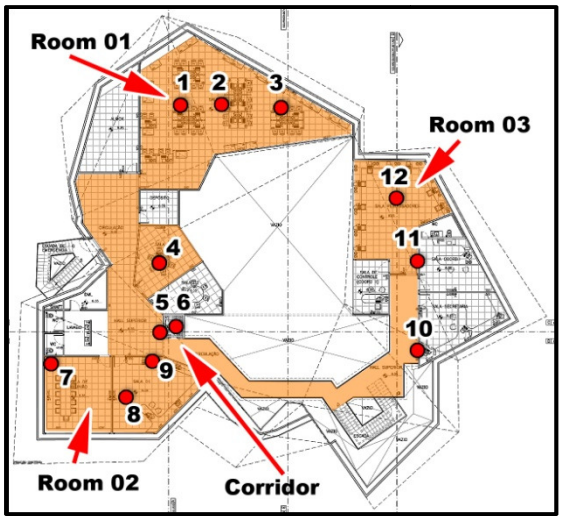


Figure 1. LABCOG Second Floor, Red Dots -> Interest Points.

Table 1 presents a list of task distributed through their respective locations on the experiment floor. The task type is identified as: Insp (Inspection), Assem (Assembly) and Maint (Maintenance). The tasks are simple and are restricted to visual inspection and simple interaction. To assist the user, the AR system provides visual information to guide and propose standards for comparison. Visual information is made available by images, diagrams, texts annotations and 3D models.

Grp	N.	Task	Type
R1	1	Installed Air Duct	Insp.
	2	Missing Air Vent	Asse.
	3	Adjust Color Button	Main.
Cor r.	4	Whole Repair on Floor	Inps.
	5	Exposed Structure	Asse.
	6	Missing Emergency Intercom	Asse.
R2	7	Read Equipment Values	Insp.
	8	Pending Lighting Installation	Asse.
	9	Pending Energy Connector	Asse.
R3	10	Fire Extinguisher Installation	Insp.
	11	Missing Electric Panel	Asse.
	12	Ajust Button on Wall	Main.

Table 1. Task distribution for AR Experiment.

Each interest point is marked with a printed panel which is unique and can be identified by the number on the top left square. The panel allows the tool to identify the interest point and overlay the correct visual information. Each panel is composed of four symbols, also known as fiducials, which provide redundant information for obstructed or partial visibility. For the identification to work, at least one symbol on the panel must be entirely visible. Figure 2 presents two panels used in the experiment.

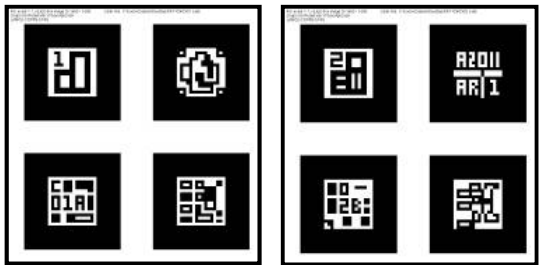


Figure 2. Symbols for interest points one and two.

Each task has an instruction set based on its category. The inspection guides the user to observe certain visual points and request feedback through the checklist. Figure 3 shows the user inspecting a fire hose compartment and observing key points marked by augmented reality content.

Figure 3. Fire hose box; top: application screenshot; bottom: user holding the tablet.

The assembly tasks are presented to the user as part numbers and part position. The user must choose the correct element based on photograph, 3d model or only identification number. Afterwards the user must place the part in the marked position seen through the

AR display. So the user is not overwhelmed by specific assembly steps the elements are represented by printed images which must only be placed correctly. Figure 4 shows the collection of parts used in assembly tasks during the experiment, there are a total of 15 elements which are divided in groups of three. Figure 5 shows an assembly scene where the user is observing the position indication on the tablet screen.

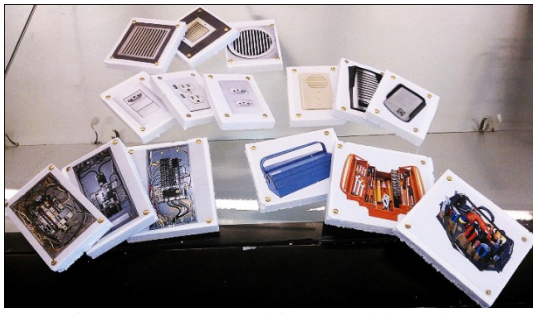


Figure 4. Parts used in assembly tasks.

Figure 5. Top: Assembly task screenshot; Bottom: user inspecting installed part in place.

The maintenance task is presented as a rotation button which needs to be correctly set to the value displayed through the AR screen. A button used in the experiment is shown on Figure 6, while the AR overlay display can be seen on Figure 7.

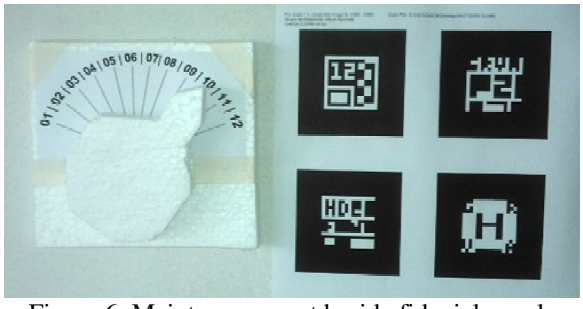


Figure 6. Maintenance part beside fiducial panel.

Figure 7. Maintenance task screenshot.

The last step to each task is the checklist, which can be accessed through the AR menu of each interest point while it's visible. The checklist has mostly boolean answers to identify if the observations where made correctly. Figure 8 shows a simple checklist for task one accessed through the tool browser. All the checklists are based on a web php platform which saves responses to a local file on the tablet for data analysis.

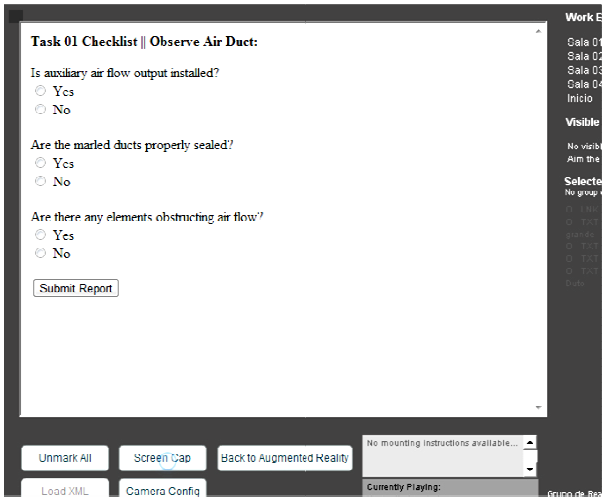


Figure 8. Task checklist screenshot.

1	App. presentation vs actual functionality
2	Camera Screen Visibility
3	Fiducial panel detection quality
4	Side menu visibility and touch screen interaction
5	Tablet comfort and ergonomical aspects
6	AR content quality
7	Inspection tools quality
8	Assembly tools quality
9	Maintenance tools quality
10	The whole experience of finishing the experiment
11	The prospect of future developments based on this tool

Table 2. Questionnaire: Tool usage aspects being observed by volunteers.

When the user completes the twelve interest point circuit he is presented with a short questionnaire to evaluate the tool and give feedback on key features. Table 2 presents the questionnaire. Each item can score from 0 to 7. At the end of the questionnaire the user is invited to give written feedback of aspects not covered by other evaluation methods. This personal feedback is presented further ahead.

3. FIELD OBSERVATIONS

The experiment was executed individually and each volunteer was presented to tool in turn. The overall perception and understanding of AR varied and the technological background was also very diverse. The volunteers can be distributed as follow:

- 2 – Graphic Designers
- 2 – Field Technicians
- 1 – Administrative Professional
- 2 – Product Designers
- 1 – CAD Professional
- 1 – Geographer

The experiment went well and each volunteer spent 5 minutes learning how to use the tool, around 11 minutes navigating the circuit, 5 answering the questionnaire and around 10 commenting the experience and providing verbal feedback. This sums a total of 30 to 40 minutes per volunteer. Most of the environment on the floor is crowded with interior finishing materials which delayed walking and proposed, in some cases, a challenge while holding the tablet. Most bad panel detection was due to low light conditions which were compensated by adjustable detection threshold values on demand. The volunteers used the tool well with almost no interruptions for questions. The tool itself worked flawlessly presenting no sudden crashes or slowdowns which could affect time or task execution.

4. COLLECTED DATA

As explained the data collected is distributed in three categories: execution time, circuit score and evaluation questionnaire. Figure 9 presents the completion time distribution. The time is distributed between 8 to 14 minutes. The variation can be associated with experience in the use of touch screen technologies, as some were able to quickly adapt to the interface while higher time scores had difficulty with screen finger pressure and finger aim. Although gender was not annotated during the experiment it should be noted that the female volunteers (3 out of 10) completed the track in the lowest times. Volunteer nine, being exposed for the first time to an AR tool, achieved the best time.

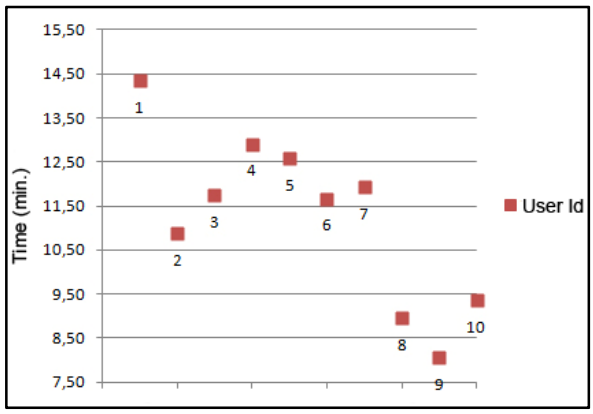


Figure 9. User time performance graph.

Almost all users achieved a perfect score and claimed that even with drawbacks they believed the instructions and visual aid to be clear and objective. Some volunteers also commented the task to be easy and could be more challenging. Following the circuit score is the questionnaire results which can be observed on Table 3 and Figure 10. Both present a good score, except for Menu and Comfort. The side menu was considered small and crowded, this affected finger interaction and aim. The letter size was considered small and in some cases the user needed to bring the display close to properly read the text. The comfort criteria was majorly affected by tablet weight and lack of grip around the edges, which made one hand hold unsafe. The users were explained that a score 7 in the fiducial criteria meant a perfect detection and tracking with imperceptible noise. In this case the fiducial was expected to receive a score around 6.

	User										
	1	2	3	4	5	6	7	8	9	10	Med.
Presentation	7	7	7	7	7	6	7	7	7	7	6,9
Screen	7	5	5	6	7	7	7	7	7	7	6,5
Fiducial	7	7	6	6	5	5	6	5	7	7	6,1
Menu	6	3	4	7	5	3	5	4	5	7	4,9
Comfort	5	6	6	7	7	5	7	3	5	7	5,8
AR Quality	7	7	6	7	7	7	7	5	7	7	6,7
Inspection	4	6	6	7	7	7	7	7	6	7	6,4
Assembly	7	7	5	7	7	7	7	6	7	7	6,7
Maintenance	5	7	7	7	7	7	7	7	5	7	6,6
Experience	6	6	5	7	7	5	7	5	7	7	6,2
Prospect	7	7	7	7	7	7	7	7	5	7	6,8

Table 3. User tool aspect evaluation and medium values.

The Inspection, Assembly and Maintenance criteria evaluates how the tool performed under the proposed tasks. The Inspection task was achieved the lowest score, as the pointing elements where considered confusing and hard to distinguish on the screen.

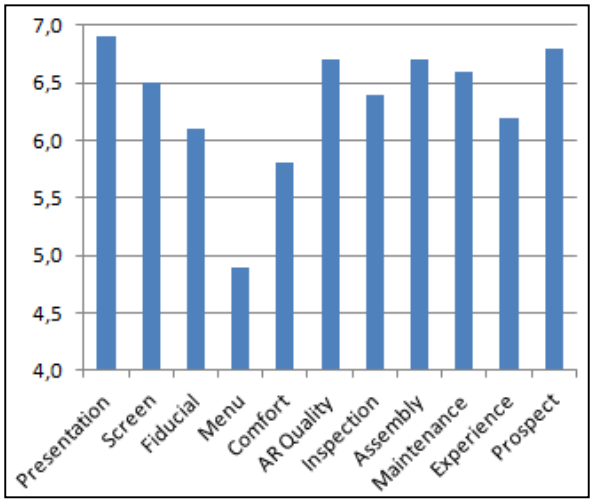


Figure 10. User tool aspect evaluation medium values.

5. USER FEEDBACK

Each user category made comments and suggestions relative to their expertise and personal experience. The following comments summarize post experiment talks: Both graphic designers commented constantly on the crowded interface, small text/buttons and tablet weight. The field technicians seemed fine with the whole tool and had no major complaints, as they observed the solution was light weighted compared to other materials normally carried. The administrative professional though the tool to be innovative and probably had much potential in field use. The product designers also made heavy comments on interface and weight; they also suggested improvements to reduce handleling discomfort. Bellow follows a list of the mains suggestion and comments:

- Fiducial ideal visibility distance indication, so the user doesn't stand too close or far from the panel.
- Tablet is uncomfortable for long period of work.
- Small interface controls make user interaction difficult.
- Interest point identification could be more evident, maybe marked on a map.
- Application tutorial within the tool guiding the between tasks.
- Computer verification of assembly activities alerting the user when the part is in the correct position or within tolerance levels.
- When the user is too close to the panel the content should auto-adjust to fit on the screen or suggest the user to take a step back.
- The inspection points could be more pronounced; currently they are small and not always entirely visible.

6. CONCLUSION

The experiment indicates that the AR tool can be useful and was well commented by the volunteers. Although well perceived, it needs adjustments in interface design and general comfort aspects. The presented tool was developed considering a wide range of applications and has many options not evaluated in current experiment, such as animation and audio feedback. The current tool would be more efficient if fragmented and task orientated. Further development and testing is in progress and could, in the future present, more impressive results. These results and further testing will help shape and direct the tool so it may reach an applicable build.

REFERENCES

Anastassova, M. and J.-M. Burkhardt (2009). "Automotive technicians' training as a community-of-practice: Implications for the design of an augmented reality teaching aid." Applied Ergonomics **40**(4): 713-721.

Fernandes, G. A., A. L. G. A. Coutinho, et al. (2010). Overview of Augmented Reality in Air Transport. Air Transport Research Society Conference. Porto, Portugal.

Fernandes, G. A. and G. G. Cunha (2009). Realidade Aumentada como Ferramenta de Acompanhamento em Projetos de Engenharia. 30 Congresso Ibero-Latino-Americano de Métodos Computacionais em Engenharia. Buzios, Brazil.

Fernandes, G. A. and G. G. Cunha (2010). Augmented Reality as a Support Tool for Engineering Projects. III The International Workshop on Applied Modelling & Simulation. Rio de Janeiro and Buzios, Brazil.

Fernandes, G. A., G. G. Cunha, et al. (2011). NAV – The Advanced Visualization Station: A Mobile Computing Center for Engineering Project Support. The 10th International Conference on Modeling and Applied Simulation. Rome, Italy.

Fuge, M., M. E. Yumer, et al. (2011). "Conceptual design and modification of freeform surfaces using dual shape representations in augmented reality environments." Computer-Aided Design(0).

Huang, S.-H., Y.-I. Yang, et al. (2012). "Human-centric design personalization of 3D glasses frame in markerless augmented reality." Advanced Engineering Informatics **26**(1): 35-45.

Lee, S. and Ö. Akin (2011). "Augmented reality-based computational fieldwork support for equipment operations and maintenance." Automation in Construction **20**(4): 338-352.

Ong, S. K. and Z. B. Wang (2011). "Augmented assembly technologies based on 3D bare-hand interaction." CIRP Annals - Manufacturing Technology **60**(1): 1-4.

Ran, Y., Z. Wang, et al. (2011). "Trends of mixed reality aided industrial design applications." Energy Procedia **13**(0): 3144-3151.
Shen, Y., S. K. Ong, et al. (2010). "Augmented reality for collaborative product design and development." Design Studies **31**(2): 118-145.
Shin, D. H. and P. S. Dunston (2010). "Technology development needs for advancing Augmented Reality-based inspection." Automation in Construction **19**(2): 169-182.

AUTHORS BIOGRAPHY

Gabriel A. Fernandes
Main AR developer and researcher for the Applied Virtual Reality Group (GRVa) in the Federal University of Rio de Janeiro. Graduation in Industrial Design (UFRJ), MSc in Computational Systems (COPPE/UFRJ) and currently concluding his DSc in Computational Systems (COPPE/UFRJ) has worked with AR since 2005.

Gerson Cunha
Researcher at. Universidade Federal do Rio de Janeiro (UFRJ). Has a DSc. degree on Civil Engineering in Federal University (PEC) at COPPE/UFRJ. Actually he is master's and doctorate professor at the same institution. He has great interesting in Computer Graphics subjects as Scientific and Computer Vision, Virtual Reality and Augmented Reality.

Maria Célia Santos Lopes
DSc. in High Performance Computing - Civil Engineering from the University Federal of Rio de Janeiro (2004). He has experience in development of simulation systems in Engineering, Scientific Visualization, Text and Web Mining, Data Mining using neural networks and other methods.

Luiz Landau
DSc. Civil Engineering from Universidade Federal of Rio de Janeiro (1983). He is currently professor of the Alberto Luiz Coimbra Institute for Graduate Studies and Research in Engineering, Federal University of Rio de Janeiro (COPPE / UFRJ), Coordinator of the Laboratory of Computational Methods in Engineering Civil Engineering Program (LAMCE / PEC / COPPE / UFRJ), Advisor to the Foundation for the Coordination of Projects, Research and Technology Studies-COPPETEC, IA CNPq Researcher and Coordinator of the PRH-02 ANP.

Alvaro Coutinho
DSc. in Civil Engineering from Universidade Federal do Rio de Janeiro (1987). He is currently a professor at the Federal University of Rio de Janeiro, a member of the editorial board Int J for Numerical Methods in Fluids, Comm Numerical Methods in Engineering, Latin American J Solids and Structures, consultant for the COPPETEC Foundation. He has experience in Civil Engineering with emphasis on structures, mainly in the following topics: finite elements, parallel computation, finite element method, computational fluid dynamics and vector computation.

AUGMENTED REALITY SYSTEM WITH CHROMA KEY FOR SIMULATORS

Mário Luiz Ribeiro [(a)], Gerson Gomes Cunha[(b)], José Luis Drummond Alves[(c)], Maria Celia Santos Lopes[(d)], Gabriel A. Fernandes[(e)], Luiz Landau[(f)], Cezar H. V. da Costa[(g)]

Universidade Federal do Rio de Janeiro / COPPE
(a) (b) (c) (d) (e) (f) (g)

[(a)] ribmalu@lamce.coppe.ufrj.br, [(b)] gerson@lamce.coppe.ufrj.br, [(c)] jalves@coc.ufrj.br, [(d)] celia@lamce.coppe.ufrj.br, [(e)] gabriel.ufrj@gmail.com, [(f)] landau@lamce.coppe.ufrj.br, [(g)] veiga@lamce.coppe.ufrj.br

ABSTRACT
This paper presents a training solution for heavy equipment operator professionals which need to improve general operation skills, exercise or wish to experiment with different risk decision making scenarios. Commercial simulators usually have high cost, demand great large installation areas or displacement to specialized training facilities. This solution combines the advanced Augmented Reality techniques, Chroma Key and Head Mounted Displays (HMD) to provide acceptable, low cost and on site training environment.

The system installed in limited space when compared to the commercial simulators. Through video composition the application adds the user's own and real equipments/controls over the virtual training environment. This work, also presents results on a comparison between a color projection and a high contrast cloth for mask filtering. The overall result provides acceptable results for general training purposes.

Keywords: Augmented Reality, Chroma Key, Immersion

1. INTRODUCTION

In some situations, the expertise and the ability are indispensable factors to the execution of the work to be accomplished.

In all areas, the training process becomes important for forming qualified professionals'. In this context, the objective of training future professionals, lies in the real sensation of being in the environment. (Crane cabin, Airplane, etc)

The basic idea was developing a simulation system which reproduced with the maximum realism, considering spatial limitations and cost factors, the true situations faced by the professional in the future, however in an augmented virtual environment.

2. SISTEMS IN THE APPROACHED AREAS

With the objective of offering a good immersion, there are several works, such as, the work of Ronald T. Azuma, 1997, which presents techniques of Virtual Reality or Augmented Reality for visualization in real time.

However, in general we face limitation like the cost of equipments and the necessary space for the use of the whole apparatus. The first part of this work presents systems used in the visualization training field and respective equipments used.

2.1. Delta 3D

The Delta 3D is a Open Source program with characteristics of a Game Engine appropriate for a wide variety of usage including simulation, formation or other graphic applications.

This program was used by the Federal Aviation Administration (FAA); its objective was to improve operational procedures and better pilot performance when following Global Positioning Satellite (GPS) routes.

Figure 1: Delta 3D using Digital Cave in “U”.

2.2. Virtualization Gate

The Virtualization Gate is an immersive environment similar to CAVE-based technology, but for use with the HMD (Head Mounted Display) and based on sophisticated Chroma key tracking system. The solution overcomes several limitations present in Occlusion based Augmented Reality applications. Allowing full body immersion and interaction with the virtual world by scanning the users’ contours from several points of view and generating a 3D avatar with similar appearance. Having its shape captured in real time allowing its insertion in any Virtual Reality application.

This system uses multiple cameras, a 3D modeling tool and real time texture generation for avatar creation

in low, but acceptable, resolution. However, the challenge of this system is the ability to render a finer quality model with a higher refresh rate, which could improve immersion.

Figure 2: Visualization of the model with low quality.

2.3. VAR-Trainer

The project VAR-Trainer was financed by the European Union, with the main objective to create and develop a simulation program for heavy machine operation. This training reduces accidents related to operator error.

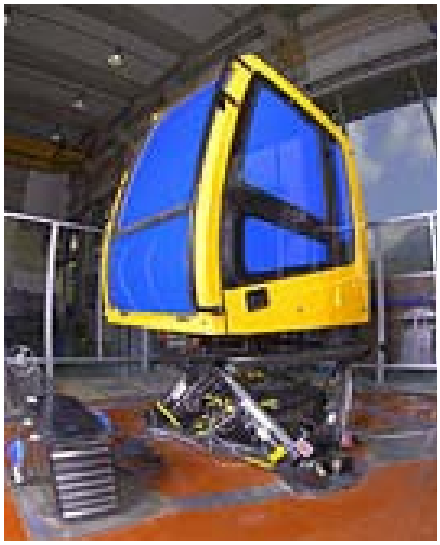


Figure 3: VAR-Trainer.

2.4. ChrAVE

The ChrAVE, is a system for helicopter pilots' training, based on Chroma Key technology. The objective is to supply a realistic simulated environment where a pilot can improve abilities which are important for general mission success.

Since it uses real helicopter structure to function it can implemented in the pilots own work environment.


Figure 4: ChrAVE in the pilot's cabin with blue screen covering the windows

3. OBJECTIVES

The objective is to develop a training environment which can, through Augmented Reality and Chrome Key, allow the user to interact with virtual content without the need for avatars. The simulator blends virtual and real content, in this case, the user will see not only himself but other relevant equipment and controls.

3.1. Augmented Reality

The use of virtual reality for training applications has proved itself useful and been validated by several previous works, such as Ayman Wasfy, 2004 and Fuhua Lin, 2002. Although Augmented Reality has been around for about the same time as Virtual Reality, only recently has it emerged as practical a tool. Augmented Reality can extend the perception of space and the information about that space.

One of the most important qualities of the Augmented Reality is its tangible properties in user interaction. AR makes it possible to mix real and virtual elements in one user interactive environment. The real world, in this case, can be the foundation for a virtual simulation and training exercise. Virtual objects can visually interact with virtual ones through object occlusion techniques, so virtual objects can pass behind real objects. Current limitations lie self user visibility, since occlusion of the users body cannot be done the same way as solid, non organic objects.

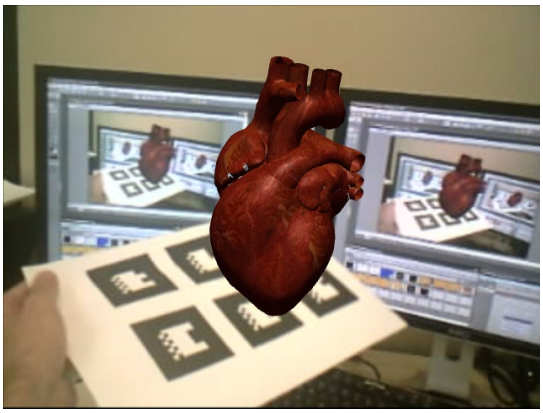


Figure 5: Model for study.

3.2. Chroma Key

Chroma Key is a technique commonly used in the film and television industry. It can be used to superimpose two video feeds, by filtering out specific colors that stand out. Usually this can done play placing people or objects in front of a luminous green, blue or red screen and later using these contrasting colors as a mask for image composing.

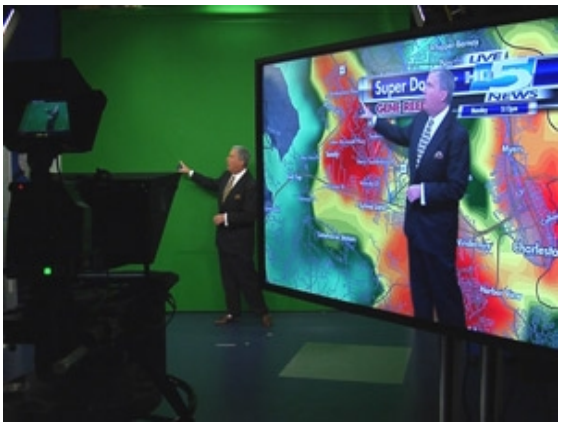


Figure 6: Chroma Key – Image Composition.

4. DEVELOPMENT OF THE SYSTEM

After an overview on Augmented Reality and Chroma Key applications, it was determined that and integration of these technologies would be possible by altering AR the rendering pipeline to include overlay of a masked video layer. The combined use of AR visualization system with Chroma Key filtering for self and instrument visibility improves immersion qualities on the crane training application case.

This application was developed and tested in the LAMCE/GRVA (Laboratory of Computational Methods in Engineering / Applied Virtual Reality Group) CAVE. This work also discusses the integration of real instruments and controls with a virtual environment by Chroma key AR training application. Below follows a summary of software and hardware used throughout the system.

4.1. Director

Adobe Systems Program, allowing to the user to develop multimedia applications, it has support for vectorial graphs and interactive three-dimensional objects with the use of Shockwave 3D.

4.2. DART

Group of tools developed to provide support for development of Augmented Reality applications. It was built an extension for Macromedia Director Environment, integrating tools like OpenAL (for sound 3D), VRPN (real team streaming dates) and ARToolkit (optical tracking).

4.3. 3D Max

A complete 3D graphic application that has sophisticated modeling tools and an extensive list of plugins, used for game design, architectural visualization, special effects , TV studio and industries in the development and study of new products.

4.4. VFX3D

It is a helmet with head phones (stereo) and an articulated visor containing two optic screens. Used for mono or stereo vision.

4.5. Intersense I-1200 Vistracker

The I-1200 has a 6-DOF (degrees of freedom) and high precision, with vision-inertial system for the movable control and autonomous control in simulation and training.

4.6. Creative Live! Cam Notebook Pro

This Webcam allows to generate pictures with 1.3 Megapixels, it possesses an adjustment of 180° and remote control.

5. THE SYSTEM PROCESS

For this system, the process is divided in three stages:

- First Stage: The image capture from a Webcam, with the use of libraries, such as, DirectShow for Windows (Videoimput Lib).
- Second Stage: It consists in converting the primary colors of the image from RGB (Red, Green, Blue) to the HSL format (Hue, Saturation, Lightness) representing the points in a model of colors RGB. This provides better tone control for frame color filtering.
- Third Stage: This stage performs the extraction of the images alpha channel and the assigning values based on comparison with the reference color pre-selected.

With the generated image in RGBA format, it is passed to Director which receives the data and saves it to a specific variable to store the image. Allowing the user to access and manipulate this image in a friendly environment through scripts (Lingo, JavaScript). The Shockwave 3D, is the 3D format based on the Director. This format is based on a complete system called Stenograph that enables you to change, manipulate and control nearly all aspects of visual and functional environment or a 3D object.

The Director 3D engine is efficient in layer composition and can easily blend two-dimensional and three-dimensional elements. In this case, the system enables Stenograph add graphics designed respectively in front and behind the 3D content, known as:

- Overlays – texture overlay, shown in front of all models on the list of a camera being used.
- Backdrops – Image appears behind the 3D content that will be processed in the projection of the camera.

The simulator developed in this work uses a Chroma Key video filtering to produce a masked overlay layer for the 3D, but the overall system is based on three composed layers which are listed below:

- The last layer drawn (Backdrop), is the actual video or captures image without transparency Webcam (camera image unchanged).
- The core layer or the second layer has a three-dimensional environment, in which case the environment of the port and the crane used for training.
- In the first layer (overlay) or top layer, receives the manipulated video with alpha channel.

The top overlay layer distinguishes this work from conventional Augmented Reality, by allowing the user to see the equipment, himself and the virtual environments.

For correct virtual orientation and display precise head tracking is required. Orientation tracking for this application was produced by the use of a high response and precision sensor called Viz-Tracker or IS-1200 (Intersense). This sensor is a small and box shaped which is fixed on the HMD.

Sensor access inside Director is available through an extension compiled using Intersense SDK, this extension binds the "isense.dll" with the Director Script interface (Lingo). Through a few simple access methods it is possible to monitor sensor state, calibrate and convert between Euler and Quaternion angle representation. Quaternion orientation was preferred over Euler due to unresolved gimbal lock problems apparently presented in Euler transforms available in Director.

5.1. Development

Taking into account the large number of cameras on the market, including the ability to control their settings.
The wide variety of software and libraries that would perform such a task. The first step in the implementation of this technique is to correctly capture a video feed. To accomplish this task the videoInput library was used.

This system uses a video feed which is spitted in two separate processing pipelines. The first is an unaltered version of the video and the second is the masked overlay video. The first step was the creation of a virtual camera system which retrieves video from a real camera and converts the color system to HSV for later control. This system creates the first virtual camera which will be accessed by videoInput extension.

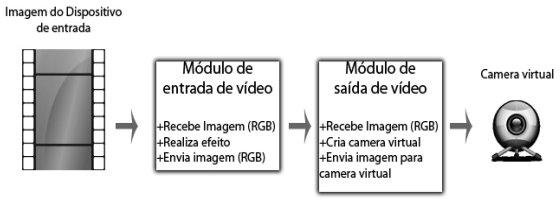


Figure 7: Real Camera to Virtual Camera

Once the video is properly converted from RGB to HSV, a simple algorithm can filter the selected Chroma color and adjacent lightness values. Below follows a step-by-step description of the filtering and masking process:

1. Capture the current RGB frame of the video and then store the video on an image created by video module.
2. In RGB color space, each pixel is converted to the HSV color space (hue, saturation and value - hue, saturation and value). The possibility of a better description of the relationship between colors. That is, this conversion is an alternative to the perceptual non-uniformity of colors from RGB space.
3. As for the HSV color space, the pixels are compared with the color chosen by the user and then with his color components in RGB color space, or alpha component (RGBA being) are changed when you have the same color or a color that is the desired tolerance range.
4. Applications of linear blur filter in order to smooth the edges that the effect of Chroma Key can affect.
5. The picture that has changed color is sent to the output module, which appears as a virtual camera.

The Director multimedia platform was chosen for its capacity to easily manage masked layers inside its 3D engine. Through its extension it was possible to import in real time the unaltered and masked video into planar textures to be used in simple overlay and backdrop composing.

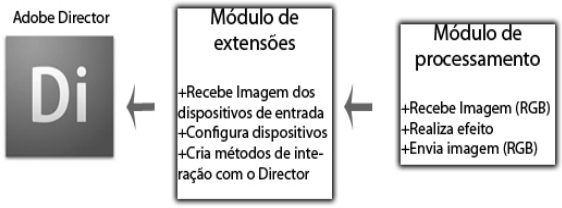


Figure 8: Image filtering and sending to Director.

5.2. AR Solution

Director provides more than one solution to work with Augmented Reality. AR is mostly based on tracking systems, which can be based on video processing or sensor data. As mentioned before, positioning sensors where used, for their stability and precision, none the less, experiments where made with fiducial markers and are commented below. Communication between sensor and Director is achieved by using the Intersense Xtra. The challenge is to integrate position information into virtual environment considering scale and coordinate system alignment.

6. RESULTS

6.1. Test-1

In the first tests, the fiducial marker did not produce good results. The tracking process presented a constant inaccuracy in marker recognition, this produced a leg in virtual camera positioning. It was also observed errors

in tracking, even when the marker was clearly visible and visual detection thresholds where within acceptable tolerances.

6.2. Test-2

The first Chroma Key test the CAVE environment with the HMD. The experiment placed the within a crane cabin while he the user was sitting on the real control chair. Through the HMD the user could see himself and the real controls composed over a virtual dock. The video user image was captured by a simple webcam and tracking was done a position/orientation sensor. The camera captured user images over a green contracting background. The green background was generated by projection green light over the CAVE walls.

This experiment produces the equivalent result of using several projectors to generate the virtual dock immersive environment. The main problem was green reflected light leaking into on the user image and producing distorted colors on the final composition. This was due to the reflection of green light projected on to the walls. Unfortunately this hard to correct without distorting general image color, loosing contrast and altering image lightness.

Figure 9: Environment with the effect of light green.

Figure 10: Own vision of the observer in the simulation.

6.3. Test-3

For the second Chroma Key test, the same CAVE configuration was maintained, but with the addition of two spots lights directed to the equipment (mobile base) and the user, to try to counteracting the green light effect. Some improvement can be observed by comparing Figure 10 with Figure 11.

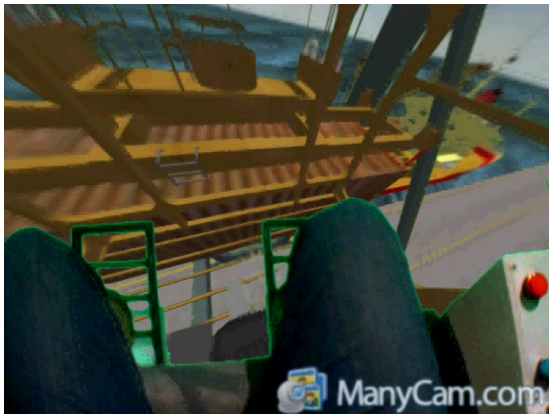


Figure 11: New result in the observer's vision.

6.4. Test-4

This experiment used the same cave environment as tests 2 and 3, but without background color generated by light projection. The background color, in this case, was generated by a special green colored cloth like material, which has high contrast properties.

Figure 12: Environment to simulation with Chroma Key cloth background.

Figure 13: Testing Lighting.

In this test, the results were satisfactory as can be seen in Figure 14. The results were more consistent with the proposed objective, because the equipment and the

user are no longer influenced by the reflected light green.

Figure 14: Vision of the observer without the influence of green light.

6.5. Final Considerations

The system proved to be a functional tool, reaching all the initially established goals. Enabling the extraction of real world elements through Chroma Key technique and allows the viewing and use of real equipment for training in various types of simulators. The virtual environment then extends real world elements and allows simulation of endless training scenarios.

With the insertion of the wearer and the equipment in the foreground, we see that there was a consistent position relative to the observer with the virtual objects and real objects, allowing a more realistic interaction than a conventional process of Augmented Reality. It is worth observing that the implementation of the Chroma Key technique was achieved entirely by software and using a low cost video capture system.

With the addition of Chroma Key Techniques and Augmented Reality, this system demonstrates the feasibility of using a simulator in small areas at a reasonable cost. Allowing adaptation to a simulated environment offering an immersive feel and consistent in order to meet the needs.

6.6. Future Implementations

Future research is still required in relation to low cost video capture devices with better color correction and exposure control. These improved device qualities would help in augmenting color fidelity and filtering by computer software.

It also intends to implement the system allowing the visualization and extraction of depth map in stereo mode, which will allow the placement of virtual objects relative to the observer without the need for markers. This would greatly improve the overall user depth perception. Using two cameras to catch the real elements, with a slight viewpoint shift, providing conditions for the placement of observers in the virtual space allowing a virtual object is manipulated.

REFERENCES

Autodesk 3Ds Max.
http://usa.autodesk.com/adsk/servlet/pc/index?id=13567410&siteID=123112 [accessed 19 May 2009]

Adobe Director 11
http://www.lingoworkshop.com/Articles/history.php [accessed 26 April 2008]

Ayman Wasfy, Tamer Wasfy, Ahmed Noor, Intelligent virtual environment for process training, June 2004.

ChrAVE - Chroma Keyed Augmented Virtual Environment.
http://www.delta3d.org/article.php?story=20041201225135603&topic=projects [accessed 24 September 2009]

Chroma Key - The Blue / Green Screen Page
http://www.seanet.com/~bradford/blue_green_screen_visual_effects_1.html [accessed 28 April 2009]

Creative Live! Cam Notebook Pro
http://us.creative.com/products/product.asp?category=218&subcategory=553&product=14809&nav=0 [accessed 20 October 2008]

DART – Group of Tools for Augmented Reality Project
http://www.lamce.ufrj.br/grva/data/realidade_aumentada/dart.htm [accessed 12 February 2009]

Delta3D - Open Source Gaming & Simulation Engine
http://www.delta3d.org/ [accessed 21 January 2009]

Director - Xtra Development Kit
http://kb2.adobe.com/cps/160/tn_16064.html [accessed 26 April 2009]

Fuhua Lin, Lan Ye, Vincent G. Duffy, Chuan-Jun Su, Developing virtual environments for industrial training, January 2002

InterSense IS-1200 VisTracker / Inertia Hawk
http://www.inition.co.uk/inition/product.php?URL_=product_mocaptrack_intersense_IS-1200&SubCatID_=19
http://cb.nowan.net/blog/2007/07/20/intersense-is-1200-vistracker [accessed 12 February 2009]

Ronald T. Azuma, A Survey of Augmented Reality, Hughes Research Laboratories, Malibu, CA - August 1997.

VFX3D Virtual Reality Helmet
http://www.digit-life.com/ [accessed 7 June 2009]

Virtualization Gate, Immersive Environment.
http://grimage.inrialpes.fr/vgate/VGate/VGate.html [accessed 19 August 2009]

AUTHORS BIOGRAPHY

Mário Luiz Ribeiro
Postgraduate Course in Civil Engineering, Master of Science in Civil Engineering in 2009, in area of Computing Systems at Federal University of Rio de Janeiro (UFRJ - Brazil) in 2009. Working as a 3D modeler in the areas of Virtual Reality and Augmented Reality since 1997 at LAMCE / GRVA, at COPPE/ UFRJ.

Gerson Cunha
Researcher at. Universidade Federal do Rio de Janeiro (UFRJ). Has a DSc. degree on Civil Engineering in Federal University (PEC) at COPPE/UFRJ. Actually he is master's and doctorate professor at the same institution. He has great interesting in Computer Graphics subjects as Scientific and Computer Vision, Virtual Reality and Augmented Reality.

Maria Célia Santos Lopes
DSc. in High Performance Computing - Civil Engineering from the University Federal of Rio de Janeiro (2004). He has experience in development of simulation systems in Engineering, Scientific Visualization, Text and Web Mining, Data Mining using neural networks and other methods.

José Luis Drummond Alves
is an Associate Professor at the Engineering Graduate Center of the Federal University of Rio de Janeiro, Brazil; Specialist in Nonlinear Finite, Element Analysis, Finite Element Technology and Code Development; Consulting Services in Engineering for Oils & Gas Industry at COPPE / UFRJ - Program of Civil Engineering, LAMCE - Laboratory of Computational Methods in Engineering.

Luiz Landau
DSc. Civil Engineering from Universidade Federal of Rio de Janeiro (1983). He is currently professor of the Alberto Luiz Coimbra Institute for Graduate Studies and Research in Engineering, Federal University of Rio de Janeiro (COPPE / UFRJ), Coordinator of the Laboratory of Computational Methods in Engineering Civil Engineering Program (LAMCE / PEC / COPPE / UFRJ), Advisor to the Foundation for the Coordination of Projects, Research and Technology Studies-COPPETEC, IA CNPq Researcher and Coordinator of the PRH-02 ANP.

Gabriel A. Fernandes
Main AR developer and researcher for the Applied Virtual Reality Group (GRVa) in the Federal University of Rio de Janeiro. Graduation in Industrial Design (UFRJ), MSc in Computational Systems (COPPE/UFRJ) and currently concluding his DSc in Computational Systems (COPPE/UFRJ) has worked with AR since 2005.

Cezar Henrique Veiga da Costa
Graduated in Computer Science with emphasis in Systems Analysis from the University Estacio de Sa (2005) and Masters in Civil Engineering with specialization Computing Systems from Federal University of Rio de Janeiro (2010). He has experience in Computer Science in Computing Systems, mainly in the following areas: virtual reality and augmented reality, computer networks and peripheral interfaces.

ADVANCED ETCHING ALGORITHMS FOR PROCESS SIMAULTION OF 3D MEMS-TUNABLE LASERS

A. El Boukili

Al Akhawayn University in Ifrane, Morocco

Email:a.elboukili@aui.ma

ABSTRACT

We are proposing new and advanced etching algorithms for the process simulation of three dimensional (3D) micro electro mechanical systems (MEMS)-tunable vertical cavity semiconductor optical amplifiers (VCSOAs). These algorithms are based on advanced domain decomposition methods, Delaunay meshing algorithms, and surface re-meshing and smoothing techniques. These algorithms are simple, robust, and significantly reduce the overall run time of the process simulation of 3D MEMS-tunable laser devices. The description of the proposed etching algorithms will be presented. Numerical simulation results showing the performances of these algorithms will be given and analyzed
for realistic 3D MEMS devices and MEMS-tunable laser devices.

Keywords: advanced etching algorithms, domain decomposition, process simulation, MEMS-tunable lasers

1. INTRODUCTION

Vertical cavity semiconductor optical amplifiers (SOAs) represent a low-cost alternative to existing amplifier technologies. They could be used in fiber-optic communication systems such as metro and access networks (Garett et al. 2005; Garrett 2005).

The surface-normal operation of vertical-cavity (SOAs) gives rise to many advantages as high coupling efficiency to optical fibers, polarization insensitive gain, the potential to fabricate high fill-factor two-dimensional arrays, and the ability to test devices on wafer (Garett et al. 2005; Garrett 2005; Wu et al. 1995; Larsson, Massengale, and Harris 1996; Tayebati et al. 1998; Christensen et al. 1997; Rabinovich et al. 2003).

Understanding MEMS-tunable device fabrications is essential in optimizing design and enabling a more rapid production of mature devices. The 3D MEMS process simulation is a complex and challenging issue. The structures are geometrically complicated and inherently three-dimensional.

The main goal of this paper is to improve, enhance, extend from two dimensions (2D) to three dimensions, and optimize the existing geometrical etching algorithms for process simulation of 3D MEMS and 3D MEMS-tunable vertical cavity semiconductor optical amplifiers (MT-VCSOAs) as given in Figure 1. Most of commercial process simulators used worldwide are based on the industry standard two dimensional (2D) process simulator SUPREM-IV (Hansen and Deal 1993) or 3D process simulator FLOOPS (Law 1994). SUPREM-IV has been developed by Stanford University in California. In this paper, firstly, we are focusing on the improvements of the 2D etching algorithms used in SUPREM-IV. And, secondly, we did extend these 2D etching algorithms to 3D which was another challenging task.

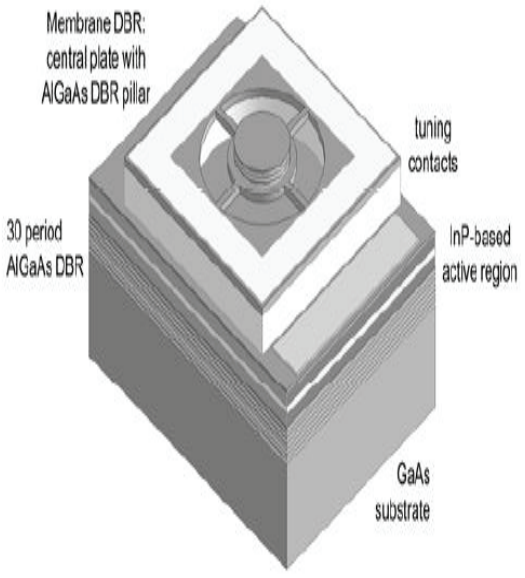


Figure 1: Schematic representation of MT-VCSOAs

We were, exactly, motivated by solving the following three difficulties:

1. In case of a 2D refined mesh using SUPREM-IV algorithms, a strange ripped nylon mesh is generated after etching as seen in Figure 2.
2. Extending the 2D geometrical etching algorithms of SUPREM-IV to 3D for process simulations of complex geometries of MEMS and MT-VCSOAs, while keeping the trade-off

between computational efficiency and the simulation's run time.

3. Creating an accurate and suitable 3D geometry and mesh for fabrication, mechanical, and electrical simulations of MT-VCSOAs and MEMS.

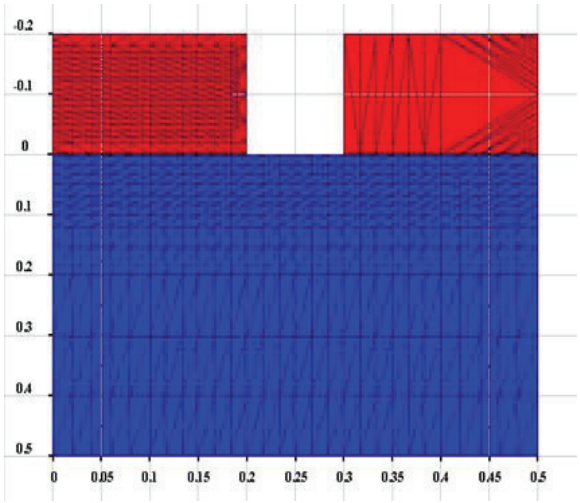


Figure 2: Ripped nylon mesh after SUPREM-IV etching

The traditional geometrical etching algorithms (Sahul 1996; Hansen and Deal 1993; Law 1994) are based on Boolean operations as region subtractions, region intersections, and region unions. They also incorporate additional supporting algorithms as, geometry validation, mesh quality control after etching, de-looping, removing holes, and making regions convex. In case of moving boundaries, as in physical etching, moving boundaries algorithms (Oldham et al. 1980; Glenn 2000) are used to calculate the etching rate (or the velocity of the boundary). In our case, we are not dealing with moving boundaries.

The main step in these etching algorithms is the Boolean subtraction operation. Our contribution is firstly to solve the difficulty 1. By applying boundary smoothing and re-meshing idea after subtraction operation as can be seen in Figure 3. Secondly, we used domain decomposition method (DDM) together with geometrical etching algorithms for the etching simulation of the complex 3D MEMS and MT-VCSOAs. The use of DDM is an excellent remedy to the difficulty in 2. The application of DDM significantly increases the efficiency and reduces the run time of etching simulation and the overall run time of the other process simulations including deposition, implantation, oxidation, and diffusion.

The DDM consists in decomposing the whole 3D structure into different 3D blocks. We could then apply etching algorithms to each block and then merge all the blocks to get the whole 3D structure. Advanced merging algorithms have also been used after etching.

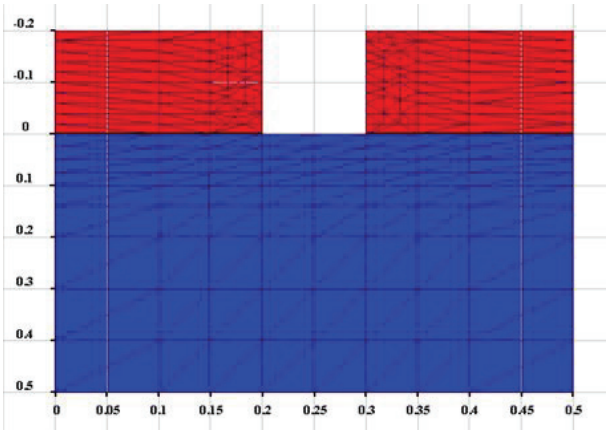


Figure 3: Boundary smoothing helped repair the ripped mesh

They are based on region unions and intersections. The DDM is also very convenient for parallelization.

This paper is organized as follows. The second section describes in details the new etching algorithms. The third section presents the 2D and 3D numerical results that validate qualitatively and quantitatively the introduced geometrical etching algorithms. The performance of these etching algorithms are also discussed and analyzed. A comparison of the obtained 3D numerical results with other results found in literature is also presented.

Section 4 contains conclusions, future work, and recommendations stemming from the work presented in this paper.

2. DESCRIPTION OF THE NEW ETCHING ALGORITHMS

Every 2D or 3D semiconductor structure could be defined in terms of geometry, surface mesh, and volume mesh. Geometry represents the gross outline of the structure and the different materials in the structure. We could look at the geometry as a set of regions. And each region is defined by a material. Surface or boundary mesh refers to a set of planar elements (points, edges, or triangles) whose union form the boundary of the geometry. The volume mesh refers to a set of volumetric elements (as tetrahedral elements) whose union defines the interior and the exterior of the whole device. The main focus of this paper is to smooth and re-mesh properly the surface of each region of the geometry after a geometrical etching. By doing so, we were able to solve the difficulty 1. in 2D and in 3D.

In this section, we will describe in details the algorithms developed and used to perform the geometrical etching simulations in 2D. The extension of these algorithms to 3D has also been achieved and used.

The geometrical etching algorithm first defines the geometry of the region to be etched. This region is then subtracted from each region of the whole structure. The pseudo-code of the algorithm is given in the Algorithm 1 given below. The main idea behind the etching

algorithm is as follows: the intersection **i** between the etch region **e** and a region **r** in the geometry is first computed (line 3 in Algorithm 1). If the etch region is fully enclosed within the geometry region, then, the etch is completed by defining the geometry region to have two boundaries (lines 4-5). The outer boundary is the existing boundary of the geometry region and the inner boundary is the etch region's boundary. Our new implementation of the etching algorithm consists in adding a new procedure **SmoothBoudary()** that will smooth properly these inner and outer boundaries according to some given data from the user. If the etch region completely contains the geometry region, then, the geometry region is completely deleted (lines 6-7). Otherwise, the intersection between the etch region and the geometry region is considered. The geometry region **r** is to be replaced with the regions that are outside the etch region **e**. Let **I** be the set of all the points of the intersection **i**. Let **R** be the set of all the points of the geometry region **r**. Let **E** be the set of all the points of the etch region **e** and let **P = R - I**. The Algorithm 1 below, then, works as follows.

First a point p0 of the region **r**'s boundary that is not part of the intersection (member of the set **P**) is found (line 13). Staring from this point p0, all the other subsequent points of **R** that are member of **P** are collected to form a new boundary **B** until the intersection with the etch region is found (lines 15-19). The collection continuous along the intersection from the set **I** until the region **r** is found (lines 20-25). The points of the region **r** in the set **R** are then collected until the first point is encountered (lines 26-29). This closed loop then generates a new boundary. This new boundary is then smoothed and re-meshed by calling the new procedure **SmoothBoundary()** and finally added to the geometry data structure. The procedure continues until all the points of the geometry region r's boundary that belong to **P** are part of the new boundary (lines 14, 31-32).

The pseudo-code of the new geometrical etching algorithm (Algorithm 1) is given in the following subsection.

2.1. Pseudo-code of the new etching algorithm

Algorithm 1: new geometrical etching algorithm
Procedure Etch (**e**, **r**)
Inputs: **e**: etch region, **r**: geometry region
Outputs: new smoothed boundaries and new regions

begin
1. **e**:= Etch Region;
2. **r**:= Geometry Region;
3. i:= RegionIntersection(**e,r**);
4. **if** (**i=e**) **then** (etch region inside region)
5. add boundary of **e** as internal boundary of **r**
SmoothBoundary(); (new improvement to etching algorithms)
6. **elseif** (**i=r**) **then** (geometry region inside etch region)
7. delete **r**;
else
8. **E**:= Points of **e**;
9. **R**:= Points of **r**;
10. **I**:= Points of **i**;
11. **P:= R - I**;
12. **B**:=**0**;
13. first:=p:= MEMBER(0,**P**);
14. **while** (**P** ≠ **0**) **do** (find the points that are outside)
15. INSERT(p,**B**);
16. np:=FIND(**R**, point next to p);

17. **if** (np ∈ **P**) **then**

18. DELETE(p,**P**);
19. p:=np;
else (find the points that belong to **i**)
20. np:=FIND(**I**, point on edge from p to np);
21. INSERT(np,**B**);
22. **repeat**
23. np:=FIND(**I**, point next to np);
24. INSERT(np,**B**);
25. **until** np is on **r**;
(loop back to get all the other points of **r**)
26. p=FIND(**R**,point next to np);

27. **while** (p ≠ first) **do**

28. INSERT(p,**B**);
29. DELETE(p,**P**)
30. add new region formed by points in the set **B**;
SmoothBoundary(); (new improvement to etching algorithms)
31. **B**:=0;
32. first:=p:=MEMBER(0,**P**);
33. delete region **r**;
SmoothAllBoundaries(); (new improvement to etching algorithms).
This procedure is optional. It could be called to smooth globally all the boundaries.
end

We should note that if the boundary smoothing procedures **SmoothBoundary()** and **SmoothAllBoundaries()** are called outside the geometrical etching Algorithm 1, then, the analysis of the complexity in time of the Algorithm 1 shows that the Algorithm 1 is O(**N+K**). The number **N** represents the number of points of the region **r**. And **K** is the number of the points of the etch region **e**. Then, the complexity of the Algorithm 1 has the advantage to be linear.

2.2. Pseudo-code of the new smoothing boundary procedures

During etching algorithm, we call the smoothing boundary procedure **SmoothBoundary()** to smooth and re-mesh the new boundary that we just get. This procedure will split up some or all the long edges of the new boundary according to some criteria and to the user defined parameter α. The shape of the geometry is not altered if the long edges are split up. However, what constitutes a long edge? In this algorithm, long edges are judged according to the original perimeter of the boundary under hand. If an edge of length l_s is bigger than the user specified percentage α of the perimeter or bigger than 3 times the smallest edge of length l_min, then, this edge will be split up. The algorithm 2, given bellow, first finds edges in the boundary and checks to see if it is bigger than α times perimeter or bigger than 3 times l_min. If it is, then, the procedure **SplitEdge()** is called to split up the edge.

Algorithm 2: Boundary Smoothing

Procedure SmoothBoundary ()

Inputs: **B**: region boundary to be smoothed according to user α: user defined parameter used as a smoothing factor

Outputs: new smoothed and re-meshed boundary **B**

begin

1. p:=perimeter of boundary **B**;
2. l_min=length of the smallest edge of boundary **B**;
3. **for** each edge **e** of **B** **do**
4. l_s = length of **e**;
5. **if** (l_s > 3 times l_min or l_s > α times p) **then**
6. **SplitEdge**(**e**);

end

The etching and smoothing algorithms rely on many other geometry utility algorithms such as bulk and boundary mesh generation algorithms, algorithms to remove holes and make all the regions convex, algorithms assuring boundary orientation in counter-clockwise manner and other similar utilities.

3. 3D NUMERICAL RESULTS AND ANALYSIS

Using the proposed new geometrical etching algorithms, we were able to solve the 3 difficulties 1., 2. and 3. described above. On the other hand, we were successful to etch efficiently all the complicated 3D structures under hand. Figure 4 shows the mesh of a realistic MEMS Tunable VCSOAs after improved etching. This Figure 4 shows also the mesh quality after improved etching. The Figure 5 shows the structure of a realistic Radio Frequency MEMS device before etching.

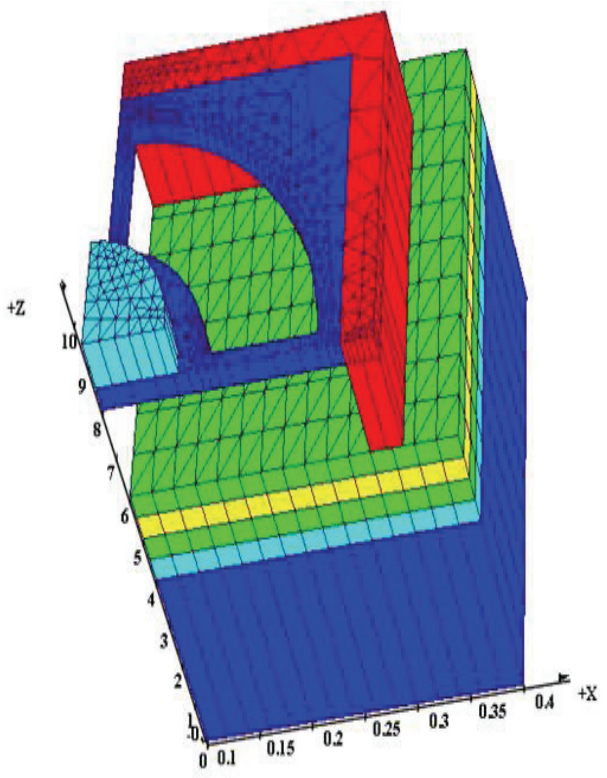


Figure 4: 3D MEMS Tunable VCSOAs after improved etching

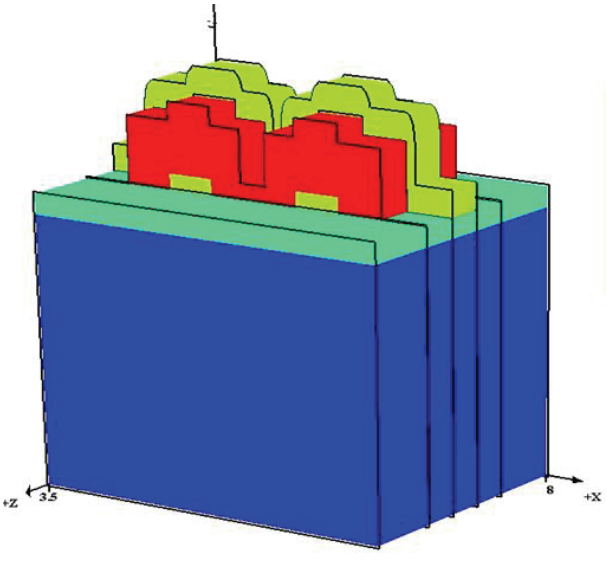


Figure 5: Radio frequency MEMS before 3D etching

The Figure 6 shows the mesh and the geometry of a Radio Frequency MEMS device after improved etching. The quality of the mesh and the geometry shown in all these Figures (Figure 3 to 6) do validate the excellent

performance and the qualitative and the quantitative behavior of the proposed etching algorithms. These performances are comparable with those find in (Glenn 2000) for different 3D structures. The size of our obtained meshes is even smaller and better optimized than those in (Glenn 2000). This is due to the use of the domain decomposition method and the local mesh adaptation.

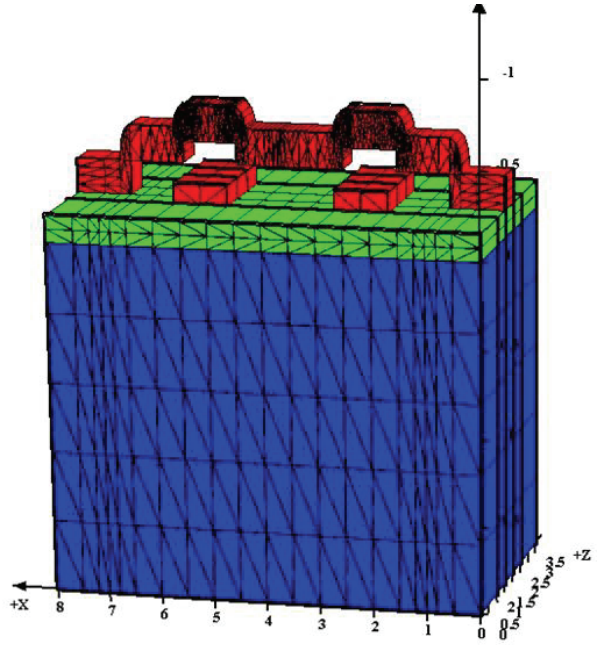


Figure 6: Radio frequency MEMS after 3D improved etching

4. CONCLUSION

In conclusion, by using these new geometrical etching algorithms, it is possible to generate accurately complex geometries and meshes for mechanical and electrical simulations of 3D MEMS and MEMS-tunable laser devices. Parallelization of these algorithms could be investigated in future work.

REFERENCES

Christensen, G.L., Tran, A.T.T., Zhu, Z.H., Lo, Y.H., Hong, M., Mannaerts, J.P., Bhat, R., 1997. Half Symetric Cavity Tunable Microelectromechanical VCSEL with Single Mode. *IEEE Photon Technology Letter*, Vol.10, 1679-1681.

Garett, D.C., Qi, C., Chaung-Yeng, C., Shaomin, W., Chad, S.W., Noel, C.M., Bowers, J.E., 2005. MEMS-Tunable Vertical Cavity SOAs. *IEEE Journal of quantum electronics*, Vol. 41, No 3, March.

Garett, D.C., 2005. *MEMS-Tunable Vertical Cavity SOAs*. P.H.D. Thesis, University of California, Santa Barbara, CA, USA.

Garett, D.C., 2000. *A Three-Dimensional Front tracking Algorithm for Etching and Deposition processes*. P.H.D. Thesis, State University of New York, NY, USA.

Hansen, S.E., Deal, M.D., 1993. *User's Manual of SUPREM-IV.GS: Two dimentional Process Simulation for Silicon and Galllium Arsenide*. Integrated Circuit Laboratory, Stanford Unievrsity, Stanford, California, USA.

Larson, M.C., Massengale, A.R., Harris, J.S., 1996. Continuously Tunable Micromachined Vertical-Cavity Surface Emitting Laser. *Electron Letter*, Vol. 32,330-332.

Law, M.E., 1994. *User's Manual of FLOOPS: Florida Objected Oriented Process Simulator Manual*. Dept. of Elect. Eng., University of Florida, Florida, USA.

Oldham, W.G., Neureuther, A.R., Sung, C., Reynolds, J.L., Nandgaonakar, S.N., 1980. A General Simulator for VLSI Lithograpgy and Eteching Process: PartII-Application to Deposition and Etching. *IEEE Trans Electron Devices*, Vol.ED-27, 1455-1459.

Rabinovich, W.S., Stievater, T.H., Papanicolaou, N.A., Katzer, D.S., Goetz, P.G., 2003. Half Symetric Cavity Tunable Microelectromechanical VCSEL with Single Mode. *IEEE Photon Technology Letter*, Vol.10, 1679-1681.

Sahul, Z.H., 1996. *Grid and Geometry Servers for Semiconductor Process and Simulation*. P.H.D. Thesis, Stanford University, Stanford, CA, USA.

Tayebati, P., Wang, P.D., Vakhshoori, D., Lu, C.C., Azimi, M., Sacks, R.N., 1998. Half Symetric Cavity Tunable Microelectromechanical VCSEL with Single Mode. *IEEE Photon Technology Letter*, Vol.10, 1679-1681.

Wu, M.S., Vail, E.C., Li, G.S., Yuen, W., Chang-Hasnain, C.J., 1995. Tunable Micromachined Vertical Cavity Surface Emitting Laser. *Electron Letter*, Vol. 31,1671-1672.

AUTHORS BIOGRAPHY

Abderrazzak El Boukili received both the PhD degree in Applied Mathematics in 1995, and the MSc degree in Numerical Analysis, Scientific Computing and Nonlinear Analysis in 1991 at Pierre et Marie Curie University in Paris-France. He received the BSc degree in Applied Mathematics and Computer Science at Picardie University in Amiens-France. In 1996 he had an industrial Post-Doctoral position at Thomson-LCR company in Orsay-France where he worked as software engineer on Drift-Diffusion model to simulate hetero-junction bipolar transistors for radar applications. In 1997, he had European Post-Doctoral position at University of Pavia-Italy where he worked as research engineer on software development for simulation and modeling of quantum effects in hetero-junction bipolar transistors for mobile phones and high frequency applications. In 2000, he was Assistant Professor and Research Engineer at the University of Ottawa-Canada.

Through 2001-2002 he was working at Silvaco Software Inc. in Santa Clara, California-USA as Senior Software Developer on mathematical modeling and simulations of vertical cavity surface emitting lasers. Between 2002-2008, he was working at Crosslight Software Inc. in Vancouver-Canada as Senior Software Developer on 3D Process simulation and Modeling. Since Fall 2008, he is working as Assistant Professor of Applied Mathematics at Al Akhawayn University in Ifrane-Morocco. His main research interests are in industrial TCAD software development for simulations and modeling of opto-electronic devices and processes.
http://www.aui.ma/perosnal/~A.Elboukili.

MODELING OF WATER RETENTION IN SUBSTRATES USED IN PRODUCTION OF TOMATO "SWEET GRAPE"

[1]Honorato C. Pacco, [2]Delvio Sandri, [1]Sebastian Avelino Neto, [1]Marco A. Amaral Jr. [1]Ananda H. N. Cunha.

[1]Faculty of Agricultural Engineering, UnUCET-UEG, Anapolis-GO, honor122@yahoo.com.br

[2]College of Agriculture and Veterinary Medicine UNB, Brasilia-GO.

honor122@yahoo.com.br

ABSTRACT

The objective of this work was the modeling of the water retention curves in three different substrates used in production of tomato variety "Sweet grape". The experiments were carried out within the existing greenhouse on the campus of UEG-UnUCET. Is saturated with water three types of substrates (sand, 40% with a commercial substrate, natural coconut fiber and commercial substrate by 80% pine bark and 20% coconut fiber) for 48 hours contained in two plastic pots for each substrate, was instrumented with three pressure sensors in each plastic pots for 59 days, where pressure measurements was performed using a digital tensiometer. Was constructed wetlands curves and pressure for each substrate in our study, each reading pressure current humidity existing on the substrates. As a result it was observed that the substrate with coconut fibers has the best ability to retain water in the assay compared to sand substrates and commercial. The model allowed us to determine the choice of substrate with improved ability to retain water for tomato production and reduce leaching loss of macro and micro nutrients in a hydroponic culture, and also allows the automation system.

Keywords: Modeling, tensiometers, Digital tensiometer, water content.

1. INTRODUCTION

Modeling is important in determining soil moisture, allows the correct irrigation crop plants, as well as to identify soil management techniques, more appropriate, enables the automation system. Water is a key factor in crop production. Their lack or an excess affects decisively in plant development and its rational management is very important in agricultural production, RABELLO et al. (2005).

The humidity sensors allow monitoring of soil water, avoiding excess irrigation and assisting in the rational use of water, control soil salinization, and erosion bioremediation, BRASEQ (2011).

The tensiometer is a device used to measure the pressure at which water is retained by the soil particles and can help control the irrigation.

The pressure of the water retained in the soil is correlated with the moisture existing in the profile forming a curve of water retention in the soil which can be determined in the laboratory.

The occurrence of water deficit undermines the productivity and quality of vegetables produced. Thus, the vegetable, except in areas or during seasons with regular distribution of rainfall is usually performed under irrigation.

Adequate irrigation minimize environmental impacts, as a rule, allow to reduce the expenditure on water and energy, nutrient losses by leaching and the incidence of diseases, with consequent reduction in the use of agrochemicals.

The objective was to model the water retention curves in three different substrates distributed in plastic pots for growing tomato variety "Sweet grape", using pressure sensors to record the data.

2. MATERIALS AND METHODS

The experiments were carried out within the existing greenhouse on the campus of Universidade Estadual de Goiás-UEG UnUCET in the months from May to July 2011. We have worked with three types of substrates is the more common vegetable crop (A: sand, with 40% a commercial substrate, Fc: coconut natural fiber and, S: commercial substrate composed of 80% pine bark and 20% coconut fiber) which were divided into two plastic pots (greater diameter, 24 cm, and base diameter 14 cm and a height of 22.5 cm) for each substrate.

Is saturated with water for 48 hours and instrumented with three pressure sensors (tensiometers puncture Fig. 1) in each plastic pots for 59 days, where pressure measurements was performed using digital tensiometer (Tensimeter – Fig. 2). On each day of readings took samples of each type of substrate for its determination of moisture content by placing in an oven at 60 ° C

temperature until constant weight the samples contained in the crucibles by triplicate.

Then with the results of tension readings and the results of moisture, was performed to model the moisture curves (%) and tension (kPa) each substrate in a study used in the production of tomatoes of the variety “Sweet grape" inside the greenhouse, using the hydroponic system and drip with water and wastewater natural water. Obtaining the graphical humidity and pressure for each type of substrate.

This type of methodology is a procedure generally quite simple and extremely important because it helps in monitoring the dynamics of soil water, since it evaluates the potential of existing water.

The modeling was performed using the software Microsoft Excel 2010 Microsoft Office, and Matrix Laboratory software version MatLab version 7.10-R2010a, and software Statistic version 8, yielding equations of the models with each Software to compare the results.

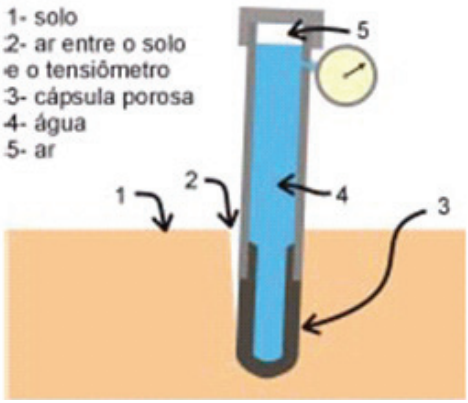


Figure 1: Tensiometer sensor. Source: CALBO (2006).

Figure 2: Digital tensiometer

3. RESULTS AND DISCUSSIONS

In this work, from coconut fiber (Fc) showed the best ability in water retention as observed in Figure 3, and validated the work of Lima et al. (2011) using natural coconut fiber as substrate for the production of tomato in hydroponic systems, and Rubio et al (2011) who also worked with coconut fiber in Sweet pepper production in substrate in response to salinity, nutrient solution management and training system. The commercial substrate (S) also showed better water retention about sand (A) the substrate.

Obtained is a linear model for the curves of moisture (%) and tension (kPa) for three substrates under study (Figure 3), where it is observed that the substrate consists of coconut fibers (Fc), has the best ability retention of water in the experiments, followed by a commercial substrate (S), indicating that could be used for the cultivation of vegetables reducing water loss, reduce leaching of nutrients loss, controlling the time of irrigation and reasonable amount of water used for the proper development of the plant, to obtain good yield and quality in the products of culture. It can be observed in the curve of Figure 3, the water evaporated very easily in the substrate composed of 60% sand and 40% of pine bark.

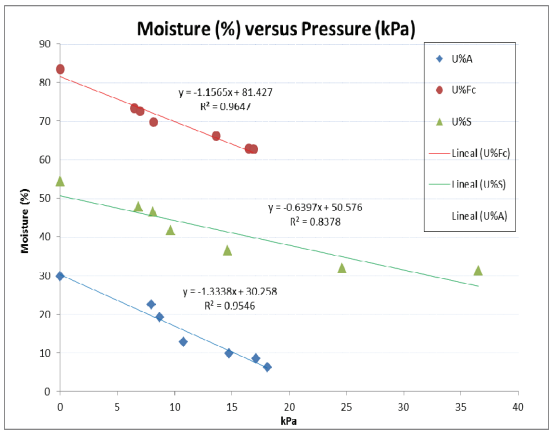
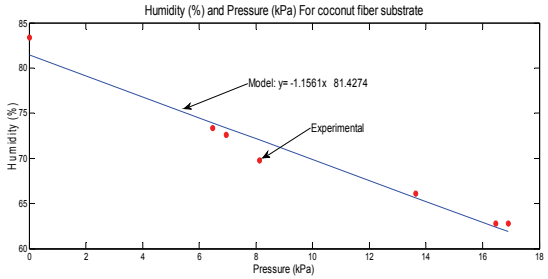
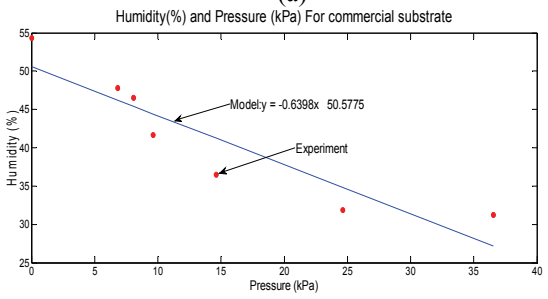


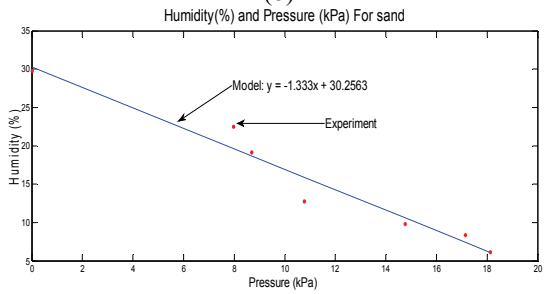
Figure 3. Modeling for Microsoft Excel 2010 of water retention curves for different substrates (A: sand: Fc: coconut fiber, S: commercial substrate) used in the cultivation of tomatoes "Sweet grape" in a greenhouse.



(a)



(b)



(c)

Figure 4: Modeling by MatLab version R2010a of water retention curves in different substrates ((a) Fc: coconut fiber, (b) S: commercial substrate and (c): Sand) used in the cultivation of vegetables in a greenhouse.

Comparing models obtained using the software Microsoftware Excel 2010 and MatLab version R2010a, the equations of the model for each type of substrate being studied are the same as shown in Figures 3 and 4 (a, b, c).

When the mathematical modeling was performed using Statistic software version 8 with the data of humidity (%) and stress (kPa) of the middle support (substrate) gave the following results:

For the medium support coconut fiber (Figure 5) was obtained in the modeling with an equation different in coefficient values (y = 78 741 - 0.9527x) compared with those obtained by MatLab software version 7.10 - R2010a (y = 81.4274 - 1.1561x) and software Microsoft Office Excel 2010 (y =81.427 - 1.1565x), where the coefficient values are almost equal to those obtained by MatLab software and Microsoft Office Excel. The best curve fit was found with the correlation coefficient when modeling was using the software Statistic version 8 (R^2 0.9793) regarding the Microsoft Office Excel (R^2 0.9647).

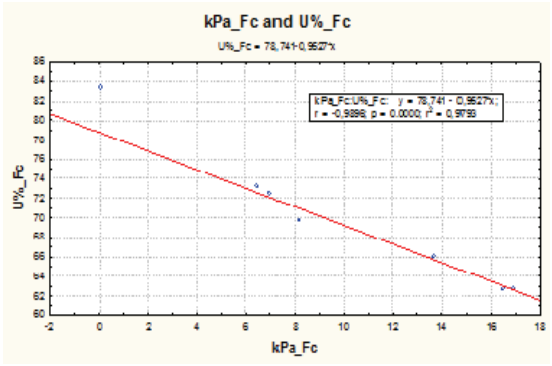


Figure 5. Modeling by Software Statistic version 8, of water retention for coconut fiber substrate (Fc), used in the cultivation of vegetables in a greenhouse.

For the commercial substrate support medium (Figure 6) was obtained with an equation modeling also different in coefficient values (y = 47.0270 - 0.4958x) compared with those obtained by MatLab (y = 50.5775 - 0.6398x) and software Microsoft Office Excel (y =50.576 - 0.6397x), where also the coefficient values are almost equal to those obtained by MatLab software and Microsoft office Excel. And the best curve fit was found with the correlation coefficient when using the modeling software was Microsoft Office Excel (R^2 0.8374) regarding the software Statistic version 8 (R^2 0.7802).

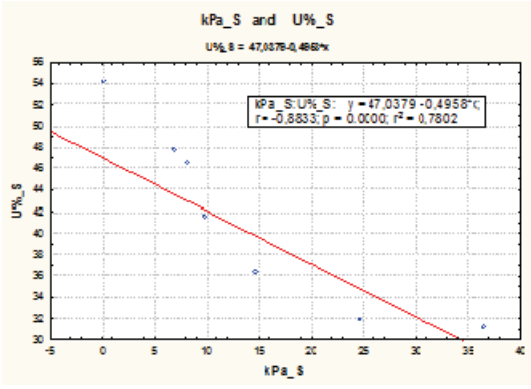


Figure 6. Modeling by Software Statistic version 8, of water retention for commercial substrate (S), used in the cultivation of vegetables in a greenhouse.

To support medium sand (Figure 7) was obtained in the modeling equation, with values nearly equal in coefficient (y = 30.1823 - 1.3292x) compared with those obtained by MatLab software version 7.10 - R2010a (y = 30.2563 - 1333) and software Microsoft Office Excel (y = 30258 - 1.3338). The best curve fit was found with the correlation coefficient when using the modeling software was Microsoft Office Excel (R^2 0.9546) regarding the software Statistic version 8 (R^2 0.9084).

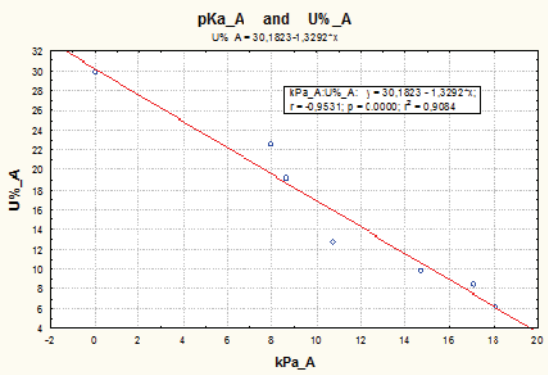


Figure 7. Modeling by Software Statistic version 8, of water retention for sand (A), used in the cultivation of vegetables in a greenhouse.

In this research the substrate coconut fiber (Fc) used as support showed the best ability in retaining water as shown in Figure 8, and validated by Lima et al. (2011) using natural coconut fiber as a substrate for the production of tomatoes hydroponically, and RUBIO et al (2011) who also worked with coconut fiber in sweet pepper production in substrate in response to salinity, nutrient solution management and training system. The commercial substrate (S) also showed better water retention respect to the substrate mixed sand (A) used as a support.

As can be observed in the curve of Figure 8, the water evaporates very easily and quickly on the substrate composed of 60% sand and 40% pine bark or commercial substrate. The model that best fit to the curves for the behavior of the substrates used as the base for retention of water, the polynomial model, yielding the following correlation coefficients for coconut fiber with R^2 0.8716, a commercial substrate for (pine bark) with R^2 0.996 and sand mixed value for R^2 was 0.8833.

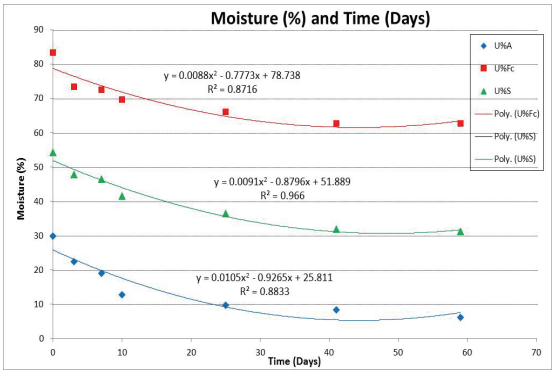


Figure 8. Curves moisture (%) versus time (days) for different substrates (A sand + 40% commercial substrate; Fc: coconut fiber and S: commercial substrate) contained in plastic pots used as supporting medium in the culture vegetables in a greenhouse.

4. CONCLUSIONS

It can be concluded that modeling has determined that the substrate with coconut fiber had the best ability of water retention time compared to the other two substrates in a study for the production of greenhouse tomato variety "Sweet grape" by hydroponic. Controlling the system set point pressure for each crop of vegetables, we can obtain a good yield and good quality end product.

The modeling shows that the software allowed Microsoft Office Excel 2010 version best fit the curves of water retention in the substrates used for production of vegetables in a greenhouse. Water evaporates quickly in sand mixed compared with coconut fiber and commercial substrate as support used to the production of vegetables.

ACKNOWLEDGMENTS

To the National Post-Doctoral - PNPD Capes – Brazil

REFERENCES

Braseq. 2011. Sustainability in agricultural production. Boletin BrasEq. August 2011.

Calbo, A. G. Irrigas: gaseous system of irrigation control. In: State of water in soil and plant, 2006. Available at: <http://www.cnph.embrapa.br/novidade/prelancamento/irrigas/irrigas.html>.

Lima, A.A; Alvarenga, M.A.R; Rodrigues, L; Chitarra, A.B. Productivity and quality of tomatoes produced on substrates by applying and humic acids. Journal of the Brazilian Association for Horticultural Science. 29: 269-274 p. Brazil, 2011.

Rabello, L. M; Vaz, C. M. P; Torre Neto, A. Capacitive sensor for probing the moisture in the soil profile. San Carlos: EMBRAPA, 2005. 2 p. (Technical Communication, 71).

Rubio J.S; Pereira W.E; Garcia-Sanchez F; Murillo L; Garcia A.L; Martinez V. 2011. Sweet pepper production in substrate in response to salinity, nutrient solution management and training system. Brazilian Horticulture 29: 275-281.

AUTHORS BIOGRAPHY

Honorato Ccalli Pacco:
Food Engineering from the National University "Jorge Basadre Grohmann" Tacna Peru - UNJBG, Master in Food Engineering from the State University of Campinas UNICAMP - Brazil, PhD in Agricultural Engineering from the State University of Campinas UNICAMP - Brazil. Research Scientist and Professor at the State University of Goiás - Brazil. Fields of action, Food Engineering and Agricultural Engineering.

CROSS-DOCKING TRANSSHIPMENT PROBLEM APPROACHED BY NON LINEAR PROGRAMMING AND SIMULATION ANALYSIS

Aiello Giuseppe[(a)], Enea Mario[(b)], Muriana Cinzia[(c)]

[(a) (b) (c)]Dipartimento di Ingegneria Chimica, Gestionale, Informatica e Meccanica, Università degli studi di Palermo, Edificio n°8, 90128 Palermo – Italy

[(a)]giuseppe.aiello03@unipa.it, [(b)]mario.enea@unipa.it, [(c)]cinzia.muriana@unipa.it

ABSTRACT

The need for fast product delivery causes the attention of supply chain is addressed to strategies able to optimize the distribution process. In this field the cross-docking seems to be an efficient strategy which makes possible to reduce or eliminate the storage phase by meeting customer demand. In this paper a transshipment problem for cross-docking strategy is considered by means of a deterministic model studied through the non linear programming technique. The solution found allows to determine the optimal quantities to ship, the number of routes activated and the optimal truck number when the constraint on truck capacity is enforced. The influence of the demand fluctuation is also addressed through a simulation tool representing the cross-docking system. Finally the comparison between the cross-docking strategy and the direct delivery one is considered in terms of cost efficiency and trucks utilization.

Keywords: cross-docking, simulation, non linear programming model

1. INTRODUCTION

The increasing customer demand for fast product delivery leads business managers to improve the supply chain especially with reference to the distribution strategy. The minimization of the total cost of products delivered is related to the possibility of implementing an efficient control of the physical flow of products transferred. The specific importance of the distribution process is due to the fact that it can affect up to a 30% of an item price (Apte and Viswanathan, 2000). In order to implement new distribution strategies able to properly handle their products, industries nowadays look at the cross-docking as an efficient distribution strategy. Cross-docking can be defined as a continuous transportation where products are transshipped from the supplier, collected in the cross-dock, then are aggregated on the basis of their destination and finally shipped to the destination. The main objective of a cross-docking process is to avoid intermediate storage phases, thus eliminating inventory holding cost and labour intensive picking operations (Vahdani and Zandieh, 2010). Cross-docking attempts to lessen or even eliminate such burdens by reducing warehouses to purely trans-shipment centers where receiving and shipping are its only functions (Li et al., 2004). In other words, in the cross-docking network, the warehouses, as cross-docks, are transformed from inventory repositories to points of delivery, consolidation and pick-up (Chen et al., 2006). This allows to achieve a second objective consisting in the reduction of product cycle time (Li et al., 2009a). As observed by Yu and Egbelu (2008), the cross-docking systems operate best for companies which distribute a large amount of items and/or serve a large number of stores in a short time. With respect to the traditional warehouse systems the cross-docking allows to increase the inventory turnover, reduce the inventory level and operational costs and improve the customer responsiveness. As reported in Li et al. (2008), not all products are suitable for cross-docking and anyway the selection of the distribution strategy depends upon a number of factors such as product volume, product value, product life cycle, facility space constraint, stockout cost, etc. In particular the stockout costs are of greater importance in defining the products must be managed in the system due to the fact that generally in a cross-dock there is not inventory. For this reason usually cross-docking is suitable for fast moving items with stable demand, such as perishable products and agricultural products (Apte and Viswanathan, 2000). Groceries and agricultural products are also characterized by low stockout cost and for this reason they could be effectively managed with the cross-dock system. These products must be fast delivered to customer in order to preserve their freshness and because of the short period of circulation.

The implementation of a cross-docking system relates to the need to take decisions at different levels such as operational, tactical, and strategic levels. At the strategic point of view decisions address the determination of the optimal number of cross-dock and the number of trucks which must be disposed in the network considered. In such context Musa et al., (2010), proposed a Heuristic Algorithm to minimize the total transportation cost when each arch of the network can be satisfied through a direct link or one cross-dock

center. Charkhgard and Tabar (2011), faced the cross-docking problem in the case of determination of optimal truck capacity. They formulated a mixed integer non-linear programming model and solved it through a heuristic Algorithm such as the simulated annealing.
As reported in Agustina et al., (2010), the tactical level mainly relates to the determination of the best layout in the cross-docking. In this field Gue (1999), proposed a material flow model in order to minimize the flow inside the cross-dock, while Heragu (2005), proposed a model to simultaneously optimize the areas put in for storage, forward and cross-docking and the product allocation in order to minimize the total material handling cost.
The operational level is grouped in five research areas (Agustina et al., 2010) namely the scheduling problem, the transshipment problem, the dock door assignment problem, the vehicle routing problem and the product allocation problem. The scheduling problem usually aims at determining the optimal sequence of inbound and outbound trucks which in turn will minimize the makespan and then the total cost related to the products and trucks management. Larbi et al., (2007), studied the scheduling of transshipment operation in order to minimize the total inventory cost and truck replacement cost. He used a dynamic programming model and solved the model through a heuristic method. Li et al., (2009b), developed a truck scheduling with dock door assignment problem solved through a Genetic Algorithm. Boloori Arabani et al., (2011), proposed a multiobjective approach in order to minimize the makespan and the total lateness by means of a Genetic Algorithm. The transshipment problem concerns the determination of how much to ship, between which locations, on which routes and at what times. In this research field Lim et al., (2005), formulated a problem considering the inventory, the capacity of cross-docking and the time window constraints. Further studies have been conducted by Miao et al., (2008), who considered the transshipment problem where the transportations have fixed schedule and shipping and delivery can be only executed within time windows. The model's objective is minimizing the shipping and inventory holding cost by means of a Genetic Algorithm. The assignment problem deals with the proper assignment of inbound and outbound trucks within origins and destinations respectively. The first work in this field was realized by Tsui and Chang (1992). They developed a model to determine the assignment of receiving doors to the origins and shipping doors to the destinations. The objective of the model is to minimize the travel distance of the forklifts. Lim et al., (2006), considered an assignment problem with capacity of cross-dock and time window constraints. The objective of the model is to minimize the total shipping distance of transferring cargo from inbound to outbound dock. They solved the problem by means of a Genetic Algorithm.
In this paper the cost tradeoff between direct shipping and cross-docking systems is investigated referring to a numerical application solved by means of an Integer Non Linear Programming (INLP) approach. This approach is usually employed to simply represent complex systems and solve NP-hard problems as in the case of the cross-docking ones. In such context it is generally employed to determine a set of best candidate solutions that can be subsequently studied under disturbance conditions. In this study the optimal solution provided by the INLP approach is further tested by means of a post-optimality analysis performed through a simulation model in order to take into account the effects of the uncertainty of the demand on the Total Cost function and on the Utilization Coefficient of trucks. The robustness of the solution has therefore been evaluated.
Simulation can be defined as the process of designing a model of a real system, implementing the model as a computer program, and conducting experiments with the model for the purpose of understanding the behavior of the system, or evaluating strategies for the operation of the system (Smith, 1999). The simulation model takes the form of a set of assumptions concerning the operations of the system. These assumptions are expressed in mathematical, logical, and symbolic relationships between the entities, or objects of interest, of the system. Some of these assumptions can comprise those situations in which one or more inputs are random variables of the model and then they represent uncertainty elements that affect the system performances. The use of simulation allows to incorporate the randomness of such elements in the system, by representing the randomness through properly identified probability distributions arisen from the study of data related to the real processes of the system. In this case the outputs provided by the model can be considered only as estimates of the true characteristic of the model.
On the other hand there are some limitations affecting the use of simulation models and that must be taken into consideration when performing a simulation. First of all the real system could be very complex and several decisions must be taken in order to decide what details must be included in the model. Thus some details will be omitted and their effects lost or aggregated into other variables that are included in the model. In every case this representation will lead some inaccuracy sources. Another issue is the availability of data needed to describe the system behavior. In fact it is a common experience to describe a system by having few data. This issue must be considered prior to design the model in order to minimize its impact on the model itself.
The simulation-based approach played a significant role in analyzing performance at cross-docking centers. There are several studies where simulation modeled a cross-docking system. For example references to this application can be found in Rohrer (1995), that studied the importance of hardware and software system in the cross-docking systems. He describes how simulation helps to ensure success in cross-docking systems by determining optimal hardware configuration and

software control, as well as establishing failure strategies before cross-docking problems are encountered. Magableh and Rossetti (2005), studied a generic cross-docking facility with the aim at analyzing operational risks associated to individual cross-docking facility within a company's distribution network. Aickelin and Adewunmi (2008), proposed an assignment problem solved with the combined use of simulation and Memetic Algorithm. Liu and Takakuwa (2009), focused on the personnel planning of materials handling at a real cross-docking center in order to minimize the total personnel expenses at a cross-docking center. The approach employed includes the adoption of a simulation model together with integer programming. Arnaout et al., (2010), proposed a cross-docking simulation model in which the orders size and the due dates are represented by stochastic variables. Liu (2010), proposed a discrete event simulation model for non-automated cross-docking center with the aim of providing a decision making tool for logistic managers. Liu and Takakuwa (2010), studied the just-in-time shipments in a non-automated retail-cross-docking center. They proposed a simulation-based approach to analyze the material handling operation.

As you can see from the previous mentioned literature the simulation model has recently adopted to study the cross-docking problem under the view point of the variable affecting its performances. One of the variables that are usually poorly considered in cross-docking simulation is the variation of demand. The reason of the poor use of simulation tool in this field is due to the fact that usually the cross-docking problem is faced with reference to the deterministic behavior of the system modeled by considering that the customer demand is related to products having a stable demand and no fluctuations are considered. Furthermore it must be taken into account that the complexity of simulation models increases considerably as the number of suppliers, cross-docks and products increase as well. In fact as the number of nodes in the network increases the number of arches to be considered increase as number of suppliers* number of cross-docks + number of cross-docks*number of clients. In this paper a simulation tool has been employed to study the transshipment problem with cross-docking facilities where the hypothesis of deterministic behavior of the demand is relaxed and it is modeled as a stochastic variable. The comparison of the cross-docking strategy with the direct delivery one in terms of Average Total Cost and Trucks Utilization has been carried out by comparing the results of the two corresponding simulation models.

The remainder of the paper is hence organized as follows: Section 2 deals with the proposed methodology by presenting the two INLP models and the corresponding simulation models, thus the experimental application is showed in Section 3 and the main results are summarized. Finally the Section 4 reports the conclusions.

2. THE PROPOSED METHODOLOGY

The aim of the paper is to evaluate the performance of a traditional direct shipping transportation system and a cross-docking system, taking into account the effects of the uncertainty in the customers' demand by means of a simulation approach. However, solving an optimization problem by means of a simulation approach requires an excessive computational effort, therefore it is generally preferred to apply a two steps optimization procedure, where the best candidate solutions are determined first by means of a simplified model and a simulation approach is subsequently applied to select among the pre-determined best solution candidates.

The methodology here proposed, hence, consists in formulating a deterministic INLP model to determine the optimal solution of each problem neglecting the effects of uncertainty, and subsequently to perform a post-optimality simulation analysis.

2.1. INLP model for cross-docking transshipment problem

The INLP model for the cross-docking system has been formulated under the following notations:

i, number of suppliers
j, number of clients
n, number of cross-docks
k, number of products
C, maximum truck capacity
p_{jk}, demand of product *k* for the customer *j*
a_{ik}, availability of product *k* at the supplier *i*
d_{ij}, distance between the source *i* and the destination *j*
c_k, variable transport cost of the product *k* per unit distance from the origin (supplier or cross-dock) to the destination (cross-dock or client)
f_{ij}, fixed transport cost between the source *i* and the destination *j*. Such cost is proportional to the number of trucks routed between the source and the destination.
$M = 100{,}000$, upper bound
$N = 100{,}000$, upper bound

The following assumptions have been considered:

1. The suppliers and the cross-docks have very high capacity in order to ensure that products are always available and no stockout will occur.
2. Each supplier manufactures a single product.
3. The customer demand for each product is deterministic and constant and never exceeds truck capacity.
4. Trucks are always available and trucks have the same capacity.
5. Trucks have single destinations in a tour. They do not go from one destination node in the network to another but only from a origin (supplier or cross-dock) to a destination (cross-dock or client).
6. Trucks capacity and demand are expressed in terms of product units.

Decision variables:

q_{ikn}, transported quantity of product *k* from *i* to *n*
q_{knj}, transported quantity of product *k* from *n* to *j*
N_{in}, non negative variable representing the number of trucks on arc *i-n*
N_{nj}, non negative variable representing the number of trucks on arc *n-j*
x_{in} and x_{nj},binary variables

$$x_{in} = \begin{cases} 1, if\ the\ route\ i-n\ is\ active \\ 0 \qquad\qquad otherwise \end{cases}$$

$$x_{nj} = \begin{cases} 1, if\ the\ route\ n-j\ is\ active \\ 0 \qquad\qquad otherwise \end{cases}$$

The objective function:

$$\min z = \sum_{i,k,n} q_{ikn} * c_k * d_{in} + \sum_{k,n,j} q_{knj} * c_k * d_{nj} + \sum_{i,n} f_{in} * x_{in} * N_{in} + \sum_{n,j} f_{nj} * x_{nj} * N_{nj} \quad (1)$$

s.t.:

$$\sum_i a_{ik} \geq \sum_j p_{jk} \qquad \forall\, k \quad (2)$$

$$\sum_n q_{ikn}\, x_{in} \leq a_{ik} \qquad \forall\, i,k \quad (3)$$

$$\sum_{i,n} q_{ikn}\, x_{in} = \sum_j p_{jk} \qquad \forall\, k \quad (4)$$

$$\sum_i q_{ikn} x_{in} = \sum_j q_{knj}\, x_{nj} \qquad \forall n,k \quad (5)$$

$$\sum_n q_{knj} x_{nj} = p_{jk} \qquad \forall j,k \quad (6)$$

$$\sum_{i,n,k} q_{ikn} = \sum_{k,n,j} q_{knj} \quad (7)$$

$$\sum_k q_{ikn} * x_{in} \leq C * N_{in} \qquad \forall i,n \quad (8)$$

$$\sum_k q_{knj} * x_{nj} \leq C * N_{nj} \qquad \forall n,j \quad (9)$$

$$x_{in} * M \geq \sum_k q_{ikn} \qquad \forall i,n \quad (10)$$

$$x_{nj} * M \geq \sum_k q_{knj} \qquad \forall j,n \quad (11)$$

$$x_{in} * N \geq N_{in} \qquad \forall i,n \quad (12)$$

$$x_{nj} * N \geq N_{nj} \qquad \forall j,n \quad (13)$$

All the variables must be non negative.

The objective function (1) is formulated to minimize the total transshipment cost consisting in both variable and fixed costs. The variable costs are proportional to the distance traveled and the quantity of products shipped, while the fixed costs are proportional to the number of trucks routed.
Constraint (2) ensures that the availability of product *k* at the suppliers is greater than the customer demand. Constraint (3) states that the quantity of product *k* sent by each supplier to the cross-docks is less than the availability of product k. Constraint (4) ensures that the total demand of the product k will be satisfied by the total quantity of product k shipped from *i* to *n*. Constraint (5) states that the quantity of product *k* which arrives at the cross-dock *n* is equal to the quantity of that product which leaves that cross-dock. Constraint (6) expresses the concept that the demand placed by each customer *j* must be entirely satisfied through the quantity which leaves the node *n*. Constraint (7) says that the total quantity picked from all the suppliers must be equal to the total quantity shipped to all clients. Constraints (8) and (9) ensure that the total quantity shipped respectively from a supplier or a cross-dock is equal to the capacity of a single truck multiplied for the number of trucks traveling the arc *i-n* or *n-j*. Constraints (10) and (11) ensure that the route *i-n* or *n-j* will be active only if at least a unit of product will be shipped. Finally Constraints (12) and (13) ensure that the number of truck is greater than zero only if the correspondent route is active.

2.2. INLP model for direct delivery transshipment problem

The direct delivery system consists in a network composed by *i* origins (suppliers) and *j* destinations (clients) in which each origin serves all the destination with a direct link. By assuming that each client will require a products mix and each origin makes only a product type there will be as many direct links from origin to destination how many clients will there be in the network. The model has been formulated under the following notations:

i, number of suppliers
j, number of clients
k, number of products
C, maximum truck capacity
p_{jk}, demand of product *k* for the customer *j*
d_{ij}, distance between *i* and *j*
c_k, variable transport cost of the product *k* per unit distance from the origin (supplier) to the destination (client)
f_{ij}, fixed transport cost between the source *i* and the destination *j*. Such cost is proportional to the number of trucks routed between the source and the destination.
M = 100,000, upper bound
N = 100,000, upper bound

The assumptions made for the cross-docking model result valid also for the present model by considering that the assumption 1is referred to the relation supplier-client.

Decision variables:

q_{ijk}, transported quantity of product *k* from *i* to *j*
N_{ij}, non negative variable representing the number of trucks on arc *i-j*
x_{ij}, binary variable

$$x_{ij} = \begin{cases} 1, if\ the\ route\ i-j\ is\ active \\ 0 \qquad\qquad\qquad otherwise \end{cases}$$

The objective function:

$$min\ z = \sum_{i,j,k} q_{ijk} * c_k * d_{ij} + \sum_{i,j} f_{ij} * x_{ij} * N_{ij} \quad (14)$$

s.t.:

$$\sum_i a_{ik} \geq \sum_j p_{jk} \qquad \forall\ k \quad (15)$$

$$\sum_j q_{ijk}\, x_{ij} \leq a_{ik} \qquad \forall\ i,k \quad (16)$$

$$\sum_i q_{ijk}\, x_{ij} = \sum_j p_{jk} \qquad \forall\ j,k \quad (17)$$

$$\sum_k q_{ijk} * x_{ij} \leq * C * N_{ij} \qquad \forall i,j \quad (18)$$

$$x_{ij} * M \geq \sum_k q_{ijk} \qquad \forall i,j \quad (19)$$

$$x_{ij} * N \geq N_{ij} \qquad \forall i,j \quad (20)$$

All the variables must be non negative.

The objective function (14) is the same of the corresponding function in the cross-docking model. Similarly to the previous model constraint (15) ensures that the availability of product *k* at the suppliers is greater than the customer demand, while constraint (16) says that the quantity of product *k* sent by each supplier to the clients is less than the availability of product *k*.
Constraint (17) ensures that the total demand of the client *j* for the product *k* will be satisfied by the total quantity of that product shipped from *i* to *j*. Constraint (18) shows that the total quantity shipped from a supplier is equal to the capacity of a single truck multiplied for the number of trucks traveling the arc *i-j*. Constraint (19) ensures that the route *i-j* will be active only if at least a unit of product will be shipped. Finally constraint (20) ensures that that the number of truck is greater than zero only if the correspondent route is active.

2.3. Simulation Model-Building :Conceptual model definition and conceptual model translation for cross-docking and direct delivery systems

In order to built the simulation models the Model-Building step is performed consisting in the conceptual model definition and the conceptual model translation. The conceptual model definition can be expressed either formally (e.g. Activity Cycle Diagram) or informally (e.g. a list of assumptions) (Robinson, 1997). It aims at representing the actual operations carried out by the network (i.e. demand receiving, collection, shipping). The conceptual model definition for the cross-docking system is characterized by the assumptions 1-6 yet seen in the sub-sections 2.1. It must be pointed out that the assumption 1 implies that there will not occur queuing delays at the suppliers facilities neither stockout costs will be incurred. For the purpose of simulation the customer demand will be modeled as a random variable. Finally it must be added that the lead times between suppliers and cross-docks and cross-docks and clients will be considered null.
The conceptual models based on the previous discussed assumptions have been translated into simulation models through a C++ code. The assumptions have been accurately reproduced and the simulation models do not differ substantially from the conceptual models. Concerning the cross-docking system the conceptual model translation has been realized by considering the effort required to the simulation model. It arises from the deterministic model represented by the INLP model which is a NP hard problem whose complexity increases as the number of suppliers, cross-docks and clients increases as well. Consequently the effort required to determine the optimal solution of the INLP models and that required to the simulation models are very heavy. In order to reduce such effort the cross-docking network has been configured with only a cross-dock. This makes the INLP model simple to solve as well as the consequent simulation model. On the basis of this configuration of the network the simulation model has been realized.
The assumptions formulated in section 2.1 have been translated into the two simulation models in the following way:

1. The decision variables are constituted by the quantities transshipped, the number of routes activated and the number of trucks which travel along each route arisen from the optimal solution of the two NILP models.
2. The random demand which is an uncontrollable input variable of the model has been generated according to a normal distribution.
3. In the cross-docking simulation model each client in the system places an order of *k* different products. Thus the number of trucks needed is determined and the requested quantity will be sent from the suppliers to the cross-dock facility. Here the products are unloaded and collected on the basis of the customers products mix, the number of trucks needed to the shipment is determined and the products are sent to the clients.
4. For the direct delivery simulation model each client in the system places an order of *k* different products to each supplier. On the basis of the requested quantity the number of trucks needed will be calculated. Thus the requested quantity will be sent to the client.
5. The Total Cost and the Utilization Coefficient of trucks representing the output measures are determined.

2.4. Input data validation

Once the simulation model has been defined the next step to be realized consists in the validation of the input data consisting in the verification of the correspondence

between data collected with those achievable by a real system. In our case the data are represented by the number of nodes present in the network, the number of trucks routed, the distances between nodes, the customer demand and the variable and fixed costs.
The number of nodes of the network have been defined by starting from a real system and the distances between them have been arisen by starting from the Cartesian coordinates of each suppliers and clients. The localization of the cross-dock corresponds to the origin of the Cartesian plane. A qualitative representation of the cross-dock network is reported in Figure 1.

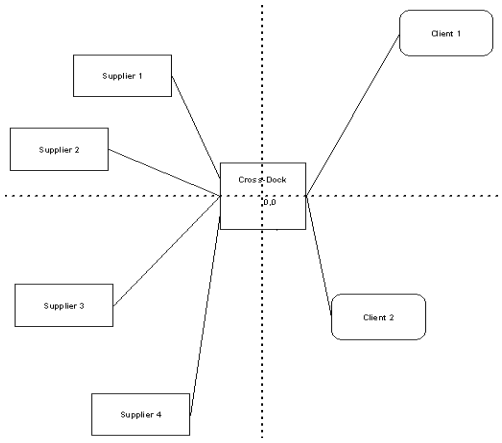


Figure 1. Cross-docking network

The active nodes as well as the number of trucks for each active route arise from the optimal solution of the INLP models. As regards the customer demand it has been modeled through a normal distribution to take into account the demand fluctuations for the products managed. This normal distribution is the result of the fitting of a time series analysis of food market demand.
Finally the variable and fixed costs have been defined by considering the data reported in a logistic review (TIR, 2010).

3. EXPERIMENTAL APPLICATION

In this section an experimental application of the discussed methodology is addressed consisting in running the two INLP models in order to get the optimal candidate solution of the two distribution systems considered and in the subsequent post-optimality analysis of the solutions carried out by means of the simulation tool. In both cases the models have been run by considering a general case of a network with four suppliers (S1-S4), two clients (C1-C2), one cross-dock (CD1) and four products (P1-P4). The input parameters consisting in the customers demand, the products availability and the distances are reported in Tables 1-5. The trucks capacity has been fixed equal to 70 units, the variable costs for products P1-P4 are respectively of 0.007€/unit*km, 0.005€/unit*km, 0.006€/unit*km, 0.003€/unit*km, while the fixed cost is equal to 250€ for the cross-docking strategy and 350€ for the direct delivery strategy.

Table 1. Customer demand for products P1-P4

Client/Product (unit/client)	P1	P2	P3	P4
C1	30	25	18	27
C2	40	25	30	25

Table 2. Products availability

Supplier/Product availability	P1	P2	P3	P4
S1	1,000	-	-	-
S2	-	1,000	-	-
S3	-	-	1,000	-
S4	-	-	-	1,000

Table 3. Distance Supplier-Cross-Docks

Distance Supplier/Cross-dock (km)	CD1
S1	100
S2	150
S3	170
S4	150

Table 4. Distance Cross-Docks-Clients

Distance Cross-Dock/Client (km)	C1	C2
CD1	150	110

Table 5. Distance Supplier-Client

Distance Supplier/Client (km)	C1	C2
S1	180	200
S2	220	250
S3	210	225
S4	190	160

At first the two INLP models have been solved by using LINGO software. Results show that a feasible solution can be found in both cases and the routes, the quantity to ship, the number of trucks that must travel along each route and the Total Cost have been determined. Results are reported in Table 6, 7 and 8.

Table 6. Results of the INLP model for Cross-Docking

Cross-Docking-INLP Model Solution	
Q_IKN(S1, P1, CD1)	70
Q_IKN(S2, P2, CD1)	50
Q_IKN(S3, P3, CD1)	48
Q_IKN(S4, P4, CD1)	52
X_IN(S1-CD1)	1
X_IN(S2-CD1)	1
X_IN(S3-CD1)	1
X_IN(S4-CD1)	1
N_IN(S1-CD1)	1
N_IN(S2-CD1)	1
N_IN(S3-CD1)	1
N_IN(S4-CD1)	1
Q_KNJ(P1, CD1, C1)	30
Q_KNJ(P1, CD1, C2)	40
Q_KNJ(P2, CD1, C1)	25
Q_KNJ(P2, CD1, C2)	25
Q_KNJ(P3, CD1, C1)	18

Q_KNJ(P3, CD1, C2)	30
Q_KNJ(P4, CD1, C1)	27
Q_KNJ(P4, CD1, C2)	25
X_NJ(CD1-C1)	1
X_NJ(CD1-C2)	1
N_NJ(CD1-C1)	2
N_NJ(CD1-C2)	2

Table 7. Results of the INLP model for Direct Delivery

Direct Delivery-INLP Model Solution	
Q_IJK(S1, C1, P1)	30
Q_IJK(S1, C2, P1)	40
Q_IJK(S2, C1, P2)	25
Q_IJK(S2, C2, P2)	25
Q_IJK(S3, C1, P3)	18
Q_IJK(S3, C2, P3)	30
Q_IJK(S4, C1, P4)	27
Q_IJK(S4, C2, P4)	25
X_IJ(S1-C1, S1-C2)	1
X_IJ(S2-C1, S2-C2)	1
X_IJ(S3-C1, S3-C2)	1
X_IJ(S4-C1, S4-C2)	1
N_IJ(S1-C1, S1-C2)	1
N_IJ(S2-C1, S2-C2)	1
N_IJ(S3-C1, S3-C2)	1
N_IJ(S4-C1, S4-C2)	1

Table 8. Output of interest measures for Cross-Docking and Direct Delivery

	Cross-Docking		Direct Delivery
Total Cost (€)	2,310.06		3,043.12
Total Variable Cost (€)	310.06		243.12
Total Fixed Cost (€)	2,000		2,800
	Route S-CD	Route CD-C	Route S-C
Number of routes active	4	2	8
Number of trucks	4	4	8
	S-CD	CD-C	S-C
Mean Utilization Coefficient	0.7857	0.7857	0.3928

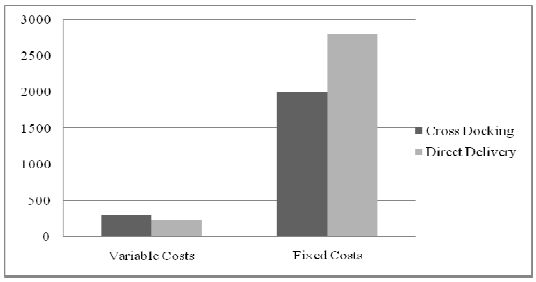


Figure 2. Comparison of Variable an Fixed Costs between the two strategies

As shown in Table 8 the cross-docking outperforms the direct delivery. In particular by analyzing the Total Cost (see also Figure 2) it can be observed that the cross-docking has lower fixed costs compared to the direct delivery, while the variable costs are greater that the direct delivery. This can be explained by considering that the fixed costs depend on the number of trucks routed and on the routes activated that in the case of cross-docking are less than the direct delivery case. The variable costs are lower in the direct delivery strategy because they depend on the total distances traveled which are lower in the this case. Finally the Utilization Coefficient is greater in the cross-docking strategy than the direct delivery one. This can be explained by considering that in the cross-docking strategy the number of routes activated between the suppliers and the cross-dock and the cross-dock and the client are lower than the total number of routes activated in the direct delivery strategy between suppliers and clients. This involves a less number of trucks will be routed in the cross-docking strategy in each supplier-cross-dock and cross-dock-client route as compared to the total number of trucks managed in the direct delivery strategy.

Successively the best candidate solutions found for the two INLP models have been employed to perform a post-optimality analysis consisting in running the two simulation models yet discussed in section 2.3. For the purpose of the present study it aims at showing the sensitivity of solutions found in the case in which the customer demand is subject to fluctuations. The only uncontrollable input variable of the simulation models is the normal distribution representing the customer demand whose mean is equal to the customer demand of the deterministic case (Table 1) while the standard deviation is reported in Table 9.

Table 9. Standard deviation of demand

Client/Standard Deviation of Product	P1	P2	P3	P4
C1	8	9	5	6
C2	7	5	4	9

For the cross-docking model once the customer's order for each product is placed each supplier will send to the cross-dock a supply equal to the total customer demand for the single product considered. Each of the routes which join the suppliers with the cross-dock results active. The number of trucks traveling along each route is determined by dividing the quantity shipped by each supplier for the truck capacity. Once the products arrive at the cross-dock they are consolidated on the basis of the customer requests. Thus the number of trucks leaving the cross-dock is determined by dividing the total product demand of each customer for the truck capacity.

For the direct delivery model the only difference is that at the arriving of the customer demand each supplier will determine the trucks number by dividing each customer demand for the truck capacity and it will send

directly to the customer the required quantity. Each route joining the suppliers with the customers results active.
Ten thousand replications of the two simulation models have been carried out to ensure a 99.5% of accuracy of the interest measures which are the Average Total Cost and the Average Utilization Coefficient of the trucks routed. The number of runs needed to ensure the accuracy has been determined by running at first ten replications of the two simulation models and by calculating the number of runs needed to ensure the desired precision by means of the following formula:

$$n_\alpha(\varepsilon,\alpha) = \frac{4\sigma^2 z_{\frac{\alpha}{2}}^2}{\varepsilon^2} \quad (21)$$

where:
σ is the standard deviation of the interest measures on the basis of the initial ten run,
$z_{\alpha/2}$ is the normal random variable of a standard normal distribution corresponding to the precision required 1-α,
ε is the absolute width of the confidence interval referred to the ten run and determined as:

$$\varepsilon_\alpha(\alpha) = \frac{2z_{\alpha}\sigma}{\sqrt{n}} \quad (22)$$

where *n* is the initial number of runs.
For detailed discussion about statistic aspect refer to Whitt (2005). The results are showed in Tables 10 and 11.

Table 10. Results of simulation for Cross-Docking

Cross-Docking Model			
	Mean	Standard Deviation	Precision
Average Total Cost (€)	2,457.59	189.762	0.996
Average Total Variable Cost (€)	310.14		
Average Total Fixed Cost (€)	2,147.45		
Route S-CD	4		
Route CD-C	2		
	S-CD	CD-C	
Mean Number of Trucks	(4.54)	(4.04)	
Mean Utilization Coefficient	0.697	0.77880	0.996

Table 11. Results of simulation for Direct Delivery

Direct Delivery Model			
	Mean	Standard Deviation	Precision
Average Total Cost (€)	3,037.56	37.6239	0.999
Average Total Variable Cost (€)	239.94		
Average Total Fixed Cost (€)	2,797.62		
Route S-C	8		
	S-C		
Mean Number of Trucks	8		
Mean Utilization Coefficient	0.3931		

The stochastic scenario shows that the Average Total Cost of the cross-docking strategy is less than the direct delivery one similarly to the case of deterministic configuration. In the stochastic case the average percentage of saving cost is about of 19%. It is worth to underline that the standard deviation of the Average Total Cost is greater in the cross-docking strategy as compared to the direct delivery one as you can see also from Figures 3 and 4. Such figures report the frequency of the Total Cost for the 10,000 replications of the two simulation models. They underline that the Total Cost function in the cross-docking strategy ranges between 1,950€ and 3,750€, while in the direct delivery case ranges from 2,600€ and 3,500€. The greater standard deviation of the cross-docking system can be explained by considering that the Utilization Coefficient value is greater in the cross-docking strategy compared to the direct delivery one. In fact when the customer demand increases due to the demand fluctuation the number of trucks in the cross-docking system tends to increase as well by causing the increasing of the Average Total Cost, while in the direct delivery strategy the number of trucks tends to be always the same due the low Utilization Coefficient and the Average Total Cost increasing is very low.

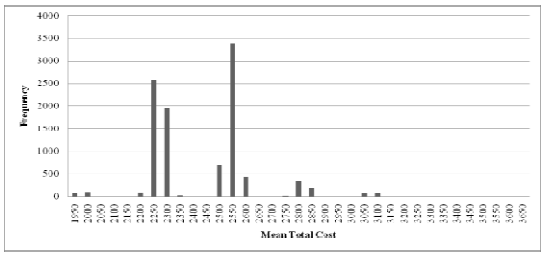


Figure 3. Frequency of Total Cost for Cross-Docking

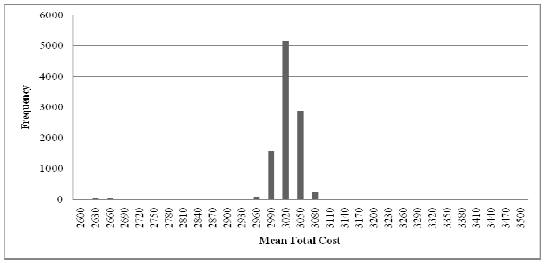


Figure 4. Frequency of Total Cost for Direct Delivery

At the end of the experimental analysis a sensitivity analysis has been conducted by varying the Average

Customer Demand of ±5% and of ±10% to show the impact of demand fluctuation on the Average Total Cost, the Average Standard Deviation and the Average Utilization Coefficient. Results are reported in Tables 12 and 13.

Table 12. Sensitivity analysis for Cross-Docking

Sensitivity analysis-Cross-Docking Model (-10%, -5%,+5%, +10%)		
	Mean	Standard Deviation
Average Total Cost (€)	(2,324.7, 2,380.94, 2,539.5, 2,634.7)	(162.64, 170.5, 213.9, 238.50)
	S-CD	CD-C
Average Number of Trucks	(4.26, 4.37, 4.72, 4.90)	(3.92, 3.97, 4.13, 4.27)
Average Utilization Coefficient	(0.667, 0.686, 0.7050, 0.711)	(0.723, 0.75, 0.80, 0.814)

Table 13. Sensitivity analysis for Direct Delivery

Sensitivity analysis-Direct Delivery Model (-10%, -5%,+5%, +10%)		
	Mean	Standard Deviation
Average Total Cost (€)	(3,010.4, 3,024.0, 3050.5, 3048.2)	(50.41, 44.2, 31.8, 29.6)
	S-C	
Average Number of Trucks	(7.98, 7.98, 8, 7.99)	
Average Utilization Coefficient	(0.354, 0.373, 0.412, 0.43)	

Results show that in both the case of cross-docking and direct delivery strategies an increasing in the customer demand causes the Average Total Cost, the Average Number of trucks and the Average Utilization Coefficient increase as well. On the other hand the Average Standard Deviation increases with the customer demand in the case of cross-docking strategy while decreases in the case of direct delivery one. This means that the cross-docking strategy is more sensitive to the fluctuation of the demand than the direct delivery. The percentage of variation of the Average Total Cost, the Average Number of Trucks and of the Average Utilization Coefficient for a variation of ±5% and ±10% of the Average Customer Demand for the two strategies is reported in Table 14. The variation is calculated to respect to the average values already seen in Tables 10 and 11.

Table14. Percent variation of the interest measures

		Absolute variation Cross-Docking	Absolute variation Direct Delivery
Average Total Cost	-10%	5.71%	0.9022%
	-5%	3.21%	0.4484%
	5%	3.22%	0.4242%
	10%	6.72%	0.3491%
	Average variation	4.72%	0.5310%
		S-CD	CD-C	S-C
Average Number of Trucks	-10%	6.57%	3.06%	0.251%
	-5%	3.89%	1.76%	0.251%
	5%	3.81%	2.17%	0.000%
	10%	7.34%	5.38%	0.125%
	Average variation	5.40%	3.09%	0.157%
Average Utilization Coefficient	-10%	4.50%	7.72%	11.05%
	-5%	1.60%	3.84%	5.39%
	5%	1.13%	2.65%	4.59%
	10%	1.97%	4.32%	8.58%
	Average variation	2.30%	4.63%	7.40%

Results show that the average variation of the Average Total Cost is equal about to 4.72% in the case of cross-docking and only to 0.53% in the case of direct delivery strategy. However the Average Total cost of the cross-docking strategy results always less than that of direct delivery. The average variation of Average Number of Trucks is equal about to 5.40% and 3.09% in the case of cross-docking and only to 0.15% in the case of direct delivery. This confirm the greater sensitivity of the cross-docking strategy to respect these two measures. In fact when the average demand increases/decreases the Average Number of Trucks increases/decreases as well and the Average Total Cost consequently (Tables 12 and 13). On the contrary the Average Utilization Coefficient varies about of 2.30% and 4.63% in the case of cross-docking and of 7.40% in the case of direct delivery. This substantially confirms that in the case of cross-docking strategy the number of trucks tends to increases as the average demand increases as well, while in the case of direct delivery when the Average Utilization is low, the number of trucks tends to be always the same and consequently the Average Utilization increases (Tables 12 and 13).

4. CONCLUSIONS

In this paper a transshipment problem for the cross-docking system has been addressed. At first a deterministic solution of the problem has been found by means of a INLP model and subsequently by starting from this optimal solution a post-optimality analysis has been conducted by means of the simulation approach in order to take into account possible fluctuations in the customer demand. The cross-docking system has been also compared to the direct delivery strategy. Results show that the cross-docking strategy outperforms the direct delivery one as regard the optimization of the Total Cost in both the deterministic and the stochastic case.

REFERENCES

Agustina D., Lee C.K.M., Piplani R., 2010. A Review: Mathematical Models for Cross Docking Planning.

International Journal of Engineering Business Management 2(2): 47-54.

Aickelin U., Adewunmi A., 2008. Simulation Optimization of the Crossdock Door Assignment Problem. *Computer Science. Neural and Evolutionary Computing*.

Apte, U. M., Viswanathan, S., 2000. Effective cross-docking for improving distribution efficiencies. *International Journal of Logistics* 3: 291–302.

Arnaout G., Rodriguez-Velasquez E., Musa R., 2010. Modeling Cross-Docking Operations using Discrete Event Simulation, *Proceedings of the 6th International Workshop on Enterprise & Organizational Modeling and Simulation* 601: 113–120. 2010, Hammamet, Tunisia.

Boloori Arabani A., Zandieh M., Fatemi Ghomi, (2011). Multi-objective genetic-based algorithms for a cross-docking scheduling problem. *Applied Soft Computing* 11(8): 4954–4970.

Charkhgard H., Tabar Y., 2011. Transportation problem of cross-docking network with three-dimensional trucks. *African Journal of Business Management* 5(22): 9297-9303.

Chen P., Guo Y., Lim A., Rodrigues B., 2006. Multiple crossdocks with inventory and time windows. *Computers and Operations Research* 33: 43–63.

Gue, K.R., 1999. The effects of trailer scheduling on the layout of freight terminals. *Transportation Science* 33(4): 419-428.

Heragu, S.S., Du L., Mantel R.J., Schuur P.C., 2005. Mathematical model for warehouse design and product allocation. *International Journal of Production Research* 43(2): 327-338.

Larbi, R., Alpan G., Baptiste P., Penz B., 2007. Scheduling of Transshipment Operations in a Single Strip and Stack Doors Crossdock. *19th International Conference on Production Research ICPR*.

Li, Y., Lim, A., Rodrigues, B., 2004. Crossdocking-JIT scheduling with time windows. *Journal of the Operational Research Society* 55: 1342–1351.

Li, Z., Low M.Y.H., Lim, R.Y.G. 2009a. Optimal decision-making on product allocation for crossdocking and warehousing operations. *International Journal of Services Operations and Informatics*. 4 (4): 352-365.

Li Z., Low M. Y. H., Shakeri M., Lim Y.G., 2009b. Crossdocking planning and scheduling: Problems and algorithms. *SIMTech technical reports* 10, 3.

Li Z., Sim C. H., Low M., Lim V. G., 2008. Optimal Product Allocation for Crossdocking and Warehousing Operations in FMCG Supply Chain. *Service Operations and Logistics and Informatics*. *IEEE*, 2: 2963-2968.

Lim, A., Ma H., Miao Z., 2006. Truck Dock Assignment Problem with Time Windows and capacity Constraint in Transshipment Network Through Crossdocks. *Lecture Notes in Computer Science* 3982: 688-697.

Lim, A., Miao, Z., Rodrigues, B., Xu, Z., 2005. Transshipment through Crossdocks with Inventory and Time Windows. *Naval Research Logistics* 52: 724-733.

Liu, Y., S. Takakuwa, 2009. Simulation-based personnel planning for materials handling at a crossdocking center under retail distribution environment. *In Proceedings of the 2009 Winter Simulation Conference*, 2414-2425.

Liu Y., 2010. Simulation Modeling and Non-automated Cross-docking Center. Proceedings of the 17Th *International Conference on Industrial Engineering and Engineering Management (IE&EM),*1928-1932, 29-31 Oct. 2010, Nagoya, Japan.

Liu Y., Takakuwa S., 2010. Enhancing simulation as a decision-making support tool for a crossdocking center in a dynamic retail-distribution environment. *Proceedings of the 2010 Winter Simulation Conference*. 2089-2100. 2010, Japan.

Magableh, G. M., Rossetti M., 2005. Modeling and analysis of a general cross-docking facility. *In Proceedings of the 2005 Winter Simulation Conference*, 1613-1620, 2005, Piscataway, New Jersey.

Miao Z., Fu K., Fei Q., Wang F., 2008. Metaheuristic Algorithm for the Transshipment problem with Fixed Transportation Schedules. *Lecture Notes in Computer Science* 5027: 601-610.

Musa R., Arnaout J., Jung H., 2010. Ant colony optimization algorithm to solve for the transportation problem of cross-docking network. *Computers & Industrial Engineering* 59: 85–92.

Robinson S. 1997. Simulation Model Verification and Validation: Increasing the Users' Confidence. *Proceeding of the 1997 Winter Simulation Conference*, 53-59, 1997, San Diego, CA,.

Rohrer M., 1995. Simulation and cross docking. *In Proceedings of the 1995 Winter Simulation Conference*, 846-849, 1995, New Jersey.

Smith R., 1999. Simulation: The Engine Behind The Virtual World. *Simulation 2000 Series*. 1.

TIR: La rivista dell'autotrasporto, Luglio 2010, n°128.

Tsui L.Y., Chang C-H., 1992. An optimal solution to dock door assignment problem, *Computer & Industrial Engineering* 23 (1-4): 283-286.

Vahdani B., Zandieh M., 2010. Scheduling trucks in cross-docking systems: Robust meta-heuristics. *Computers & Industrial Engineering* 58: 12–24.

Whitt W., 2005. Analysis for the design of simulation experiments. In: Henderson, S:, Nelson, B:, Simulation. Elsevier series of Handbooks, *Operations Research and Management Science*, Chapter 13.

Yu Wooyeon, Egbelu Pius J., 2008. Scheduling of inbound and outbound trucks in cross docking systems with temporary storage. *European Journal of Operational Research* 184: 377–396.

THE EFFECT OF TOE MECHANISM FOR SIMULATION OF SMALL BIPED WALKIN ROBOT BY GAIT GENERATION

Krissana Nerakae[(a)], Hiroshi Hasegawa[(b)]

[(a)]Functional Control Systems – Graduate School of Engineering and Science
Shibaura Institute of Technology, Japan
[(b)]Department of Systems Engineering and Science
Graduate School of Engineering and Science
Shibaura Institute of Technology, Japan

[(a)]m710505@shibaura-it.ac.jp, [(b)]h-hase@shibaura-it.ac.jp

ABSTRACT
The researches of biped robot have a long history and continuation. One important research and very basic movement is walking. However, the present research of biped robot is still far from proposing a solution which generates a level of flexibility and reliability gait pattern that would enable practical walking. To solve its problem, in this paper, we consider the feet of the robot which is one of the most important points as human feet in order to improve the flexibility of robot movement from heel to toe. The design of the various components of the robot feet, its flexibility and stability are excellent although this gait is not suitable for a quick walk. In this study, we introduce feet which have the toe mechanism to a small biped robot through inspiration from its adaptive walk. In the toe mechanism, we want to reduce the power consumption of the robot, therefore spring-damper is used instead of the sensor.

Keywords: biped robot, gait, toe mechanism

1. INTRODUCTION

The researches of humanoid robot and biped robot have a long history and continuation. One important research and very basic movement is walking. However, the research present of biped robot is still far from prosing a solution which generates a level of flexibility and reliability gait pattern that would enable practical walking on the variety of rough ground humans negotiation with the easiness on a regular basis.

To solve its problem, in this paper, we consider the feet of the robot which is one of the most important points as human feet in order to improve the flexibility of the robot movement from heel to toe. The design of the various components of the robot foot such as heel tiptoe or big toe, its flexibility and stability are excellent although this gait is not suitable for a quick walk.

Until recently, the studies of gait analysis for walking biped robot are incessant. Zhe Tang et al. have proposed an optimization for humanoid walking based on Genetic Algorithm (GA) base optimization for humanoid walking (Zhe et al. 2006). Lingyun Hu et al. have presented bipeds gait optimization using spline function based on probability model (Lingyun et al. 2006). These studies are about the gait optimization of biped robot. Furthermore, a natural human walking was proposed in tiptoe mechanism for biped robot. Y. Xiang et al. have presented optimization based dynamic human walking prediction (Xiang et al. 2007) which were studied an optimization-based approach for simulation of the motion of a digital human model. A model had 55 degrees of freedom which included tiptoe joints. Nandha Handharu et al. have proposed gait pattern generation with knee stretch motion for biped robot using toe and heel joints (Nandha et al. 2008). Cheol Ki Ahn et al. have proposed development of a biped robot with toes to improve gait pattern (Cheol et al. 2003), the gait pattern of the robot with toes was compared to the robot without toes by 3D graphical simulation. Shuuji Kajita et al. have proposed zero of moment point (ZPM) based running pattern generation for a biped robot equipped with toe spring (Shuuji et al. 2007). Abovementioned research was based on flat plate and there is no comparison between results obtained from a different foot.

In addition, there are some studies which related to framework for biped robot locomotion. S. Ali A. Moosavian et al. have proposed the introduction of a cartesian approach for gate planning and control of biped robots and implementation on various slopes (S. Ali A. Moosavian et al. 2007). Naoya Ito and Hasegawa Hiroshi have presented a robust optimization uncertain factors of environment for simple gait of biped robot (Naoya and Hasegawa 2007), to optimize the gait for biped robot by using Simulated Annealing (SA). The robust optimization considered random values as floor of fiction and restitution. Yu Zheng et al. have proposed a walking pattern generator for biped robots on uneven terrains (Yu et al. 2010), these approaches were more general and applicable to uneven trains as compared with prior research methods based on the ZMP criterion. Abovementioned were used without toes mechanism model.

In this study, to get flexibility and reliability of gait pattern, we introduce feet which have the toe mechanism to a small biped robot through the inspiration of its adaptive walk. In the toe mechanism,

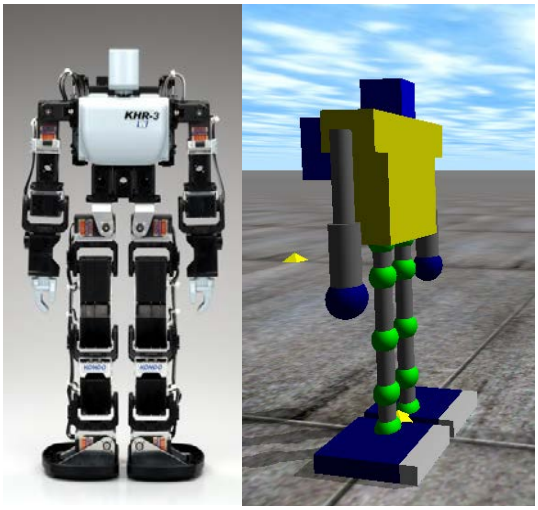
we want to reduce the power consumption of the robot therefore spring-damper is used instead of the sensor.

This paper presents the design of foot mechanisms of biped robot which have effects on walking by gait generation method. This study is the prior step for the optimization step.

2. METHODOLOGY

2.1. Simulation Model

In this paper, a robot model is simulated on flat plate as friction constant by basing on KHR-3HV model. The robot is shown in Figure 1. It has 10 RC-Servo motor under the hip and has a mass of 1.5 kg. The simulation uses the same degree of freedom on these joints.



KHR-3HV Robot Simulation Model

Figure 1: KHR-3HV Robot And Simulation Model

Several researches of toe joint utilization in bipedal locomotion have been proposed such as, 1) passive toe joints in order to achieve stable feet lifting, 2) toe joints that are both actively and passively control for less energy consumption walking and 3) active toe joints for stepping up stairs. In this study, to get flexibility and reliability of gait pattern, we introduce a passive toe mechanism to a small biped robot which is shown in Figure 2.

2.2. Simulation Based Design (SBD)

The design concept of the toe mechanisms to create models from the bone structures of the human foot and support force area on the foot are shown in Figure 3 and Figure 4, respectively.

In Figure 3, the bones of the toes are called the phalanges. The phalanges are jointed to the 5 metatarsal bones. Behind the metatarsal bones are a series of smaller bones known as the tarsal bones. The heel bone is called the calcaneus, which is connected to the talus bone (the largest bone of the ankle) walk. (http://www.chakras.org.uk)

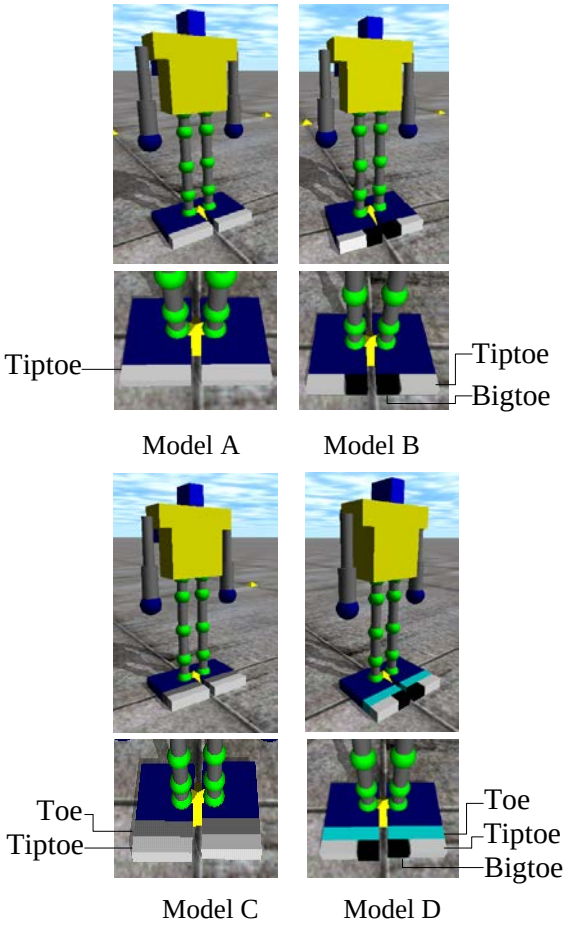


Figure 2: The Simulation Models With Toe Mechanism
Model A: The Simulation Model With Tiptoe
Model B: The Simulation Model With Tiptoe And Big Toe
Model C: The Simulation Model With 2-Tiptoes
Model D: The Simulation Model With 2-Tiptoes And Big Toe

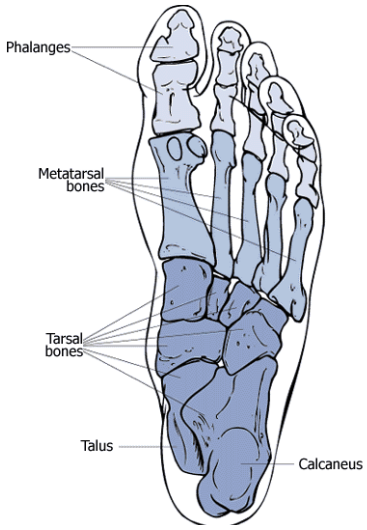


Figure 3: A Foot Bone Structures

In Figure 4, Perry J. (Perry, 1992) shows the sequence of foot support areas during stance phase. The

black area is the position where supports forces areas. We used in a conceptual design to make toe mechanism model.

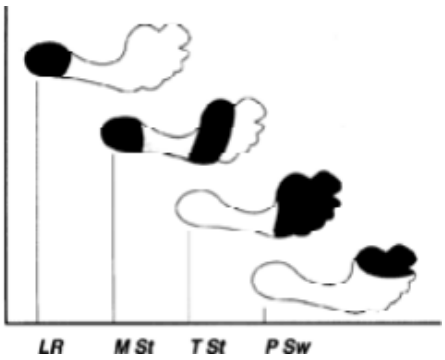


Figure 4: Human Sequence Of Foot Support Areas During Stance Phase (Perry, 1992)

2.2.1. Definition of the gait function

This paper is assumed the robot walks based on the gait function. Therefore, the function is defined based on a human gait pattern that focused on the walk cycle. To express this periodic cycle, the function which generated the gait is defined as follows:

$$\theta_i(t) = a_i + b_i \cos(\omega t) + c_i \sin(\omega t) + d_i \cos(2\omega t) \quad (1)$$

Where t is time, ω is angular velocity, i is number of each joint a, b, c, and d are coefficients of generating the gait for various wave. The gait for biped robot is changed by operating these coefficients.

2.2.2. Adaptation to simulation

The sampling time of the function to generate the gait is quarter a gait cycle. The generated angle data is allocated the joint for position control value. A joint moves with a constant velocity between control points. In this simulation, *1 cycle* of walking is defined *1.2 seconds*. Thus, angular velocity is given as follows:

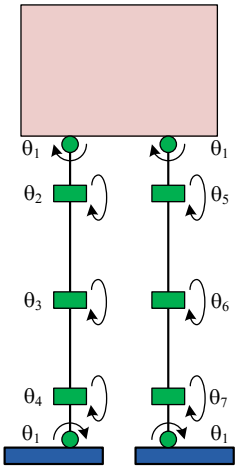
$$\omega = \frac{2\pi}{1.2} \quad (2)$$

3 cycles of walking time is *3.6 seconds*. And the total time is *7.0 seconds* taking *3.4 seconds* in order to check after walking stability. In this simulation, *1 step* takes *0.001 seconds*, thus the number of total step is *7000 steps*. For example, a gait pattern of a joint angle which is made by the gait function.

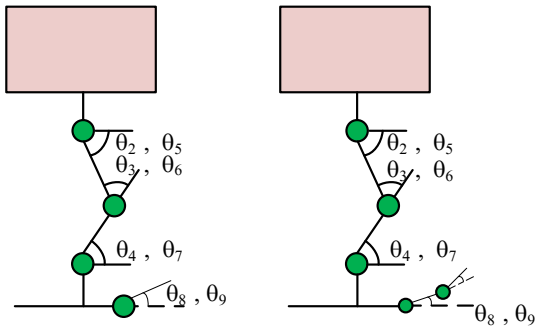
The position of the joint is shown in Figure 5. In addition, because of the servomotor of the joint of the biped robot, its joint can be rotated by *60 degree* every *0.14 seconds*. The rotate directions for each joint are shown in Table 1. Knee joints do not rotate to backward direction from standing. Thus, these joint are stricter rotating to minus angle as follows:

$$\theta_3 = \begin{cases} 0 & if \ \theta_3 < 0 \\ \theta_3 & if \ 0 \leq \theta_3' \end{cases} \quad (3)$$

$$\theta_6 = \begin{cases} 0 & if \ \theta_6 < 0 \\ \theta_6 & if \ 0 \leq \theta_6' \end{cases} \quad (4)$$



Font side view



Right side view

Figure 5: The Link Of Model

Table 1: Parameter And Rotation Direction

Parameter	Leg	Joint	Rotation Direction
θ_1	Both	Hip and Ankle	Side-to-Side
θ_2	Right	Hip	Backward-and-Forward
θ_3	Right	Knee	Backward-and-Forward
θ_4	Right	Ankle	Backward-and-Forward
θ_5	Left	Hip	Backward-and-Forward
θ_6	Left	Knee	Backward-and-Forward
θ_7	Left	Ankle	Backward-and-Forward
θ_8	Right	Toe	Backward-and-Forward
θ_9	Left	Toe	Backward-and-Forward

The horizontal surface is applied for the ground surface of the simulation. Moreover, for the ground surface, friction and restitution coefficients are defined as *1.0* and *0.0*, respectively.

2.2.3. The design of simulation experiments

The design of simulation experiments, design variable vectors can generate form gait function Eq. (1) to each joints expect for the toe joint because we are determined to be passive toe mechanism, design variable vectors shown in Eq. (5) and have degree of freedom (DoF) is 28.

$$\theta_i = [a_i, b_i, c_i, d_i]; \qquad (i = 1,2,3,4,5,6,7) \qquad (5)$$

$$\theta_{all} = [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7]$$

In order to the model move forward and on the path, by the end of the simulation we define the following conditions as Eq. (6) and (7). When X_d are the distances at the side under ± 30 *[mm]* in Eq. (6) and in Eq. (7) R_d is the angle to the rotation direction under ± 5 *degree* and distance Y_d not exceeding *200 [mm]* to prevent the slip. When X_d, Y_d *and* R_d are denoting the distances and rotation of model's center of mass (CoM) shown in Figure 6.

$$-30 \leq X_d \leq 30 \ [mm] \quad if \quad t = 7.0 \ [sec] \qquad (6)$$

$$-5.0 \leq R_d \leq 5.0 [deg] \quad if \quad t = 7.0 \ [sec] \qquad (7)$$

$$Y_d \leq 200 \quad [mm] \quad if \quad t = 7.0 \ [sec] \qquad (8)$$

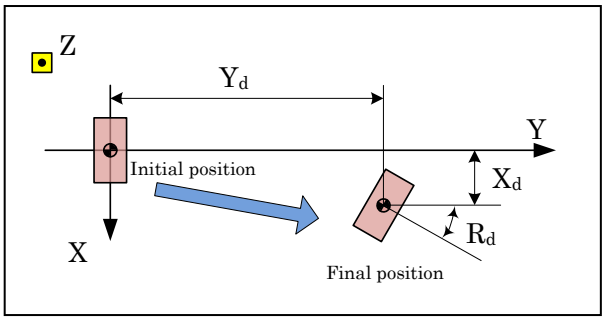


Figure 6: Overview Of The Simulation

3. RESULTS OF THE SIMULATION

All simulation models of small biped robot with passive toe mechanism can walk. The results are shown in Table 2.

The results, model A and model B are according to the condition. The maximum distance (Y_d) is model B *119 [mm]*, minimum distance side (X_d) is *-5 [mm]* and rotation (R_d) *= 3.1 [deg]*. The latter is the model A which distance (Y_d) is *86 [mm]*, distance side (X_d) is *-25 [mm]* and rotation (R_d) *= 4.2 [deg]*. On the other hand model C and model D, The results are not as good as expected. The trajectory of the robot's CoM is compared model A, B, C, and D as shown in Figure 7, 8 and 9.

Table 2: The Simulation Results

	Distance		Rotation
Model	**Y_d (mm)**	**X_d (mm)**	**R_d(deg)**
A	86	-25	4.2
B	119	-5	3.1
C	49	-11	12.0
D	0	-64	23.2

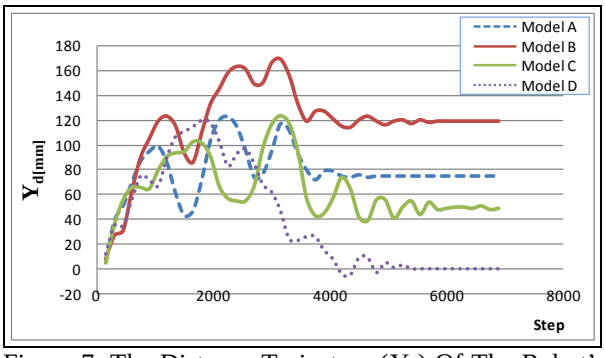


Figure 7: The Distance Trajectory (Y_d) Of The Robot's Center Of Mass (CoM)

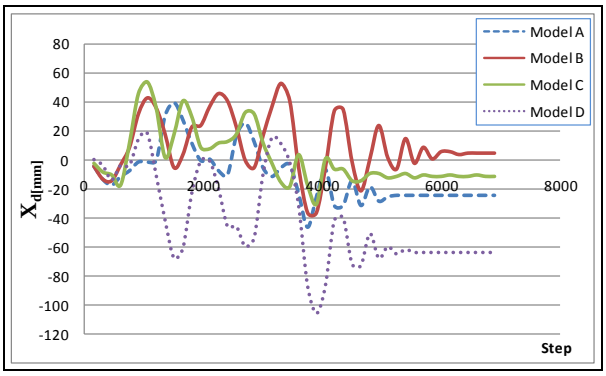


Figure 8: The Side Trajectory (X_d) Of The Robot's Center Of Mass (CoM)

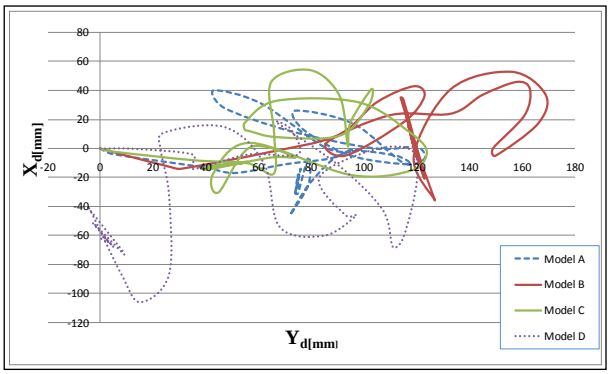


Figure 9: The Trajectory Of The Robot's Center Of Mass (CoM)

The trajectory of the robot's CoM in Figure. 7, 8 and 9. In model A, when we consider the walking step found that the trajectory of model B is larger and similar to human walking trajectory (Gait Analysis Based on Joint Moment 1997*)* but during into stable also to swing a lot. The other side, the trajectory of model C and model D the trajectories are small and walking is awkward.

Waveforms of the gait functions assigned to joints are compared model A, model B, model C and model D as shown in Figure 10-16.The widely of waveform of hip and ankle roll-joints $\theta_1(t)$ less changed and that similar to cosine function. However, the values of the other joints (Figure 11-16) that are distribute to adapt walking ability without falling. This is a one thing from the effects of the difference toe mechanisms in this study.

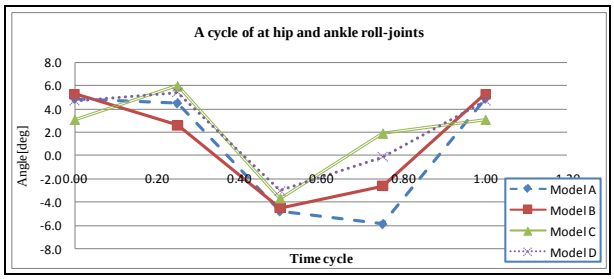


Figure 10: A Cycle Of Gait Function θ_1(t) @ (Hip And Ankle)

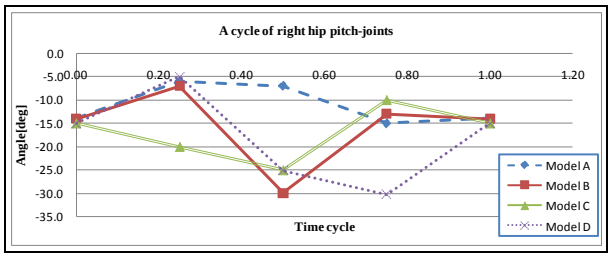


Figure 11: A Cycle Of Gait Function θ_2(t) @ (Hip Pitch-joint)

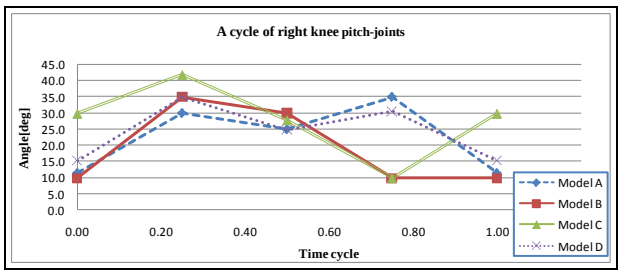


Figure 12: A Cycle Of Gait Function θ_3(t) @ (Knee Pitch-joint)

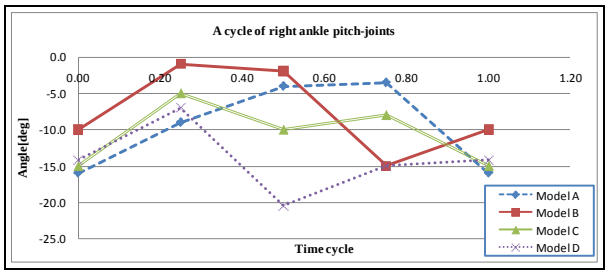


Figure 13: A Cycle Of Gait Function θ_4(t) @ (Ankle Pitch-joint)

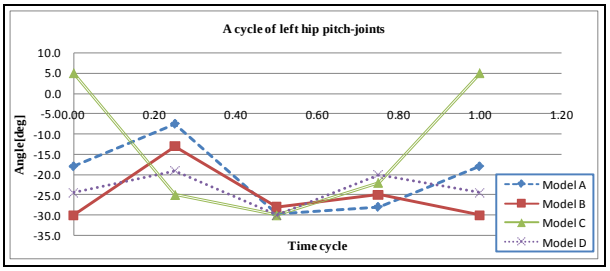


Figure 14: A Cycle Of Gait Function θ_5(t) @ (Hip Pitch-joint)

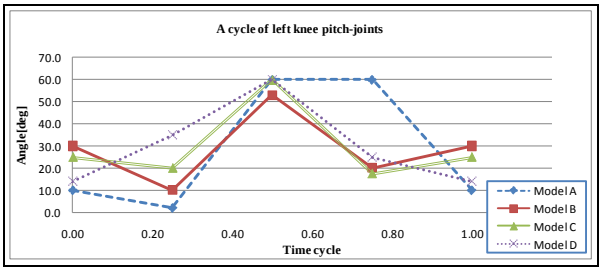


Figure 15: A Cycle Of Gait Function θ_6(t) @ (Knee Pitch-joint)

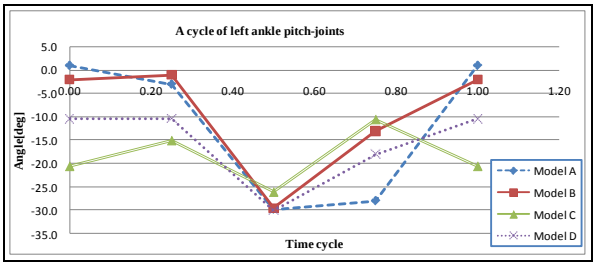


Figure 16: A Cycle Of Gait Function θ_7(t) @ (Ankle Pitch-joint)

4. CONCLUSION

We discussed a simulation base design (SBD) of small biped robot with toe mechanism walking on flat ground. The simulation of robot can walk. The simulations model A and B obtained a good gait pattern while model C and model D the results not as expected especially in model D which composites than other model and there have a strong impact. Therefore we will continue to experiment, improve and resolved for applied to real small biped robot.

As our next target, we will resolve in simulation of each model and used optimization method in order to achieve the objective function. Finally, applied to real small biped robot walking will be test as an experiment.

ACKNOWLEDGMENTS

This work was supported by Shibaura Institute of Technology Japan Scholarship. I would like to extend my heartfelt thanks to the Shibaura Institute of Technology, Professor. Hiroshi Hasegawa and everyone for their suggest and help in making this paper.

REFERENCES

Zhe, T., Zengqi, S., and Changji Z., 2006, GA Based Optimization for Humanoid Walking, *ICGST-ARAS journal, 5(2),* pp.1-10, June

Lingyun, H.,, Changjiu, Z., and Zengqi S., 2006, Biped Gait Optimization Using Spline Function Based Probability Model, *Proceeding of the 2006 IEEE International Conference on Robotics and Automation*, pp. 830-835, Orlando, Florida

Xiang, Y., Chung, H.J, Mathai. A., Rahmatalla, S., Kim, J., Marler, T., Beck, S., Yang, J., Arora, J.S., Abdel-malek, K., 2007, Optimization-based Dynamic Human Walking Prediction, *SAE Human Modeling for Design and Engineering Conference*, June 12–14, Seattle, Washington

Nandha, H., Jungwon, Y., Gabsoon, K., 2008, Gait Pattern Generation with Knee Stretch Maotion for Biped Robot using Toe and Heel Joints, *International Conference on Humanoid Robots*, pp. 265-270, Daejeon, Korea

Cheol. K.A., Min, Cheol, L., and Seok, J.G., 2003, Development of a biped robot with toes to improve gait pattern Advanced Intelligent Mechatronics, *IEEE/ASME International Conference, pp. 729-734*

Shuuji Kajita, Kenji Kaneko, Mitsuhasu Morisawa, Shinichiro Nakaoka and Hirohisa Hirukawa, 2007, ZMP-based Biped Running Enhanced by Toe Springs, *IEEE International Conference on Robotics and Automation Roma*, pp. 3963-3969, Italy

Moosavian, S.A.A., Alghooneh, M. Takhmar, A., 2007, Introducing a Cartesian Approach for Gate Planning and Control of Biped Robots and Implementing on Various Slopes, *Proceedings of the 2007 7th IEEE-RAS International Conference on Humanoid Robots*, pp. 545-550, Pittsburgh, United States of America.

Naoya, I., Hiroshi, H., 2007, Robust optimization under uncertain factors of environment for simple gait of biped robots, *Proceedings of the 6th EUROSIM Congress on Modeling and Simulation*, Ljubljana, Slovenia

Yu, Z., Ming, C.L., Dinesh, Manocha, Albertus, Hengrawan, Adiwahono., Chee-Meng, C., A Walking Pattern Generator for Biped Robots on Uneven Terrains, 2010, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, October 18-22, Taipei, Taiwan

Holland, J., 1992, Adaptation in Natural and Artificial Systems. *The University of Michigan 1975*, MIT Press

Chockalingam, N., Ashford, RL., 2007, Selected foot length ratios in a non-pathological sample. *Revista Internacional de Ciencias Podologícas* , 1(2), pp. 25-30

Masayuki, S., Naoya, I., and Hiroshi, H., 2009, The Simulation of Tiptoe mechanism for Biped Robot, *Asia Simulation Conference 2009 (JSST 2009),* October 7-9 , Shiga, Japan

Perry, J., 1992, Gait Analysis: Normal and Pathological Function, Slack Inc., N.J.,

The Clinical Gait Analysis Forum of Japan.1997, *Gait Analysis Based on Joint Moment. Ishiyaku publisher,* pp. 19-22, Japan

AUTHORS BIOGRAPHY

Krissana Nerakae received her BE and ME from Suranaree University of Technology, Thailand, in 2005 and 2009, She study in Graduate School of Mechanical Engineering in Shibaura Institute of Technology. Her research interests are optimization methods and robotics technology.

Hiroshi Hasegawa received his B.E. (1992) and M.E. (1994) from Shibaura Institute of Technology, Japan. He received PhD (1998) in Mechanical Engineering from Tokyo Institute of Technology, Japan. He has been working at Shibaura Institute of Technology, and currently is a Professor. He is member of JSME, ASME, JSST, JSCES and JSDE. His research interests include Creativity of Design, CAX (Computer Aided eXploration) and Biped Robot.

Simulation of House Prices for Improved Land Valuation

Terry Bossomaier[a], Md Zahidul Islam[a,b], Rod Duncan[a,c], Junbin Gao[a,d]
[a]Centre for Research in Complex Systems (CRiCS)
Charles Sturt University, Australia
[a]Email: tbossomaier@csu.edu.au [b]Email: zislam@csu.edu.au
[c]Email:rduncan@csu.edu.au [d]Email: jbgao@csu.edu.au

***ABSTRACT*—The housing market is an important part of most economies in the developed world, but is also a significant factor in setting land valuation. Thus good housing models serve multiple purposes. Many studies use powerful statistical techniques to study pricing, but these are not so effective for handling large scale social or demographic shifts. Agent based modelling (ABM) is more flexible and therefore this paper describes a housing market simulation using ABM. A standard economic model is used for estimating utility, combined with a second price auction model and a decision forest for linking house features to price estimates at the beginning. Simulations are presented for a range of market temperatures, revealing different levels of price inflation.**

***Index Terms*—house price, agent based model, decision tree**

1.. INTRODUCTION

One of the challenges facing government is the accurate valuation of land parcels, for a variety of purposes, including taxation. Many factors go into valuation, but previous sale values are of paramount importance. In some ways the land may itself become more or less valuable. It may be rezoned, allowing for development, and thus increase in value. On the other hand it may become polluted or some other environmental hasard, previously not recognised may become apparent. In this case the value would go down. In some situations the value may change as a result of change of use. If a petrol station is built on top of a piece of land where a petrol station had already existed, then the land value may not change much. But if the land, were to become some other sort of retail outlet, then contamination by fuel would decrease its value relative to another adjacent parcel which had not been exposed to fuel. Sometimes external factors such as a change of flight route over a land and development of new areas close by can also change the land price.

Accurate valuation is important for equity and general satisfaction with the process. But errors may lead to costly legal challenges, making accuracy financially desirable. This project is the first stage in a project to reduce errors in land valuation in NSW, with a particular focus on residential areas. Before tackling the complexity of the major cities, such as Sydney, we focus on a medium size (by Australian standards) regional city of around 40,000 inhabitants, Bathurst.

The Financial Review ranked suburbs and towns for housing retail price growth over the last three years. A wide variation from housing stagnation or decline to substantial growth is seen, sometimes with quite large variations in areas which are close together and similar in many respects. Bathurst came in the top 20 with an aggregate growth of 14.8% and the Council has provided support for the project. The city has several interesting dynamic features:

1) It has a significant manufacturing base, with nearly one third of jobs in manufacturing, which creates several residential foci. It also adds a special link to the performance of the economy in general.
2) It is 35km from the next nearest town, thus has a lot of land for potential development and the council strategic plan includes new areas zoned for development
3) It is Australia's oldest inland city, with a lot of distinctive old housing which is attractive to some buyers
4) It has a number of big schools of state-wide significance, also creating foci for residential development.

Accurate house price prediction is, of course, still an unsolved problem, presenting numerous difficulties. It is especially difficult because it operates at several different levels and timescales. At the coarsest timescale, the Economist regularly measures house price statistics across the developed world, using the ratio of capital investment to rental returns as an indicator. By such measures some areas, such as Hong Kong, and including Australia, have grossly overval-

ued housing. Outside of major financial catastrophes such indicators may imply long term stagnation or even decline in house prices.

Of more immediate concern are the changes in macro-economic conditions: jobs; interest rates and financing opportunities; and trends in consumer spending and taste (such as the balance between apartments and houses). Closely coupled to these factors is the availability of land for new development.

Lastly, there are factors relating to individual attributes of the houses themselves. Some are relatively static, such as proximity to schools, busy roads, while some can change rapidly as a result of owner investment in the property, or lack of it – some houses in Sydney, for example, are bought for the land only and allowed to become derelict.

These different considerations and scales have led to a very extensive range of modelling approaches. Statistical methods of many kinds (McMillen 2008) are popular at all levels, while artificial intelligence techniques, such as fuzzy logic (Lughofer, Trawiński, Trawiński, Kempa & Lasota 2011), have proved useful at the individual house level. Simulation using *Agent Based Models* (ABMs) are less widely used, but have two big advantages: they are capable of handling many disparate factors and timescales; and with suitable changes in parametrisation can easily be applied to different towns, cities and countries. Since they are intrinsically parallel, they can be scaled to very large numbers of agents with distributed computing.

The ABM comprises two parts: the housing market, comprising land availability, macro-economic factors and buyer/seller populations; and the models for buyer preferences. Magliocca et al. (Magliocca, McConnell, Walls & Safirova 2012, Magliocca, Safirova, McConnell & Walls 2011) provide a strong framework for the first of these, but there is no agreement on the best model for buyer choice. In fact the best data for buyer choice comes from statistical models and is not readily incorporated into an agent based model.

House prices can be estimated using a Fuzzy Delphi method where a number of experts are asked separately to estimate a house price based on a set of price factors such as green areas, seashore and grave yard (Damigos & Anyfantis 2011). An average of all the estimated prices of the property can then be used to re-estimate the price in a recursive manner until all individual estimates become very close to the average. The final average can then be considered as the estimated price. In order to establish a relationship between a set of price factors and the price of a property a group of buyers can be surveyed where they estimate the price of a property and also assign scores for various price factors of the property (Kusan, Aytekin & Ozdemir 2010). Records having scores on various price factors and an estimated price can be used as a training data set in order to build a classifier which is then used to estimate prices of other properties.

Determining buyer preferences is different. Hedonic modelling has been quite successful here (Bourassa, Hoesli & Zhang 2011), but one of the most interesting findings comes from quantile statistics. Quantile regression suits property valuation for a number of reasons. It produces a number of regression models supporting a decision maker with the ability to use alternate models in order to make a more accurate valuation (Narula, Wellington & Lewis 2012). Moreover, it transpires (Zietz, Zietz & Sirmans 2008) that housing preferences are dependent upon relative house value. So the size of plot of land and the number of bathrooms are more significant in higher priced homes. But the newness of the building is more significant in lower priced homes. Factors, such as commuting distance, are less dependent on value. Furthermore the distributions seem to vary according to locality. These results pertain to houses in the State of Utah in the USA and may not generalise to other parts of the world such as Australia. Thus the first stage of the Bathurst model is the quantile modelling of house price factors.

Given the impact of these different factors on house prices, we plan a soft computing model for agent behaviour. Fuzzy logic is a good choice here, since it allows linguistic expression of rules in a natural way. It has been used in a variety of housing studies. Together with the preference factors for the house and its immediate spatial location, macro-economic factors such as interest rates and limitations on loan amounts, general economic conditions and propensity to take financial risks in such conditions produce a set of fuzzy rules with a desirability value, D_n resulting for any given house. For the present paper a simpler approach is adopted, comparing a house feature vector with desired features based on socio-economic category (section 2.4.).

Given the buyer and seller properties for existing houses, new developments now have to be considered. The Magliocca et al. (Magliocca et al. 2011) model consists of farmer, f_i, developer, d_i and consumer agents, c_i, with effectively three markets: farmers selling to developers; and developers selling to consumers; and consumers selling to each other. However, the development market may be limited not only by the willingness of farmers to sell, but also council restrictions on the rezoning of land for housing development. Thus we restrict this first model to two

markets: developer and consumer.

The market then uses a Cobb-Douglas model (Magliocca et al. 2011) to determine the *utility*, U for each house, n

$$U(c_i, n) = (I_i - P_{ask})_i^{\alpha} D_i^{\beta} B_i^{\gamma} \quad (1)$$

where P is the asking price, and B the block size. From this the maximum price the consumer will pay is P_{max} is

$$P_{max} = \frac{I}{\beta_i + \gamma_i} \quad (2)$$

The offer price on any house will then be $\frac{P_{max} U}{U_{max}}$

To make the income and asking price of the same order, we use the annual payments on the property based on a fixed, simple (effective) interest rate, η, a non-variable mortgage term of M years, with M set to 20.

$$P = H(\eta + \frac{1}{M}) \quad (3)$$

where H is the actual price.

The maximum utility for each agent is set relative to *all* houses. Thus most of the time the available houses will have a lower utility and an agent will offer below their maximum price.

2.. THE FULL MODEL

In the first stage of the model our focus is on testing the underlying dynamics and thus several simplifications were made:

- Only the market for exisisng houses is modelled; there is a pool of buyers which is larger than the number of houses (representing incoming buyers and investors). The tricky issue of how to set rents is thus obviated. Thus each agent in the population may be either a buyer, seller or both.
- All housing factors (e.g. interior, proximity to non-housing spatial entities, such as schools, and shopping centres) stay constant.
- Each vendor/buyer has a separate preference, and family status that do not change during the simulation.
- The model is a small town – commuting costs are negligible.

All of these extensions will be introduced later. The software written is sufficiently flexible to embrace them easily.

A. The Market Process

In a typical housing market, sales occur by either private treaty or by auction. There is a considerable body of theory, on such processes, which allows some simplification.

Auctions have existed in many possible formats in different industries, countries and eras. Auctions may be first-price or second-price, where the winner of the auction pays respectively the highest bid or the second-highest bid submitted. Auctions may be sealed bid, where the buyer's bid is kept private, or open bid, where bids are public information. Open auctions may be English auctions, where bidders make successively higher bids, or Dutch auctions, where the auctioneer announces successively lower prices until a bidder enters a bid. For a general set of assumptions, the Theory of Revenue Equivalence stated first by Vickrey (Vickrey 1961) and then developed more formally by Myerson (Myerson 1981) establishes that the seller's expected revenue, and so also the buyer's expected price, will be independent of the design of the auction.

The following assumptions would satisfy the Theory of Revenue Equivalence.

1) The houses are auctioned individually.
2) The winner of the auction is the highest bidder. Only the winning bidder need contribute to the purchase.
3) The buyers are risk-neutral, and buyers' valuations are independently distributed and are known only to themselves.

If these assumptions are satisfied, then the expected sales price for the house will be the second-highest valuation of the buyers participating in the auction. This result holds true for all of the forms of auctions listed above.

Thus we adopt a single mechanism here which is a blind, second price auction. Each buyer makes an unseen bid. The buyer who makes the greatest offer gets the house, but pays the price of the *second-highest* bidder.

B. Overarching Framework

Every house put up for sale at each timestep gets auctioned once. If it does not sell, it is held over to the next week and re-auctioned.

1) A random number of vendors and buyers are added to the market at each timestep. Two parameters, the *market temperature,* ξ, and the *housing demand,* ζ control the number of sellers and buyers respectively. Thus there can be a glut of houses with no buyers, (high ξ, low ζ), a housing shortage (low ξ, high ζ) or a boom or

bust where both are high or low respectively. For the first simulations we set $\xi = \zeta$.
Vendors set a selling price. A buyer can also be a vendor. We make the implicit assumption that if someone buys another house without selling the first, they are acting as an investor.

2) Each buyer finds the house on the market with maximum utility, U, which determines their willingness to pay, W.
3) A second-price auction is run on each house, the ownership changes hands and vendor, buyer and house disappear from the market.

C. Houses

Houses will in the full model have a set of, say, H parameters, number of bedrooms, proximity to school etc, which will be determined from real data. To begin with we set up a random distribution of some kind. All houses have their price initialised at the start, as discussed in section 2.7.

Each parameter has a desirability value, d_i, between 0 and 1, formed as either a ratio of, say, the number of bedrooms, to the maximum number, or a qualitative value representing view, proximity to school and so on. Hedonic studies such as the one presented by Bourassa et al. (Bourassa et al. 2011) provide a starting point for the scope of such parameters.

D. Owners

All owners decide whether to buy or sell at random according to the market temperature. There is no refractory period on how soon an agent may re-enter the housing market after a sale.

Each buyer has their own values of α, β, and γ. But the calculation of the desirability of a house is formed as a dot product between the desirability vector, d_n and the weight vector w_c applied to it based upon the buyers social category. Current categories in use: families; single income, no kids (SINK); double income no kids (DINK); and retiree. For a family the weighting vector would emphasise number of bathrooms over say number of garages. The weight vector we use in our simulation is shown in Table I. A buyer belonging to the Family category has weight on House Size equal to 0.7 and weight on Mountain View equal to 0.1 meaning that they give more emphasise on house size than mountain view.

E. Vendors

Vendors set a reserve price based on what they paid for on the house to begin with (or its value at the start of the simulation). We could set a fixed increase percentage (and watch what happens as we change this), or we could measure the current inflation/deflation rate, or median value increase and produce an estimate from this. They have to sell if the reserve price is met.

F. Buyers

At each timestep each buyer bids on one house (the one with the maximum utility *of the houses on the market*). The buyer weighting vector and the house properties are combined to produce a desirability value (maybe just a weighted average).

G. Decision Tree Estimation of Initial Asking Price (Reserve Price)

In order to determine the utility function of a property a buyer uses the Cobb-Douglas equation as mentioned before in Eqn. 1. The buyer needs to know the asking price $P_{ask|n}$ of a property n in order to estimate the utility of the property for him/her. The vendors set a reserve price of the property based on what they paid. They also estimate the asking price based on their experience on the current market situation. That is, they make an estimate of their property price based on the recent sale prices of similar properties. In this study we consider that the reserve price and asking price of a property are same. We also assume that the reserve price of a property is a public knowledge as people know the recent sale prices of the similar properties.

In order to simulate the experience based asking price (reserve price) estimation by a vendor agent, we consider that we have access to a data set having various information on the properties and their recent sale prices. The data set can be considered as a two dimensional table where the records represent the properties and the columns (attributes) represent various information of the properties. One of the attributes is the Sale Price of a property and the other attributes are on various information such as Lot Size, Floor Size and Number of Bed Rooms. The attribute representing the sale price is considered as the label or class attribute, and all other attributes are considered as non-class attributes.

We then can build a decision tree (Quinlan 1993, Quinlan 1996, Islam 2012) or a decision forest (i.e. a set of decision trees) (Islam & Giggins 2011, Abellan & Masegosa 2009) from the data set in order to explore various logic rules for predicting/estimating the price of a property that we are interested in.

In this study we first generate a synthetic data set and then build a decision forest from the data set in order to learn various logic rules. The data set has 2000 records and 10 attributes, out of which nine are non-class attributes and the remaining one

Category	House Size	Year	Bedroom (number)	Bathroom (number)	Garage (number)	Landscape	Sprinkler	Mountain View
1	2	3	4	5	6	7	8	9
Family	0.7	0.3	0.9	0.9	0.5	0.1	0.2	0.1
Single Income	0.3	0.5	0.5	0.1	0.2	0.1	0.1	0.1
Double Income (no kids)	0.5	0.6	0.6	0.6	0.5	0.3	0.2	0.1
Retiree	0.1	0.9	0.3	0.1	0.3	0.7	0.9	0.3

TABLE I
WEIGHT VECTORS FOR DIFFERENT BUYER CATEGORIES

House Price ($1000)	Dist. (%)	House Size (sqm)	Lot Size (sqm)	Year	Bed	Bath	Garage	Land-scape	Sprink-ler	Mount. View
1	2	3	4	5	6	7	8	9	10	11
50-100	5	45-90	90-120	1960-2010	1	1	0	0	0	0
100-200	10	80-170	100-400	1960-2010	1-2	1	0	0-1	0	0
200-300	35	150-220	350-650	1980-2012	2-4	1-2	1	0-1	0	0
300-400	25	200-280	500-1000	1995-2012	3-5	1-3	1-2	1	0-1 (40%)	0-1 (40%)
400-500	15	250-390	700-1000	1990-2012	4-6	2-3	2-3	1	0-1 (70%)	0-1 (40%)
500-600	5	350-450	800-1000	1985-2012	4-6	2-3	2-3	1	1	0-1 (60%)
600-700	2	420-640	700-1200	1980-2012	5-7	3-4	2-3	1	1	0-1 (70%)
700-800	1	600-700	800-1000	1980-2012	5-7	3-4	2-4	1	1	0-1 (80%)
800-900	1	650-750	900-1200	1980-2012	5-7	4-5	2-4	1	1	0-1 (80%)
900-1000	.5	700-900	800-1200	1975-2012	6-7	4-5	2-4	1	1	0-1 (80%)
1000-1100	.2	700-950	800-1400	1975-2012	6-8	4-6	2-4	1	1	0-1 (80%)
1100-1200	.2	700-950	800-1500	1975-2012	6-9	4-6	2-4	1	1	0-1 (80%)
1200-1300	.034	700-1000	900-1600	1975-2012	6-9	4-6	2-4	1	1	0-1 (80%)
1300-1400	.033	800-1000	1000-1700	1975-2012	6-9	4-6	2-4	1	1	0-1 (80%)
1400-1500	.033	800-1100	1200-1800	1975-2012	6-9	4-6	2-4	1	1	0-1 (80%)

TABLE II
SYNTHETIC DATA SET GENERATION RULES

is the class attribute ("Sale Price"). The non-class attributes are "House Size (sqm)", "Lot Size (sqm)", "Year of building", "Number of Bed Room", "Number of Bath Room", "Number of Garage", "Landscape", "Sprinkler", and "Mountain View".

Table II presents the rules that we use to generate the synthetic data set. There are altogether 11 columns and 17 rows in the table. The first column shows ranges of prices, while the second column shows the distribution of the prices in the data set. For example, in the synthetic data set there are 5% records within the price range of $50K - $100K, and 35% records within the price range of $200K - $300K. Column 3 to 11 show the ranges of various other attribute values corresponding to a price range. For example, for the price range of $50K - $100K the lot size varies between 90 sqm to 120 sqm.

We generate a record of the synthetic data set using the following steps.

- Step 1: Generate Property Price.
  We first generate a property price following the distribution of price ranges.

- Step 2: Generate Lot Size.
  Considering the high correlation between a property price and lot size (Zietz et al. 2008) we next generate the lot size within a range as shown in Table II. The probability distribution within a range of lot size is uniform.
- Step 3: Calculate House Size.
  We next calculate the House Size (i.e. Floor Size) using $P = 1000 \times h + 150 \times l$, where P, h and l are the house price, house size and lot size, respectively. If the calculated house size falls outside the upper limit of the range of House Size as shown in Table II then we consider the value equal to the upper limit of the range. We do the same for the lower limit of the range as well.
- Step 4: Generate other attribute values.
  We also generate other attribute values based on the rules as shown in Table II following a price range. For example, the number of bed rooms for a property having price within the range of \$500K and \$600K can be anything between 4 and 6 with uniform probability distribution.
  The domain of the attributes Landscape, Sprinkler and Mountain View is $\{0,1\}$, where 0 means no landscaping/sprinkler and 1 means the existence of landscaping/sprinklers. A percentage value within parenthesis indicates the probability of having the feature. For example, there is a 40% probability of having sprinklers for a property within the price range of \$300K and \$400K (see Table II).
- Step 5: Check Floor Size.
  Considering an average bed room size is 10 sqm, bath room size is 4.5 sqm, and garage size is 20 sqm we calculate the house size (using the generated number of bedroom, bathroom etc.) in order to make sure that the calculated house size is not greater than the house size estimated in Step 3. If that is not the case then we reduce the number of a bedroom, and/or a garage as necessary.

From the synthetic data set we build a decision forest having four trees using SysFor algorithm (Islam & Giggins 2011). Figure 1 shows the first tree of the forest. In this study we use the first tree to estimate the initial reserve price of a property. For example, the logic rule for Leaf 1 of the tree indicates that if there is no landscaping, no garage and the lot size is ≤ 120 sqm then the property price should be within \$50K to \$100K range. As the starting reserve price of a property we estimate the lower limit of the range.

H. Implementation

The population of houses was initialised with attributes randomly generated according to the distributions from the synthetic data (i.e. 5% with prices from 50-100, size 45-90 and son on). The houses are then randomly assigned among the agents. Agents are randomly assigned a type from the four categories SINK, DINK, Family (Fam) and Retiree (Ret). Incomes are determined using a base of 20 with a random exponent from 1.0 to 1.75, yielding incomes in a range from 20 to 189 - DINKs and Families are given two incomes.

A fixed number of buyers (10) and sellers (5) is used at the start of the simulation from a total of 1000 agents – these are changeable by parameters. There are parameters for the market temperature for both buying (i.e. housing demand) and selling (5 each). These numbers are divided by the scaling factor (1000) to yield the chance that an agent enters the market. With the current numbers, at every time step, an agent has a 5/1000 (0.5%) chance of becoming a buyer and a 5/1000 chance of becoming a seller.

The second-price auction model was implemented making the assumption that if there is no second price (i.e. only one bidder), or the second price is below the reserve, then the price paid is the reserve price. If the reserve price isn't met by buyers at the auction, the reserve price is reduced by 10% for the next auction. If a house attracts no bidders, no auction is deemed to have taken place, so it remains in the available pool for the next time step.

Values for the utility function were set assuming that people spend at least half their income on housing, so α, β and γ were each set to a random value from 0.6 to 1.0 (yields P_{max} equal to 50-83% of income, I).

For desirability, the weightings are given in Table I.

3.. ILLUSTRATIVE SIMULATIONS

The model was implemented in RePast Symphony Version 2. Simulations were run for 1500 time steps, approximately 5 years with a timestep being one day. All houses being put up for sale are auctioned once and only once during a timestep. Figure 2 shows the price as a function of time. A natural inflation effect is observed. The growth in price is more rapid at the beginning of the simulation, suggesting a settling period where houses undervalued at the beginning rapidly approach a more market driven value.

Figure 3 shows the size of the market as judged by number of agents for different market temperatures (2,5,8 as before). The lower market temperature produces more of an "equilibrium" with a matched numbers of buyer and sellers. The higher market

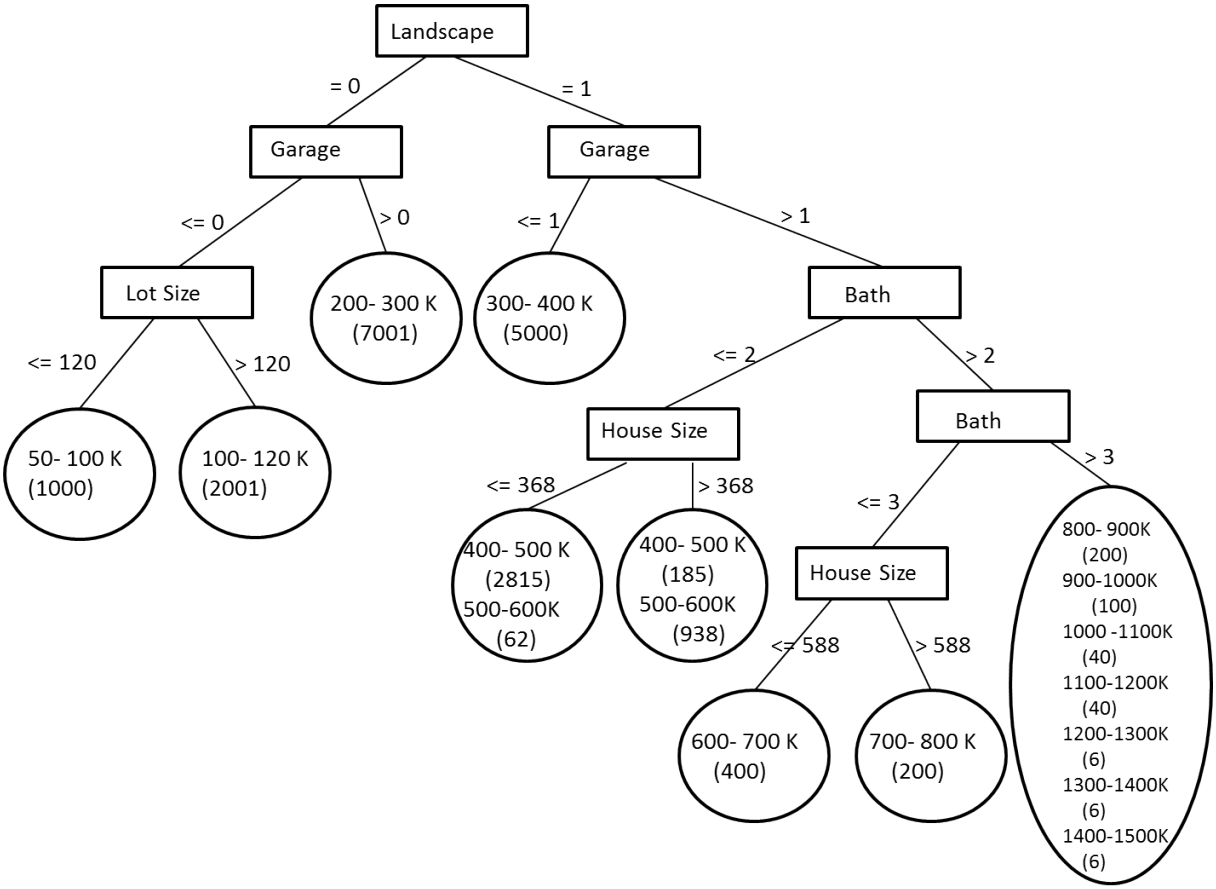


Fig. 1. A Decision Tree Obtained from the Synthetic Data Set

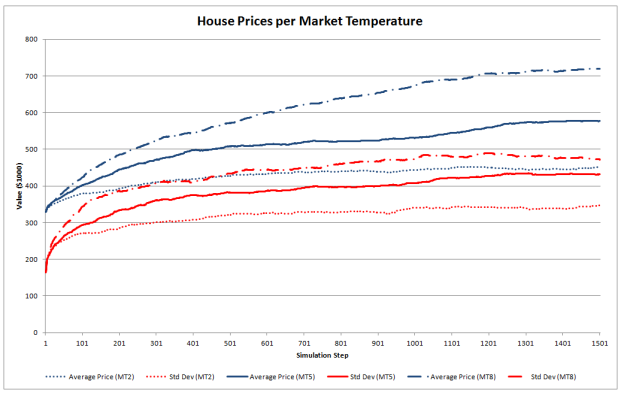


Fig. 2. Growth in house prices with time for three different market temperatures, MT2, MT5 and MT8, corresponding to temperature values of 2,5 and 8 (0.2% chance of selling, see section 2.8.

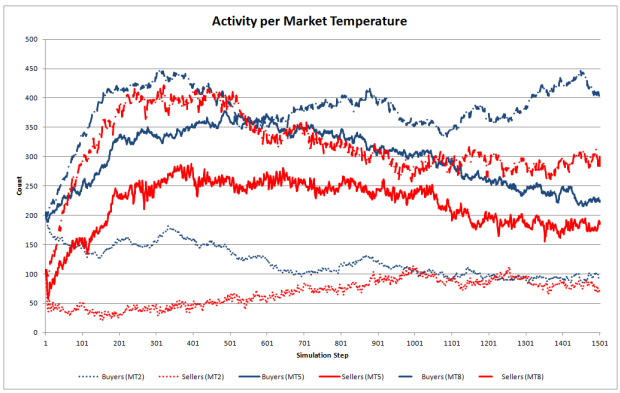


Fig. 3. Number of buyers and sellers in the market for three market temperatures MT2, MT5 and MT8, corresponding to temperature values of 2,5 and 8 (0.2% chance of selling, see section 2.8.

temperatures create an excess demand, with more buyers than sellers, consistent with the inflationary trend observed in figure 2.

4.. CONCLUSIONS AND FURTHER WORK

Agent based modelling provides a way to project beyond hedonic modelling and study house price growth under various scenarios. From house price growth, land valuation can be predicted and adjusted.

A decision tree approach has been introduced to estimate house price ranges. The next stage of the project is to determine the decision tree from real data, use census data to determine the actual probability distributions for socio-economic categories to parameterise the agents.

An interesting change in Bathurst demographics has been foreshadowed. One of the mining companies is set to dramatically increase its gold mining activity, with the expectation of around 3,000 new jobs. As almost 10% of the current population, this will present challenges to the housing market. The challenge for our simulation will be to see if we can successfully predict the outcome.

REFERENCES

Abellan, J. & Masegosa, A. R. 2009. An ensemble method using credal decision trees, *European Journal of Operational Research* .

Bourassa, S., Hoesli, M. Scognamiglio, D. & Zhang, S. 2011. Land leverage and house prices, *Regional Science and Urban Economics* 41: 134–144.

Damigos, D. & Anyfantis, F. 2011. The value of view through the eyes of real estate experts: A fuzzy delphi approach, *Landscape and Urban Planning* 101(2): 171–178.

Islam, M. Z. 2012. Explore: A novel decision tree classification algorithm, *LNCS* 6121: 55–71.

Islam, M. Z. & Giggins, H. 2011. Knowledge discovery through sysfor: A systematically developed forest of multiple trees, *in* 'Proc. of the Ninth Australasian Data Mining Conference (AusDM 11', ACS, pp. 195–204.
URL: *http://crpit.com/vol121.html*

Kusan, H., Aytekin, O. & Ozdemir, L. 2010. The use of fuzzy logic in predicting house selling price, *Expert Systems with Applications* 37(3): 1808–1813.

Lughofer, E., Trawiński, B., Trawiński, K., Kempa, O. & Lasota, T. 2011. On employing fuzzy modeling algorithms for the valuation of residential premises, *Information Sciences* 181: 5123–5142.

Magliocca, N., McConnell, V., Walls, M. & Safirova, E. 2012. Zoning on the urban fringe: Results from a new approach to modeling and housing markets, *Regional Science and Urban Economics* 42: 198–210.

Magliocca, N., Safirova, E., McConnell, V. & Walls, M. 2011. An economic agent-based model of coupled housing and land markets (chalms), *Computers, Environment and Urban Systems* 35: 183–191.

McMillen, D. 2008. Changes in the distribution of house prices over time: Structural characteristics, neighborhood, or coefficients?, *J. Urban Economics* 64: 573–589.

Myerson, R. 1981. Optimal auction design, *Mathematics of Operations Research* 6: 58–73.

Narula, S. C., Wellington, J. F. & Lewis, S. A. 2012. Valuating residential real estate using parametric programming, *European Journal of Operational Research* 217: 120–128.

Quinlan, J. R. 1993. *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, California, USA.

Quinlan, J. R. 1996. Improved use of continuous attributes in c4.5, *Journal of Artificial Intelligence Research* 4: 77–90.

Vickrey, W. 1961. Counterspeculation, auctions and competitive sealed tenders, *Journal of Finance* 16: 8–37.

Zietz, J., Zietz, E. N. & Sirmans, G. S. 2008. Determinants of house prices: A quantile regression approach, *Journal of Real Estate Finance Economics* 37: 317–333.

UNSUPERVISED ALGORITHM FOR RETRIEVING CHARACTERISTIC PATTERNS FROM TIME-WARPED DATA COLLECTIONS

Tomáš Kocyan, Jan Martinovič, Michal Podhorányi, Ivo Vondrák

VŠB - Technical University of Ostrava
IT4Innovations
17. listopadu 15/2172, 708 33 Ostrava, Czech Republic

tomas.kocyan@vsb.cz, jan.martinovic@vsb.cz, michal.podhoranyi@vsb.cz, ivo.vondrak@vsb.cz

ABSTRACT
This paper discusses possibilities of using the Voting Experts algorithm enhanced by the Dynamic Time Warping (DTW) method for improving performance of Case-Based Reasoning (CBR) methodology used with time-warped data collections. CBR, in general, is the process of solving new problems based on the solutions of similar past problems. Success of this methodology strongly depends on the ability to find similar past situations. Searching these similar situations in data collections with components generated in equidistant time and in finite number of levels is now a trivial task. The problem arises for data collections that are subject to different types of distortions (e.g. measurement of natural phenomena such as precipitations, measured discharge volume etc.). The main goal of this paper is to provide suitable mechanism for retrieving typical patterns from distorted time series and thus improve the usability of CBR.

Keywords: segmenting, case-based reasoning, voting experts, dynamic time warping, time series

1. INTRODUCTION

Many types of existing collections often contain repeating sequences which could be called as patterns (Theodoridis and Koutroumbas 2006). If these patterns are recognized they can be for instance used in data compression or for prediction. Extraction of these patterns from data collections with components generated in equidistant time and in finite number of levels is now a trivial task. The problem arises for data collections that are subject to different types of distortions in all axes. This paper will focus on processing of measured river discharge volume, especially on finding typical patterns in this data collection. In our future research, such found patterns will be used for simulation of the rainfall runoff process and for prediction of discharge volumes in basin's outlet cross section.

River discharge is an important value and is frequently monitored along many major rivers and streams in the world. River discharge is the amount of water that flows through the river bed. There are mathematical formulas to measure the volume of water that flows through a certain section of the river at a given time. River discharge is very variable value. The existence of an increasing or decreasing trend in a river discharge time series data can be induced by the change in the climatic factors such as temperature (e.g. Comani 1987) or precipitation (Maheras and Kolyva-Machera 1990). Variability in river runoff has been reported in, for example, Giakoumakis and Baloutsos (1997). For this reason, it is almost impossible to use conventional methods for patterns extraction and it is necessary to use an alternative approach (Fanta, Zaake and Kachroo 2001). Found patterns will be presented in river hydrograph, which is a graph that shows the change in river discharge over time. Different river catchments produce different shapes of hydrograph (river discharge variability).

2. CASE-BASED REASONING

Case-Based Reasoning methodology belongs to a group of artificial intelligence methods and it can be simply described as a process of solving new problems based on the solutions of similar past problems. Each of these cases consists of a specific problem, its solution, and the way it was achieved. For purposes of computer reasoning the CBR has been formalized as a four-step abstract process (Watson 1997):

- Retrieve - Retrieve the most similar cases from case database that are relevant to solving it.
- Reuse - Map the found solutions from the previous cases to the new problem.
- Revise - Test the derived solution in the real conditions and, if necessary, revise it.
- Retain - After the solution has been successfully adapted to the target problem, the resulting experience is stored as a new case in memory and can be used for prediction in the next cycle.

For achieving the best results using Case-Based Reasoning methodology, it is necessary to successfully manage the first step (Retrieve). Many supervised and unsupervised methods for looking for patterns and similar situations exist, but the most of them have a common problem: they cannot handle searching for patterns of different lengths and they are not resistant to distortion. The Voting Experts algorithm is one of the algorithms that are able to handle these problems.

3. VOTING EXPERTS

The Voting Expert Algorithm is a domain-independent unsupervised algorithm for segmenting categorical time series into meaningful episodes. It was first presented by Cohen and Adams (2001). Since this introduction, the algorithm has been extended and improved in many ways, but the main idea is always the same. The basic Voting Experts algorithm is based on the simple hypothesis that natural breaks in a sequence are usually accompanied by two statistical indicators (Cohen, Adams, and Heeringa 2007): low internal entropy of episode and high boundary entropy between episodes.

The basic Voting Experts algorithm consists of the following three main steps:

- Build an nGram tree from the input, calculate statistics for each node of this tree (internal and boundary entropy) and standardize these values in nodes at the same depth.
- Pass a sliding window of length *n* over the input and let experts vote. Each of the experts has its own point of view on current context (current content of the sliding window) and votes for the best location for the split. The first expert votes for locations with the highest boundary entropy, the second expert votes for locations with a minimal sum of internal split entropy. By this way, the votes are counted for each location in the input.
- Look for local maximums which overcome selected threshold. These points are adepts for a split of sequence.

For detailed explanation of each of mentioned steps see (Cohen, Adams, and Heeringa 2007). Tests showed that the algorithm is able to segment selected input into meaningful episodes successfully. It was tested in many domains of interest, such as looking for words in a text (Cohen and Adams 2001) or segmenting of speech record (Miller, Wong, and Stoytchev 2009).

There are several ways how to improve the basic Voting Experts algorithm. Simply we can divide these improvements into the two main groups. On the one hand, a custom "expert" can be added to voting process (for example Markov Expert by Cheng and Mitzenmacher (2005)) and receive additional point of view on your input. On the other hand, there are methods based on repeated or hierarchical segmenting of the input (Miller and Stoytchev 2008, Hewlett and Cohen 2009).

One of the simplest ways how to slightly improve performance of segmenting is two-way passing of the sliding window. It means using classic voting algorithm supplemented by segmenting of reversed input. This idea was outlined in (Hewlett and Cohen 2009) which showed the way to make high-precision cut points by selection of higher values of the threshold. Additionally, reversing the corpus and segmenting the reversed input with Voting Experts generates a different set of backward cut points. The subsequent intersection of sets of cut points offers high precision segmenting. However, on the other hand, this high precision causes loss of recall.

4. VOTING EXPERTS POST-PROCESS

Proposed solution for Voting Experts improvement takes the task of using Dynamic Time Warping algorithm (introduced below) and high precision cuts as a starting point for looking for typical patterns located in the input. The basic idea is to refine the sparse set of high precision cuts into regular sequences as correctly as possible. The mentioned refinement will be done by several types of post-processing methods and the results will be compared. Methods will differ, but they share a common principle (as shown in Figure 1). If there are high precision cuts in the input (such as cuts A, B, C and D in Figure 1) and if the shorter sequence (bounded by cuts C and D) is a subsequence of the longer one (bounded by cuts A and B), we can deduce new boundaries E and F by projecting the boundaries of common subsequence to the longer sequence. In this very simplified example the sequences were composed by definite number of values and limited length, so the evaluation is quite straightforward.

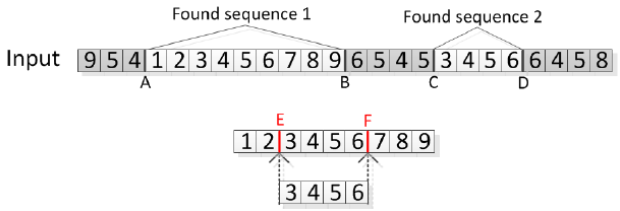


Figure 1: Refinement of sequences

In the case of application of previously mentioned process on distorted data, it is necessary to slightly modify it. Typical episodes of measurement of natural phenomena (such as precipitations, measured discharge volume etc.) are unfortunately subject to distortion in both time and value axes. For this reason, it is necessary to find out suitable mechanism that is able to deal with this deformation. The Dynamic Time Warping (DTW) algorithm can be used for this purpose.

DTW is a technique to find an optimal alignment between two given sequences under certain restrictions (Muller 2007). The sequences are warped in a nonlinear fashion to match each other. First DTW was used for comparing two different speech patterns in automatic speech recognition. In information retrieval it has been successfully applied to dealing with time deformations and different speeds associated with time-dependent data. For our purposes, the DTW algorithm will be used as a tool for finding the longest common subsequence of two sequences.

4.1. Searching the mutual subsequences

Despite the fact that the DTW has its own modification for searching subsequences, it works perfectly only in a case of searching exact pattern in some signal database. This case is demonstrated in Figure 1, where sequence

s_1 = '3456' exactly matches the corresponding subsequence in sequence s_2 = '12**3456**789'. However, in real situations exact patterns are not available because they are surrounded by additional values (Figure 2a), or even repeated several times in the sequence (Figure 2b). Unfortunately, the basic DTW is not able to handle these situations and it fails or returns only a single occurrence of the pattern. To deal with this type of situations, own DTW modification was created. This modification is able to find the longest common time warped subsequences under selected restrictions. Because this paper's topic is not focused on modifications of the DTW, our proposed solution will be presented very briefly.

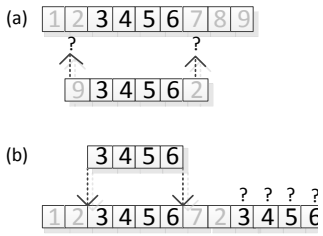


Figure 2: Basic DTW inaccuracies

The modification is based on the DTW's distance matrix (Muller 2007), in which we want to find the longest common warping paths with the lowest path cost as much as possible. The searching of warping paths starts by searching common column's and row's minimums. Formally, we are looking for points:

$$p(x,y): x = \min(c_x) \wedge y = \min(r_y), \quad (1)$$

where (x, y) are indices (coordinates) of minimum point, c_x is the x row and the r_y is the y column of distance matrix. Consider two sequences s_1 = *{4, 2, 3, 4, 5, 7}* and s_2 = *{5, 2, 3, 4, 8, 8}*, so their distance matrix will look like in Figure 3a.

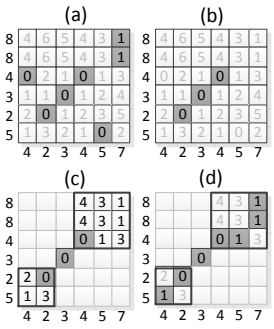


Figure 3: Distance matrices

Then, for each of these minimum cost points of the distance matrix:

1. If the actual point is contained in any other warping path, ignore this point and continue with the next one.
2. If there are some other local minimums in below, left or below left locations in the distance matrix, move there.
3. If the total cost of warping path is not overcome, add the current location to the warping path and continue with the point 2 until other minimums are available.
4. Now, no other close local minimums are available and the warping path for example looks like in the Figure 3b. However, if the maximum cost of warping is known, the warping path can be extended until the maximum path cost is reached. Extension of warping path means adding as much points as possible to the warping path from the first point of warping path towards the upper right corner of matrix, respectively from the last point of warping path towards the lower left corner of matrix.
5. For purpose of searching optimal extensions, classic DTW is used. Searching for extensions is actually searching of optimal warping paths in the new two submatrices showed in Figure 3c.
6. Once these paths are found, the original warping path is alternately extended in both directions until it reaches the maximum (see Figure 3d).

In paragraphs above, the cost of warping path was mentioned for several times. It is usual to specify the maximum cost of warping path as a constant, which cannot be overcome. However, tests showed that it is much suitable to specify the maximum cost of warping path as an average cost per warping point, because it is more immune to measurement errors and other distortions.

4.2. Post-Process Algorithm

The main idea of the Voting Experts DTW post-process is summarized into the following steps:

1. First of all, the high precision (but not complete) cuts are created by splitting the input with high level of threshold by the Two-Way Voting Experts method.
2. Let's suppose that there are m unique sequences which have been created according to the cuts from step 1.
3. A $m \times m$ distance matrix is build.
4. For each pair in this matrix, where the length of sequence s_1 is bigger than length of sequence s_2:
   (a) The optimal mapping of shorter sequence s_2 to the longer sequence s_1 is found by using DTW modified.
   (b) If the mapping cost does not overcome selected threshold, the longest sequence s_1 stores the shorter sequence s_2 into its own list of similar sequences. By this way,

every sequence gets its own list of the most similar shorter sequences.

(c) Each of the shorter sequences points to positions in the longer sequence, where it should be split. Because there is usually more than one similar shorter sequence, it is pointed to several locations whereas many of these locations are duplicated. For this reason, the votes are collected into internal vote storage.

(d) After these votes are collected, the local maximums are detected. These places are suggested as new cuts in original input.

5. The granted votes from step 4(d) are summed with votes of frequency and entropy experts in the input. Subsequently, the local maximums of votes are searched again. The cuts are made in locations where the number of granted votes is higher than the specified threshold.
6. Algorithm ends or it can continue with step 2 for further refinement.
7. After each post-process iteration, all found patterns can be received from the 4(b) step. Actually, the found patterns are groups of similar chunks.

For our algorithm improvement, several variants of each particular step were proposed and then their influences were tested on final results.

4.3. Post-Process Evaluation

Typical test of algorithm's verification of Voting Experts algorithm's performance is searching words in continuous text. In this text, spaces and punctuations (dots, dashes, new lines etc.) are removed and the goal of the algorithm is to put spaces back into correct places. Because the correct placement of spaces in the original text is known, it is very easy to quantify the algorithm's accuracy. For objective comparison we tested our proposed alternative approach in natural domain of Voting Experts – splitting continuous text (Kocyan, Martinovič, Kuchař, and Dvorský 2012). Results of the best algorithm configuration applied on various texts overcame almost all basic algorithm monitored quality indicators. Recall was improved by up to 18% and overal F-measure quality reached improvement about 5.5% in average. It is evident that our proposed alternative approach works with categorical data. A new challenge is to modify the algorithm for processing quantitative time series and to design suitable mechanism for Case-Base Reasoning's Retrieve step. And this is the main goal of this paper.

4.4. Test Collection

A very important aspect of every study is data collection. The data used in this paper consist of two parts: First of all, an artificial collection will be created and then it will be distorted. Using this collection, an algorithm will be explained, evaluated and the results will be presented. The second collection was obtained from the U.S. database - USGS (U.S. Geological Survey) Water Data for the Nation (section – surface data). Online access to this data is organized around the categories. We used discharge data (CFS - cubic feet per second) from stations located on the main rivers in the U.S.A. (years from 1986 to 2007, 30 min. step). This data was used for practical demonstration of proposed algorithm.

5. EXPERIMENTS

First of all, the algorithm was tested on artificial collection. This collection was generated from patterns displayed in Figure 4, which were randomly repeated to the total input's length of 200 elements.

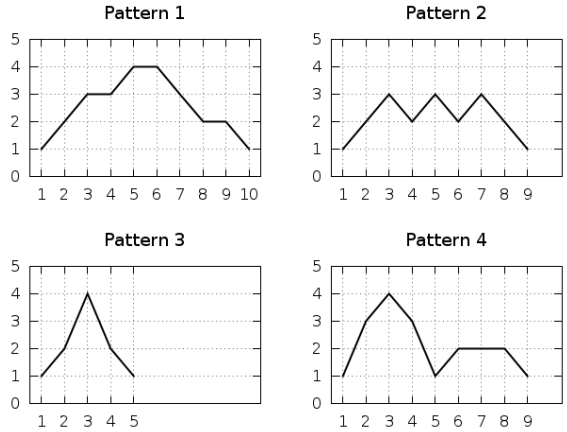


Figure 4: Set of Artificial Patterns

On this data collection, the original Voting Experts algorithm was executed with many configurations and the best result was taken. Then, with the same data, the proposed algorithm was tested in many configurations and also the best result was taken.

For the evaluation of proposed algorithm's performance, precision and recall coefficients were defined. Precision coefficient and recall coefficient rank among the most often used for the methods that are able to provide relevant documents in the information system. The precision coefficient is understood as the ratio of the amount of relevant documents returned to the entire number of returned documents. Recall represents the ratio of the amount of relevant documents returned to a given query to the entire amount of documents relevant to this query. In our case, the precision coefficient will be understood as the ratio of the amount of correct spaces induced by algorithm to the entire number of induced spaces. Recall will represent the ratio of the amount of correct induced spaces to the entire amount of spaces in input.

The basic algorithm reaches the best results with configuration listed in Table 1:

Table 1: The Best Configuration of Original VE

Parameter	Value
Sliding Window Size	10
Threshold for Cuts	4

To get high precision cuts for post-processing, the value of threshold had to be increased and followed by particular count of post-process cycles.

The best configuration of the post-process approach is shown in Table 2 and the comparison of both versions of algorithm's success is presented in Table 3.

Table 2: The Best Configuration of VE Post-Process

Parameter	Value
Sliding Window Size	10
Threshold for Cuts	9
Post-Process Cycles	13
Post-Process Threshold	8

Table 3: Evaluation of algorithms' success

Algorithm	Precision	Recall	F-Measure
Basic Voting Experts	0.71	0.68	0.70
VE with Post-Process	1	0.86	0.93

It is evident that proposed algorithm significantly outperforms the basic version of Voting Experts in all evaluation indicators and increases its performance.

Table 4 shows the progress of all indicators after each post-process cycle. The fact that the precision indicator still takes the value of 1 means that algorithm did not make any wrong cuts. Graphical representation of the table's values is in Figure 5.

Table 4: Progress of Evaluation Indicators

Post-Process Cycle	Precision	Recall	F-Measure
0 (VE output)	1	0.10	0.19
1	1	0.24	0.47
2	1	0.31	0.62
3	1	0.45	0.71
4	1	0.55	0.77
5	1	0.62	0.77
6	1	0.62	0.77
7	1	0.62	0.77
8	1	0.72	0.84
9	1	0.79	0.88
10	1	0.79	0.88
11	1	0.76	0.86
12	1	0.76	0.86
13	1	0.86	0.93

The second experiment was executed on distorted version of the previous data collection. Applied distortion randomly manipulated with length of patterns (elements of patterns were repeated or omitted) and its amplitude. Example of various distorted versions of one particular pattern can be seen in Figure 6.

The results of the experiment are shown in Table 5. It is evident that the original Voting Experts algorithm cannot work with distorted data and its accuracy rapidly decreases. Our proposed solution's accuracy decreases too, however it still provides better result with distorted input than the basic version on regular input. Examples of found patterns are shown in Figure 7.

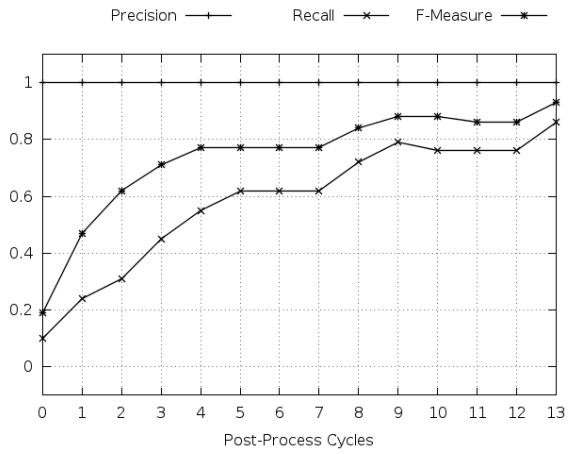


Figure 5: Progress of Evaluation Indicators

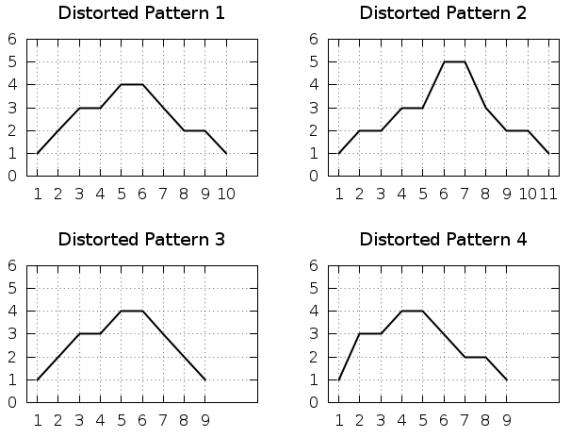


Figure 6: Example of Distorted Artificial Pattern

Table 5: Evaluation of algorithms using distorted input

Algorithm	Precision	Recall	F-Measure
Basic Voting Experts	0.48	0.58	0.53
VE with Post-Process	0.91	0.72	0.81

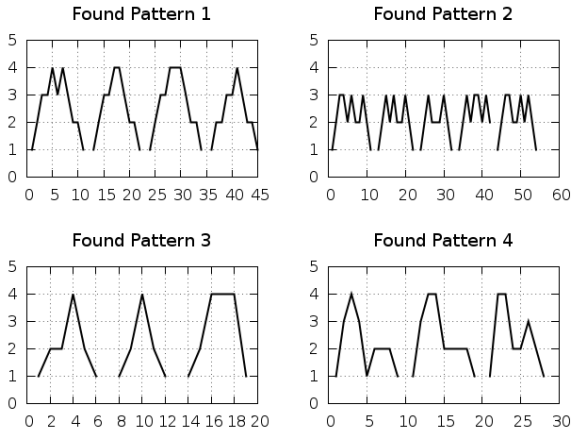


Figure 7: Example of Found Distorted Patterns

The last experiment was executed with river data described in paragraph 4.1. For this data collection, the exact real boundaries are not known, so it is impossible to numerically evaluate success of the algorithm. For this reason, the results of the algorithm will be

evaluated only visually and the examples of found patterns will be presented.

Before this data could be processed, it had to be converted to format suitable for the Voting Experts algorithm. For this reason, the data was encoded by SAX (Symbolic Aggregate approXimation) described in (Patel, Keogh, Lin, and Lonardi, 2002) and its dimension was reduced. Then, the algorithm was executed with the same configuration as in the example above. The found patterns are shown in Figure 8.

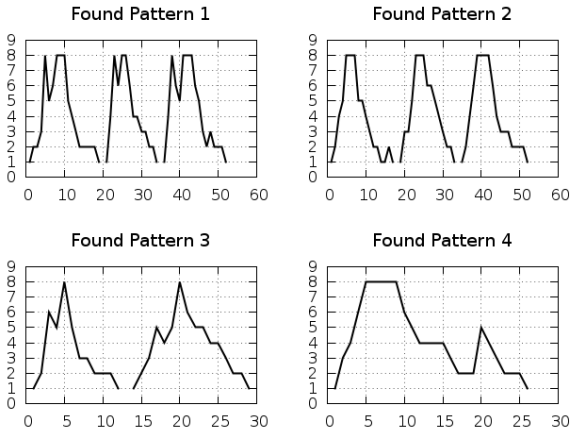


Figure 8: Example of Found River Patterns

6. CONCLUSION

The main goal of this paper was to provide suitable mechanism for searching characteristic patterns in order to improve the performance of Case-Based Reasoning. Found patterns can also be used for time series indexing or compressing.

Tests showed that proposed algorithm is able to successfully split input data into the meaningful episodes and then find characteristic patterns in both regular and distorted collections. Proposed alternative approach outperforms the original Voting Experts algorithm in all evaluation indicators (Precision, Recall, F-Measure) and moreover, it is more immune to working with distorted inputs.

Future research will be focused on optimizing and improving proposed algorithm's performance, especially on automatic settings of configuration's parameters, which have to be set manually for now. Additionally, we will focus on searching the universal encoding algorithms for transforming general time series into the format suitable for Voting Experts and on the modification of DTW for receiving characteristic patterns from a group of similar episodes.

ACKNOWLEDGMENTS

This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070) and by the grant project of the Technology Agency of the Czech Republic under number TA01021374, titled "Nove technologie ochrany zivotniho prostredi pred negativnimi nasledky pohybujicich se prírodnich hmot" (New Technologies for Environmental Protection against Negative Effects of Natural Mass Transfer).

REFERENCES

Cheng, J., Mitzenmacher, M., 2005. Markov Experts. *Proceedings of the Data Compression Conference (DCC)*.

Cohen, P. R., Adams, N., 2001. An Algorithm for Segmenting Categorical Time Series into Meaningful Episodes, *Proceedings of the Fourth Symposium on Intelligent Data Analysis, Lecture Notes in Computer Science*.

Cohen, P. R., Adams, N., and Heeringa, B, 2007. Voting Experts: An Unsupervised Algorithm for Segmenting Sequences. *In Journal of Intelligent Data Analysis*.

Comani, S., 1987. The historical temperature series of Bologna, Italy from 1716 to 1774. *Climatic Change 11(3)*: 375-390.

Fanta, B., Zaake, B.T., Kachroo, R.K., 2001. A study of variability of annual river flow of the southern African region. *Hydrological Sciences 46* (4): 513 – 524.

Giakoumakis, S. G., Baloutsos, G., 1997. Investigation of trend in hydrological time series of the Evinos River basin. *Hydrological Sciences Journal 42(1)*: 81-88.

Hewlett, D., Cohen P., 2009. Bootstrap Voting Experts. *Proceedings of the Twenty-first International Joint Conference on Artificial Intelligence (IJCAI).*

Kocyan, T., Martinovič, J., Kuchař, Š., and Dvorský J., 2012. Unsupervised Algorithm for Post-Processing of Roughly Segmented Categorical Time Series. *Proceedings of Databases, Texts, Specifications, and Objects*.

Maheras, P., Kolyva-Machera, F., 1990. Temporal and spatial characteristics of annual precipitation in the twentieth century. *Int. J. Climatol*. 10: 495-504.

Miller, M., Stoytchev, A, 2008. Hierarchical Voting Experts: An Unsupervised Algorithm for Hierarchical Sequence Segmentation. *Proceedings of the 7th IEEE International Conference on Development and Learning (ICDL)*.

Miller M., Wong P., and Stoytchev A., 2009. Unsupervised Segmentation of Audio Speech Using the Voting Experts Algorithm. *Proceedings of the Second Conference on Artificial General Intelligence (AGI)*.

Muller, M., 2007. Dynamic Time Warping. *Information Retrieval for Music and Motion*, Springer, ISBN 978-3-540-74047-6, 69--84.

Theodoridis, S., and Koutroumbas, K., 2006. Pattern Recognition. *Third Edition. Academic Press.*

Patel, P., Keogh, E., Lin, J., and Lonardi, S., 2002. Mining Motifs in Massive Time Series Databases. *In proceedings of the 2002 IEEE International Conference on Data Mining*. Maebashi City, Japan. Dec 9-12.

Watson, I., 1997. *Applying Case-Based Reasoning: Techniques for Enterprise Systems*, 15-38.

SIMULATION OF THE FLOOD WARNING PROCESS WITH COMPETENCY-BASED DESCRIPTION OF HUMAN RESOURCES

Štěpán Kuchař[(a)], Michal Podhorányi[(b)], Jan Martinovič[(c)], Ivo Vondrák[(d)]

VŠB – Technical University of Ostrava, IT4Innovations
17. listopadu 15/2172, Ostrava-Poruba, Czech Republic

[(a)]stepan.kuchar@vsb.cz, [(b)]michal.podhoranyi@vsb.cz, [(c)]jan.martinovic@vsb.cz, [(d)]ivo.vondrak@vsb.cz

ABSTRACT

This paper describes the flood warning process and its two main parts that are concerned with correctly predicting the potentially coming flood and issuing correct and quick response based on this warning. In order to correctly predict the magnitude of the flood, the hydrometeorology specialists in the process have to be adequately skilled, but process simulation models are not very concerned with accurate human resources modelling. A method for the description of human resources' skills in process simulations is proposed and used to extend the discrete event process simulation method called the BPM Method. This method is then used for modelling and simulation of the flood warning process to identify sections that cause the greatest delay and to propose improvements to the number and skill sets of specialists in the process.

Keywords: discrete event process simulation, flood warning process, human resource competency model, BPM Method

1. INTRODUCTION

Water is one of the most useful and one of the most destructive things on Earth. Most of the time, it is completely benign, but in large enough quantities it can overturn cars, demolish houses and even kill.

Flood waters are waters which escape from a watercourse in great volume and flow over adjoining lands in no regular channel. Floods kill millions of people, more than any other natural disaster. Flooding is also the world's most expensive type of natural disaster. The cost of global flood damage is hundreds of billions of euros.

Floods frequently affect the population of Central and Eastern Europe and therefore are studied by numerous scientific research institutes. Almost all large rivers in Central and Eastern Europe have experienced catastrophic flood events, e.g. the 1993 and 1995 flooding of the river Rhine, 1999 and 2002 Danube/Theiss rivers, 1997 Odra river, 2001 Visla river and 2002 Labe river. Floods, however, affect not only Central and Eastern Europe, but they represent a major problem in many regions all around the world (Horritt and Bates 2002; Knebl, Yang, Hutchinson and Maidment 2005).

In August 2002 the Czech Republic was hit by devastating floods, in what was the biggest natural disaster in modern Czech history. In some areas the floods - which affected over one third of the country - were the worst in 500 years. There were 17 deaths and thousands of people had to be evacuated from their homes. The floods caused damages of over 73 million crowns (2.5 million US dollars). One year later, life returned to normal in many of the affected areas, though the damage is still being dealt with in some places.

The growing number of losses caused by floods in countries around the world suggests that general mitigation of disasters is not a simple matter, but rather a complex issue in which science and technology can play an important role (Flowers 2003, Cheng 2006, Guo 2010). An important area of scientific study is modelling and simulations.

The main objective of this study is to model the flood warning process and simulate it to find a way to shorten the duration of the flood prediction preparation. The early flood forecast is very crucial because it allows the flood warning committee to quickly predict flood emergency and provide security countermeasures to potentially endangered areas. To adequately simulate this process it is necessary to possess information about all its components including all involved specialists and their specific competencies. Without this information it is difficult to reveal all bottlenecks of the process that delay the process execution because they can be caused not only by the wrong workflow structure but also by wrong allocation of human resources.

2. FLOOD WARNING PROCESS

The flood warning process has two distinct stages: flood warning and response (Sene 2008). The flood warning stage starts with a detection of the potential flooding threat. This activity should be done periodically based on river and precipitation monitoring and meteorology forecast. When a threat is detected, the flood warning committee is informed and it issues a request for more detailed forecasts to the institute of hydrometeorology and local catchment area offices. These organizations provide:

- information about actual river and reservoir situation;
- hydrodynamic modelling – flood simulations, flood maps, simulations of water elevation and water velocity, a real-time hydrological model for flood prediction using GIS, sediment transport, water quality analysis, etc.;
- rainfall runoff modelling – simulation of surface runoff;
- erosion modelling – simulation of water erosion;
- collection and archiving of flood data that can be used for estimating the magnitude of the flood based on historical evidence.

If these predictions identify possible emergency situations, the Emergency Manager (head of the flood committee) alerts relevant agencies and the process moves to the response stage. In this stage the countermeasures are implemented based on the forecast simulations from the flood warning stage. The following organizations and their responsibilities are part of the flood warning response stage:

1. Police:
   - assists with evacuation;
   - organizes and disseminates casualty information;
   - coordinates emergency services, local authorities, media etc.;
   - secures, protects and preserves the scene, and controls traffic;
   - helps with restoration of affected areas.

2. Fire & rescue services:
   - assesses hazards concerning evacuation;
   - rescues trapped people;
   - minimizes environmental dangers;
   - controls fires, released chemicals and other hazards;
   - cooperates with ambulance and medical services.

3. Ambulance and medical services:
   - saves life in conjunction with other emergency services;
   - assists and stabilizes injured people;
   - provides ambulances, medical staff, equipment and resources;
   - arranges transport for injured people;
   - alerts receiving hospitals.

4. Emergency coordinator:
   - prepares Emergency Plans for local resources and useful equipment;
   - issues warning messages to local authorities;
   - advises on weather, water flow, warnings and evacuation;
   - issues media statements;
   - issues situation updates.

In addition, the flood predictions are still provided even in the response stage to support the decision processes related to performed countermeasures and actions. During the response stage additional areas can be affected by the flood emergency and these predictions should be able to identify these areas in advance. This demands a number of skilled hydrology specialists that significantly influence the delay and precision of such forecasts and thus the efficiency of the whole process. We have therefore focused on this aspect of the process in our study.

3. THE BPM METHOD

A modelling and simulation method that is able to sufficiently model human-based processes is needed to model and simulate the flood warning process. For these purposes we used the discrete event modelling and simulation method called the BPM Method (Vondrák Szturc and Kružel 1999) that already provides simulation environment with stochastic parameters (Kuchař and Kožusznik 2010) and also specifies how the generic resources should be shared in concurrent instances and activities of the process (Kuchař, Ježek, Kožusznik and Štolfa 2012). This method defines three basic models of the process:

1. architecture of the process;
2. objects and resources utilized in the process;
3. behaviour of the process.

The most important one of these models for performing simulations is the behavioural model. This model is called the Coordination model and it specifies the behaviour of the process as a sequence of activities. It also specifies what resources the activities demand and which artefacts they consume and produce. Alternative flow in the coordination model is enabled by multiple activity scenarios and concurrency of the activities can also be modelled. This model can also be converted to a Petri net to provide exact semantics for performing simulations (Kuchař and Kožusznik 2010; Kuchař, Ježek, Kožusznik and Štolfa 2012).

4. HUMAN RESOURCE COMPETENCIES

Existing process simulation models are not very concerned with accurate human resources modelling and description (Rozinat, Wynn, Aalst, Hofstede and Fidge 2009; Aalst, Nakatumba, Rozinat and Russel 2008). But clearly each human resource in the process is unique with his own set of skills and experience, each one has specific working habits and performance (André, Baldoquín and Acuña 2010). In our paper we use a competency-based approach to differentiate individual resources in the process and to correctly

allocate resources to the process activities during flood warning process simulations.

4.1. Resource Competency Description

The description of the human resources' skills in the process is commonly done by using the competency models (Dreyfus and Dreyfus 1980; Sinnott, Madison and Pataki 2002; Ennis 2008) and skills frameworks (e.g. NHS Knowledge and Skills Framework (UK Department of Health 2004). Competency models describe various competencies which are important for the process. Competencies are defined as sets of knowledge, abilities, skills and behaviour that contribute to successful job performance and the achievement of organizational results (Sinnott, Madison and Pataki 2002). Skills frameworks have the same purpose, but they describe skills particular for one domain rather than general competencies. But in fact skills are just a special type of competencies.

Competency models and skills frameworks also describe how to measure and evaluate individual competencies. In most cases competencies are measured by a number of advancing stages where higher levels of competency include everything from their lower levels. There is no standard for how many levels a competency model should have and every model defines its own set of levels.

Let us look at a small example of one hydrology specialist's competencies in the flood warning process. Some of his competencies in a 10-level model could be described as:

- catchment area of the Odra river – 8. level;
- catchment area of the Opava river – 3. level;
- rainfall-runoff modelling – 4. level;
- hydrology analysis – 6. level;
- communication – 4. level, etc.

Domain specific skills (rainfall-runoff modelling, hydrology analysis), general competencies (communication) and knowledge of the environment (catchment area) are contained in this example. It is clear that competencies in the model have to be based on the process requirements and professional domain.

4.2. Activity Requirements Description

All activities in the process also have competency-based requirements that describe what competencies the human resources performing this activity should know. Each activity will therefore be defined by the set of competency levels for each resource type performing the activity specifying that only resources with given level or higher will do the activity as planned. Resources with lower competencies are able to finish the activity, but they have to spend additional time to learn how to perform the activity and their work is prone to contain more errors. A simple example of requirements for the activity of analysing results of hydrological models could have following requirements

- catchment area – 6. level;
- hydrology analysis – 7. level;
- cartography – 4. to 7. level;
- statistics – 5. to 7. level.

By comparing these requirements with the resource competencies mentioned above, the high level limits of the requirements stand out. These limits are introduced to prevent the allocation of highly skilled specialists to simple tasks that can be done by average workers. Another difference can be found in the generalization of the catchment area. When assessing the specialist's competencies, it is better to define the competency levels in specific parts of the domain so that the workers are assessed as precisely as possible. On the other hand the activity requirements should only define a level for the whole parameterized competency and relevant part of the domain will be specified by the parameters of actual process case. In other words, if the hydrology team tackles with a case concerning the catchment area of the Odra river, then the requirement in this case will be refined as the catchment area of the Odra river.

5. INTEGRATION OF COMPETENCY-BASED DESCRIPTION TO THE BPM METHOD

The BPM Method is designed as an object oriented method so introducing the competency-based description of resources and activity requirements can be done by expanding the object descriptions of activities and resources in this method.

5.1. Resource Competency Extension

Each resource with definable competencies has to be part of any shared resource pool in the process. Each such pool contains one type (role) of resources and is shared among all process instances (Kuchař, Ježek, Kožusznik and Štolfa 2012).

Each resource object in the BPM Method is then extended by the collection of competencies and their levels that are specific for this resource. Levels of parameterized competencies (as described in the previous section on the catchment area example) have to be specified individually for each possible parameter of the competency because the resource can have different levels for different parameters.

5.2. Competency Parameters Extension

Each process instance defines several sets of competency parameters that will specify the process case and influence the allocation of resources with parameterized competencies. Each parameter set comes with the percentage probability that exactly this parameter set will be chosen for this process instance. Total sum of probabilities for all parameter sets' in one process instance has to be 100% to ensure that each process case has one of these parameter sets in each simulation. For example one parameter set of the process instance could have 20% probability that it will be a process case for the Opava river catchment area and 80% that it will be for the Odra river area.

5.3. Activity Requirements Extension

Each activity in the process can be expanded by the description of its requirements for each resource type that performs this activity. Required competencies have to come from the same set as the competencies of resources to ensure their comparability. Each activity object now contains several collections of requirements – one collection for each input resource. Each requirement comes with the low and high competency limits (as described in section 4.2) and importance of this required competency. Less important competencies have weaker impact on the suitability of resources lacking these competencies.

Parameterized competencies can be specified by one competency requirement that will be refined for each process instance and its chosen parameters, or by several requirements for specific parameters that are required in every process case.

5.4. Competency-Based Resource Evaluation

Whenever any activity with requirements in any process case needs to start its execution, the simulation needs to allocate one or more resources to perform this activity. These chosen resources have to be suitable for performing this activity by having enough skills and experience to fulfil its requirements. This suitability can be evaluated by encoding the resource competencies and activity requirements to their vector representation and evaluated in the vector space model (for more information see (Kuchař and Martinovič 2012)). The resulting suitability is then compared to the referential resource of the activity that has exactly same competency levels as those required by the activity. Workers with higher suitability than the referential resource are considered suitable to perform the activity.

5.5. Resource Utilization and Unavailability

The last extension added to the BPM Method is a method for determining the utilization of each shared resource in the process. This utilization is measured by simply counting up the time when the resource is performing any activity.

Utilization is an interesting result of the simulation but it is not very useful in optimizing the performance of the process. When optimizing the number of resources in the process we are not interested in answering how long one resource was doing something in the process, but rather how long did the process have to wait for the resource when it was needed to perform another activity. One resource cannot perform two activities at the same time but activities and processes run concurrently and they very often need the same resource to be able to continue their run (e.g. one hydrology specialist is needed to model the rainfall-runoff in one process case and at the same time calibrate another model in another case). When this happens the resource has to perform these tasks sequentially by:

- finishing the first task and then starting the second one, or
- pausing the first task and returning to it after finishing the second one, or
- switching back and forth between these tasks.

In either way one task will have to wait for the completion of the other (or partial completion in the case of the third option). It is therefore important to be able to simulate and measure these waiting times. The BPM Method is only able to model the first sequencing option. Whenever an activity is enabled but the resource is not available, the BPM Method counts and notes the time needed for the resource to become available to perform the activity. Total waiting time for one resource is then computed by adding up these noted times for this appropriate resource.

6. CASE STUDY

We implemented all extensions of the BPM Method proposed in this paper into the modelling and simulation tool called BP Studio (Vondrák 2000) and used it to model, verify and simulate the flood warning process of the Moravian-Silesian region in the Czech Republic.

This process is specified in accordance with the process description presented in section 2. Simulations in this case study focus on the flood warning and forecast stage of the process that is performed by 4 worker roles – Hydrology Manager, Hydrology Analyst, Hydrology Specialist and Database Specialist. These workers are defined by a total of 20 competencies with one being further refined by 5 parameters. This parameterized competency describes knowledge of a specific catchment area and 5 major catchment areas were used in these simulations. Forecasts for each of these catchment areas are treated as individual process instances running concurrently with all other catchment area forecasts. Execution of each process instance starts 15-30 minutes after the start of the last instance according to the delay of specified requests and data coming from the emergency committee.

The first process simulation is executed for the regional flood forecast with the whole team of 1 Manager, 2 Analysts, 5 Hydrology Specialists and 2 Database Specialists each with his own acquired competency levels. 200 simulations are performed to mitigate the unpredictability of stochastic parameters on the results. The average duration of this process is 9 hours and 44 minutes with utilization and waiting times shown in Figure 1.

This simulation shows that the Manager is appropriately utilized because his waiting time is very short. The Manager is utilized in a small part of the process because the process only describes his role in acquiring data from other organizations and flood classification. He of course works through the whole process execution to lead other roles and help with problem resolution.

Waiting times of the Database Specialists are also very short but this fact opens a question if the second Database Specialist is needed in the process. Database

Specialists only play a support role in the process because their main goal is to retrieve and archive data for modelling and analysis activities in the process. Their work is therefore very scattered throughout the process and one worker in this role could manage all the requests only with a short delay. After performing another simulation without the second Database Specialist, this hypothesis proved to be true. The waiting time of the Database Specialist increased by 7 minutes but it had no influence on process duration.

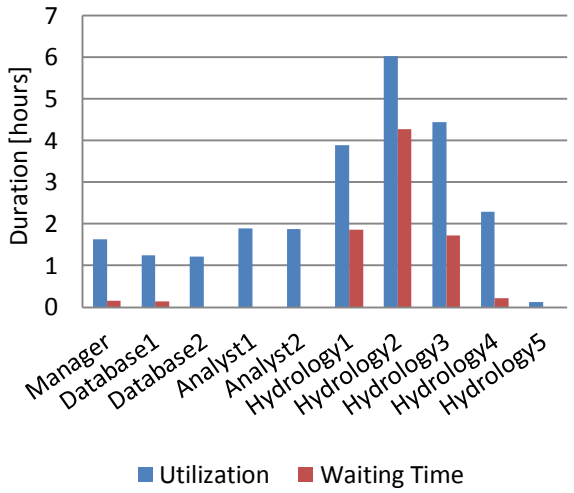


Figure 1: First Simulation Resource Utilization

The Analyst role follows a very similar pattern but the second resource in this role cannot be removed because both workers complement each other. Each Analyst is suitable for performing those activities that the other is not able to do.

The most utilized and most unavailable resources are from the Hydrology Specialist role and differences in their competencies are also very visible in this role. While the first, third and especially the second resource in this role show high waiting times, the fifth resource's competencies are so low that he cannot help with performing the standard process activities. The first step is to find the source of the high waiting time values for the second Hydrology Specialist. The simulation shows that his waiting times are caused by one activity (Calibration of the hydrodynamic model) and only this resource is able to perform this activity. But the fourth Hydrology Specialist is only a little worse than the referential resource for this activity (36% suitability of the fourth specialist as compared to the 39% suitability of the referential resource). This means that it is possible to easily gain another suitable resource by training the fourth specialist in the Hydrodynamic modelling or the Hydrodynamic calibration competencies in which he lacks the level required for performing the activity. It is enough to train him in one of these competencies to pass the suitability condition because his mastery in other required skills balances the lack of mastery in the other one.

After updating one of these competencies and executing the simulation, the results were significantly better. Total duration of the process was 8 hours and 31 minutes (more than 1 hour shorter than before) and waiting times of the second and fourth Hydrology Specialists changed to approximately 30 minutes.

To decrease the waiting times for the first and third Hydrology Specialist, a similar training is needed for the fifth specialist. There is only one possible competency to train because other competencies barely fulfil the requirements. After updating this competency, the simulations showed that several activities opened for the fifth specialist along with the one that delayed the other two specialists. This decreased the total process duration to 8 hours and 11 minutes and Figure 2 shows all resource utilizations after this change.

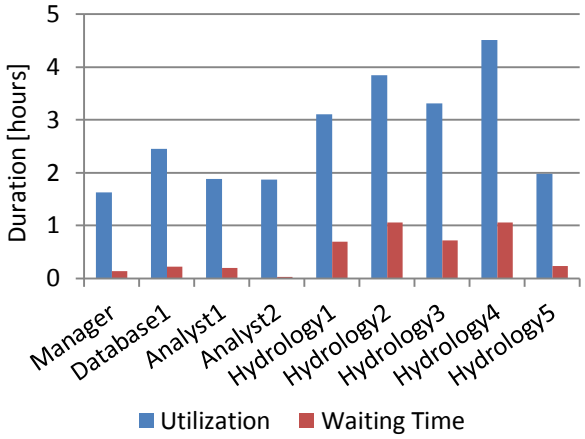


Figure 2: Updated Simulation Resource Utilization

The waiting times of Hydrology Specialists are still high and this situation cannot be solved by improving additional competences because all specialists are already unavailable throughout the whole process. Another option is to hire new hydrology specialist with the same competencies as the fourth Hydrology Specialist for him being the most unavailable worker. Total process duration of the final simulation is 7 hours and 34 minutes that is by 37 minutes lower from the previous result. By adding yet another specialist to the second most unavailable group the total process duration decreased by another 15 minutes.

7. CONCLUSION AND FUTURE WORK

This paper described a flood warning process and its importance for the safety of people's lives. To create functional countermeasures against the coming flood, it is important to predict the magnitude and location of the flood as quickly and as precisely as possible. This can be managed by employing enough experienced specialists to perform the activities of the flood prediction subprocess. To analyse the situation in the Moravian-Silesian region, we proposed to simulate the process with the current hydrology team and suggest several improvements to the team composition.

A competency-based extension for description of worker's skills was created for the BPM Method for this purpose and its results were evaluated in the case study.

The competency-based extension of the BPM Method evaluates the resources in the process in accordance to their acquired competencies and compares their suitability with this referential resource of this activity. In this version of the extension, this evaluation only specifies if the worker is competent enough to perform the activity but in reality his competencies influence the time and effort he has to spend on executing this activity. The more competent the worker is the better he performs in doing the activity (Hatch and Dyer 2004). This performance evaluation could also solve the binary nature of the suitability evaluation. By knowing their performance for given activity, even unsuitable resources could be allocated for this activity but at the cost of longer execution time and greater risk of error. This could serve well for process instances with lower priorities.

ACKNOWLEDGMENTS

This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070) and by the grant project of the Technology Agency of the Czech Republic under number TA01021374, titled "Nove technologie ochrany zivotniho prostredi pred negativnimi nasledky pohybujicich se prírodnich hmot" (New Technologies for Environmental Protection against Negative Effects of Natural Mass Transfer).

REFERENCES

Aalst W.M.P. van der, Nakatumba J., Rozinat A., Russell, N.C., 2008. Business process simulation : how to get it right? *Computer Science Report, No. 08-21*. Eindhoven: Technische Universiteit Eindhoven, 25 pp.

André M., Baldoquín M.G., Acuña S.T., 2010. Formal model for assigning human resources to teams in software projects. *Information and Software Technology* 53 (3):259-275.

Cheng, X.T., 2006. Flood and drought evaluation and management. *General report for question 87*. Barcelona: ICOLD.

Dreyfus S., Dreyfus H., 1980. *A Five-Stage Model of the Mental Activities Involved in Directed Skill Acquisition*. University of California, Berkley.

Ennis M. R., 2008. *Competency Models: A Review of the Literature and The Role of the Employment and Training Administration (ETA)*. US Department of Labor.

Flowers, R.B., 2003. Flood management: a key to sustainable development and integrated water resources. *3rd World Water Forum*, 16-23 March, Kyoto, Stiga, Osaka, Japan.

Guo, H.D., 2010. Understanding global natural disasters and the role of earth observation. *International Journal of Digital Earth*, 3 (3), 221-230.

Hatch, N. W., Dyer J. H., 2004. Human capital and learning as a source of sustainable competitive advantage. *Strategic Management Journal* 25 (12): 1155-1178.

Horritt, M.S., Bates, P.D., 2002. Evaluation of 1D and 2D numerical models for predicting river flood inundation. *Journal of Hydrology* 268(1–4),87–99.

Knebl, M.R., Yang, Z.L., Hutchinson, K., Maidment, D.R., 2005. Regional scale flood modeling using NEXRAD rainfall, GIS, and HEC-HMS/RAS: a case study for the San Antonio River Basin Summer 2002 storm event. *Journal of Environmental Management* 75, 325-336.

Kuchař Š., Kožusznik J., 2010. BPM Method Extension for Automatic Process Simulation. *8th Industrial Simulation Conference 2010*, pp. 80-85. 7-9 June, Budapest, Hungary.

Kuchař Š., Ježek D., Kožusznik J., Štolfa S., 2012. Sharing Limited Resources in Software Process Simulations. *10th Industrial Simulation Conference 2012*, pp. 33-37. 4-6 June, Brno, Czech Republic.

Kuchař Š., Martinovič J., 2012. Human Resource Allocation in Process Simulations Based on Competency Vectors. Accepted to the *7th International Conference on Soft Computing Models in Industrial and Environmental Applications 2012*. 5-7 September, Ostrava, Czech Republic.

Rozinat A., Wynn M.T., Aalst W.M.P. van der, Hofstede A.H.M. ter, Fidge C. J., 2009. Workflow simulation for operational decision support. *Data & Knowledge Engineering* 68 (9): 834–850.

Russell N., Aalst W.M.P. van der, Hofstede A.H.M. ter, Edmond D., 2005. Workflow Resource Patterns: Identification, Representation and Tool Support. *Advanced Information Systems Engineering*, 3520:216-232. Springer Berlin Heidelberg.

Sene, K., 2008. *Flood Warning, Forecasting and Emergency Response*. Springer.

Sinnott G. C., Madison G. H., Pataki G. E., 2002. *Competencies: Report of the competencies workgroup, workforce and succession planning work groups*. New York State Governor's Office of Employee Relations and the Department of Civil Service.

UK Department of Health, 2004. *The NHS Knowledge and Skills Framework (NHS KSF) and the Development Review Process*. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4090843

Vondrák I. 2000. Business Process Studio, version 3.0. Electronic manual, VŠB – Technical University of Ostrava.

Vondrák I., Szturc R., Kružel M., 1999. BPM – OO Method for Business Process Modeling. *ISM '99 Proceedings*, pp.155-163. Rožnov pod Radhoštěm, Czech Republic.

SUPERVISED TRAINING OF CONVERSIVE HIDDEN NON-MARKOVIAN MODELS: INCREASING USABILITY FOR GESTURE RECOGNITION

Sascha Bosse[(a)], Claudia Krull[(b)], Graham Horton[(c)]

[(a)(b)(c)] Otto-von-Guericke-University Magdeburg
P.O. Box 4120
39016 Magdeburg, Germany

[(a)]sascha.bosse@ovgu.de, [(b)]claudia.krull@ovgu.de, [(c)]graham.horton@ovgu.de

ABSTRACT

Hidden non-Markovian Models (HnMMs) were introduced and formalized as an extension of Hidden Markov Models for the analysis of partially observable stochastic processes. Their main advantage over HMM is the possibility to model arbitrary distributions for state transition duration, so that the unobservable stochastic process needs not to be Markovian. Besides academic examples, HnMMs were applied to gesture recognition and performed well in distinguishing similar gestures with different execution speeds. While the Proxel-Method enabled the evaluation for arbitrary HnMMs, there was no opportunity to train these models. Therefore, the models for different gestures had to be parameterized manually. This fact reduced the applicability in real gesture recognition dramatically. This paper presents a solution to this problem, introducing a supervised training approach that increases the applicability of HnMMs in gesture recognition.

1. INTRODUCTION

Hidden non-Markovian Models (HnMMs) were introduced and formalized by Krull and Horton (2009) as an extension of Hidden Markov Models for the analysis of partially observable stochastic processes. Their main advantage over HMM is the possibility to model arbitrary distributions for state transition durations, so that the unobservable stochastic process needs not to be Markovian.

Besides academic examples (e.g. Buchholz et. al 2010; Krull et. al 2010), HnMMs were applied to gesture recognition and performed well in distinguishing similar gestures with different execution speeds (Bosse et al. 2011). For that purpose, significant changes in gesture acceleration were logged while execution and the likelihood of different HnMMs to generate such a sequence was computed. The model with the highest value in this evaluation task represents the recognized gesture.

While the Proxel-Method developed by Horton (2002) enabled the evaluation for arbitrary HnMMs, there was no opportunity to train these models automatically. Therefore, the models for different gestures had to be parameterized manually. This fact reduced the applicability in real gesture recognition dramatically.

This paper has the goal to present a solution to this problem, introducing a supervised training approach that increases the applicability of HnMMs in particular in gesture recognition, but also in other application areas where a fully specified model of the hidden system is not readily available.

The next section will review some existing training methods of related paradigms and introduce HnMMs. The third section describes the steps of the training approach and the fourth section comments on implementation details. The experiments section contains two test cases. The paper is concluded by the sixth section, which evaluates the approach presented and highlights some areas future work.

2. RELATED WORK

Training a mathematical model to increase its applicability is a well addressed problem in *Machine Learning*. There are two basic forms of Machine Learning: Supervised and unsupervised learning. In supervised learning input and desired output data is used, so that the model can learn the relationship between them (e.g. Classification). In unsupervised learning, no output data is given and the model has to describe the distribution of the data (e.g. Clustering) (Marsland 2009).

For Hidden Markov Models there exist well known unsupervised training algorithms, like Baum-Welch- and Viterbi-Training (Fink 2008). These methods are iterative and guarantee a greater or equal likelihood of the model after each iteration. Buchholz (2012) adapted the Baum-Welch-Algorithm to a subclass of HnMMs also used in gesture recognition, but there is still missing a concept to train arbitrarily distributed state transition durations.

On the other hand, approaches for supervised training of Hidden Markov Models like (Mamitsuka 1998) were developed to train the evaluation probability to a specific target value. An algorithm that trains a Hidden Markov Model from sequences with desired states could not be found in literature, probably because the computation of transition and output probabilities would be trivial in that case.

For the problem of gesture recognition addressed in (Bosse et. al 2011) none of these approaches is suitable in general. Because semantic information is encoded in the model, a training method must respect this and may not change the basic structure of the model. Because of this fact, the current state of the underlying model should be computed from training data, so that the training method can respect the semantics of the model.

2.1. Hidden non-Markovian Models

An HnMM after Krull and Horton (2009) consists of a state space with transitions between the states that are time-dependent. In a specific case of HnMMs (*Eall*), every transition must emit a symbol when it fires. Buchholz (2012) named this subclass Conversive HnMM and developed algorithms for all relevant problems. CHnMM assign every transition with a discrete random variable that indicates what emission probability each output symbol has. In addition, a distribution of the initial probability of each state is given.

The evaluation task of CHnMM is solved by an approach very similar to the Forward Algorithm from HMM (Buchholz 2012). This algorithm requires a completely defined system description, including the continuous distributions of transition firing times and discrete distributions for symbol emissions. These distributions must be provided by a training algorithm.

The state space of an HnMM can be computed from arbitrary models that represent discrete stochastic processes. Furthermore, the state space of a model does not need to be computed a-priori, so the training algorithm can parameterize a discrete stochastic model to avoid problems like state space explosion. In this work, augmented stochastic Petri nets are used as the user model.

2.2. Augmented Stochastic Petri Nets

Krull (2009) defines Stochastic Petri Nets as a 7-tuple . is the set of places, a set of transitions and a set of arcs between places and transitions. must form a bipartite graph. is a set of so called inhibitor arcs while is a function that assigns integer values to each type of arc. is the initial marking of all places and can assign a so called guard function to each transition that indicates whether a transition is activated or deactivated for a specific marking.

A transition is activated if and only if every place connected to this transition has at least the number of tokens the connecting input arcs are assigned, no source of an inhibitor arc to this transition has at least the specified number of tokens and the corresponding guard function of this transition evaluates to true in the current marking. After a randomly distributed amount of time, the transition fires, deleting the required tokens and creating tokens in the places the transition is connected to. If another activated transition fires before that time and the new marking disables the transition, there are two policies: If the transition is marked as “Race Age”, the transition saves the remaining firing time and resumes the countdown when it is active again. With the policy “Race Enabled” the whole activation time is deleted and upon re-enabling, the firing distribution is sampled again.

Buchholz (2012) defines augmented stochastic Petri nets. Here each transition is augmented by the symbol emissions it can produce with probabilities for each possible symbol. The state space of such an ASPN is an HnMM. ASPN can be considered the user model corresponding to the computational model HnMM.

2.3. Example System Description

The current paper is illustrated using an example first defined by Buchholz et. al (2010). Two machines process products in randomly distributed intervals. These products are tested after both machine results are joined. The tester produces a protocol with test timestamp and state of the product (*OK* or *Defect*). Figure 1 illustrates the system and shows an example of such a system protocol.

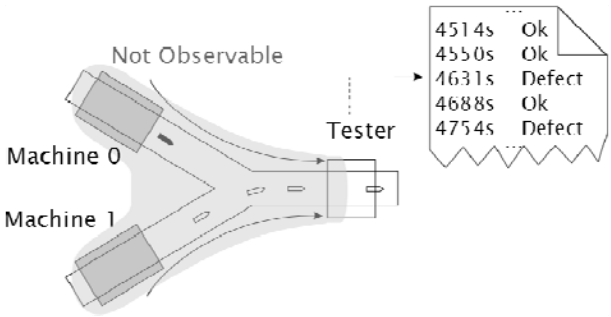


Figure 1: Schematic of tester example

This system can be converted into the augmented stochastic Petri net (Buchholz 2012) shown in Figure 2. The ASPN represents the system as a stochastic Petri net with output symbols representing the tester results attached to the state transitions.

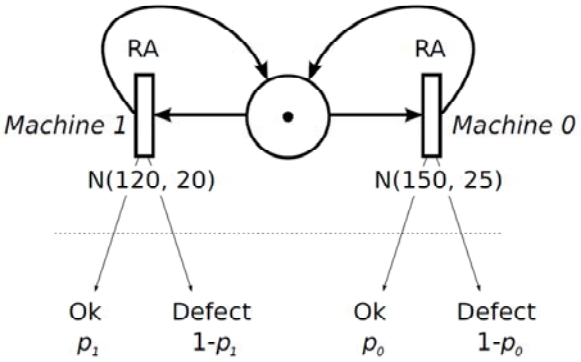


Figure 2: Augmented stochastic Petri net of the tester example

time stamp	tester result
95.755	OK
119.486	OK
225.029	Defect
236.694	OK
352.712	OK
391.634	OK
505.593	OK
559.207	OK
618.783	OK
677.500	OK
744.931	OK
803.042	OK
908.050	OK
994.619	OK

Figure 3: Example protocol of the tester system

The protocol produced by the tester (see Figure 3 for an example) does not contain information on the machine that processed a particular product, e.g. produced a certain defective item. Therefore, this part of the model can be considered unobservable. The task to be solved using HnMM is therefore to reassign the different protocol entries to the machines, thereby determining the sources of defective items.

3. APPROACH

The goal of the desired approach is a supervised training of a Hidden non-Markovian Model to improve the recognition accuracy for a given real system. For that purpose, the hidden discrete model must be adapted by using training data that contains more information than the data from the real system to be reconstructed later on. A second property of the approach should be that the trained model gets nearer to the original with increasing training data amount.

3.1. Preconditions

A symbol sequence entry of an HnMM consists of two parts: timestamp and symbol. For training purposes, those sequence entries are annotated with the transition that generated the symbol emission as shown in Figure 4. The user model to be trained is an ASPN with a known structure. The model parameters that will be trained are the continuous distribution functions of the timed transition and the corresponding symbol output probabilities. An initial state probability distribution will also be determined. This corresponds to parametric training.

time stamp	tester result	machine
95.755	OK	1
119.486	OK	0
225.029	Defect	0
236.694	OK	1
352.712	OK	0
391.634	OK	1
505.593	OK	0
559.207	OK	1
618.783	OK	0
677.500	OK	1
744.931	OK	0
803.042	OK	1
908.050	OK	0
994.619	OK	1

Figure 4: Annotated example trace of the tester system

Two main tasks can be identified to solve this problem: Firstly, for every transition samples of the firing time must be computed from the annotated symbol sequence and secondly, the distribution must be estimated from these firing time samples.

3.2. Computation of firing time samples

For every event in the protocol at time representing a firing transition, the corresponding firing time can be computed in the following way:

(1)

That means the relative firing t_{fire} time is the difference between the timestamp of the firing t and the time when the transition was last activated t_{act}. is the current age of the transition which is not equal to zero if the transition is race age and was activated before but did not fire. While the timestamp is available of course, the other two values are not from the protocol itself.

But they can be computed easily when the protocol is processed sequentially. Given the initial marking and the structure of the underlying model, at every timestamp the marking - depending on the given transition - is stored. From this marking the activated transitions can be inferred. For every transition this activation time and for all race age transitions that are not longer activated the age, i.e. the difference between timestamp and last activation time, is stored.

With this procedure, every entry in the sequence produces a relative firing time for the given transition, so that a collection of firing times arises.

3.3. Estimating the Probability Distribution

Estimating a probability distribution from a set of realizations is a problem well addressed in density estimation. The methods for this task are divided in parametric and non-parametric methods (Eggermont and LaRiccia 2001). The first mentioned is about estimating the parameters of known distribution types, the normal distribution for example. If none of these distributions fits the data, a non-parametric method can be applied. This can be a simple histogram or the more complex kernel estimation.

4. IMPLEMENTATION

To test the supervised training, the approach needed to be implemented. The computation of firing times can be done with a Petri net simulation providing the following methods: For a specific marking, a list of active transitions and for a given transition, a new marking must be returned. In addition to that, some auxiliary variables are needed. Besides the current index corresponding to the processed symbol sequence entry, the current marking of the Petri net must be saved (integer array, size). Also for every transition, the last activation times and the age times must be stored (float array, size). The latter array can be minimized, if only race age transitions are considered.

In addition to that, for every transition the symbol emissions must be counted.

From the collected data, the distributions must be computed. The estimation of the discrete probability distribution of the symbol emissions are easy to compute. Estimator is the relative frequency of the symbol emission.

For the estimation of the continuous probability distributions of the firing times the kernel estimation seems to be the most general approach. Although it can be shown that the error of this method tends to zero with enough data (Devroye and Lugosi 2001), a parametric estimation reduces the error much faster with respect to training data if the true distribution is a known one. This holds because less information is need to parameterize a known distribution than an unknown one.

Due to these facts, both approaches are considered in this concrete implementation. Firstly, an optimization for some known probability densities is carried out. Minimizing the square error between observed values and expected values in defined intervals returns the distribution fitting best.

If this distribution does not pass a chi-square-test at a specific significance level (from Banks et. al 2001), kernel estimation is performed with Gauss kernel and window size chosen according to Silverman (1998).

The needed cumulative distribution function is computed symbolically for the corresponding densities and through numerical integration for densities without symbolic integrations and kernel density.

With the estimated distributions, the augmented stochastic Petri net can be parameterized and the CHnMM evaluation algorithm is now able to compute the evaluation probabilities of other protocol sequences.

5. EXPERIMENTS

To test the developed approach, two experiments are performed. Since it is not clear yet, how to generate training data from real systems, training is tested using two academic models, where a simulation model is available to generate both training data and test sequences.

Therefore we have as reference value to estimate the quality of the trained models the evaluation probability of each sequence computed using the actual generating model. The training approach is successful if the evaluation probability of the sequence for the trained model is similar to this reference value. In particular, the difference between the values should decrease the more training data is available.

To illustrate the different effects of parametric and non-parametric estimation, both approaches are shown in the experiment results.

The computation time needed for extracting and finding the distributions from the data, as well as the runtime of the evaluation task were both under one minute and are therefore not considered in this paper.

5.1. Tester Example

The first experiment is carried out using the tester example system described in Section 2.3. For training purposes, augmented training protocols of different lengths were generated (20, 50, 100, 200, 500, 1,000, 2,500 and 10,000 entries for each transition). One model is trained using kernel estimation, and the parametric training is performed using normal distributions.

Then for ten different protocols, the evaluation probability of the protocol given the trained model is computed. Reference evaluation probabilities are computed using the CHnMM of the generating model. The mean relative difference (can also be interpreted as the error) of the trained model probabilities and the reference values using a particular amount of training data was computed using Equation (2).

$$\overline{Err_n} = \frac{1}{10}\sum_{i=1}^{10}\frac{|p_{i,n}-p_{i,ref}|}{p_{i,ref}} \qquad (2)$$

The reference evaluation probability of test sequence i is given by $p_{i,ref}$ while the evaluation probability of test sequence i obtained from a model trained with n training samples is given by $p_{i,n}$. The development of this error value with increasing training data amount is presented in Figure 5.

The first observation is that increasing the training data amount decreases the error and therefore increases the quality of the model, which is a necessary feature of a training approach. Because of the stochastic nature of the training data this decrease is not smooth for little training data. The effect is more smooth when more training date is available, which suggests that a minimum amount of training data needs to be available to obtain useful results.

The error decreases faster when training normal distributions, since the data was generated using a model with normal distributions. The models containing kernel estimations of the distributions retain a significantly higher error in the evaluation probability with the same amount of training data.

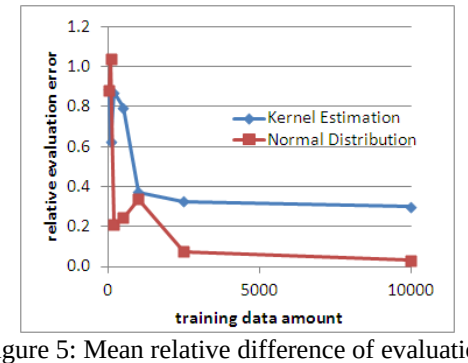


Figure 5: Mean relative difference of evaluation probabilities depending on training data amount

5.2. Non-Parametric Motion Sensor Example

In real applications, data is often assumed to meet a parametric distribution. But that is not always the case. Therefore, a second experiment was performed using a

model with more complex distributions. The example is a very simple motion sensor in a movable utility (e.g. game pad of a games console, pen or wiper on smart board). The task of the model would be to distinguish between deliberate movements the user and random influences on the utility such as jitter, jolting or draft.

The given system structure is very simple as shown in Figure 6 in the form of an ASPN. The two places of the ASPN represent the states *Idle* and *Busy*. *Idle* meaning that the utility is not in motion and *Busy* meaning that it is being used, and the corresponding motions should be registered. A speed threshold has been defined in order to distinguish between deliberate movements and random influences. The transitions between the states should cause the speed value to rise above (*Begin Moving*) or fall below (*Seize Moving*) the given threshold.

To emulate non-parametric data, transition *Begin Moving* has a firing time that is a combination of two normal distributions, N(60,10) and N(75,15), where a random process picks each of the distributions with equal probability and then samples that one. The firing time of transition *Seize Moving* is a convolution of a normal distribution and an exponential distribution, N(4,0.5) and Exp(0.5).

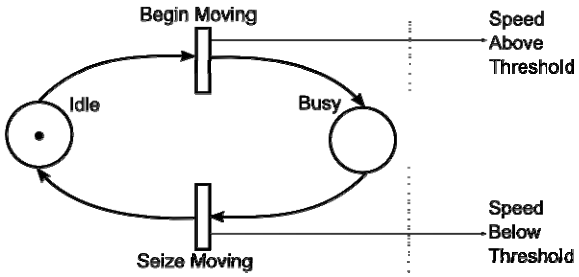


Figure 6: Augmented stochastic Petri net of non-parametric motion sensor model

For this model, training sequences of different length were generated using a discrete event-based simulation (20, 50, 100, 200, 500, 1,000, 2,500, 5,000 and 10,000 fire times for each transition). Additionally, ten test sequences were generated. As reference values for the evaluation probabilities the results of a model trained with almost 14000 items of data were used, because the non-parametric distributions could not be used directly in the analysis method. A generation of training sequences from a real motion sensor would require the users to indicate the start and end of deliberate movements e.g. by pressing a button on the game pad.

As estimate of the quality of the trained model, we again computed the mean relative error for different amounts of training data using Equation (2). The development of the error for the trained models containing kernel estimations as distributions is shown in Figure 7. It contains the mean error value as well as the maximum and minimum value obtained of the ten test sequences. The error development using parametric estimators can be seen in Figure 8 using a different scale than in the non-parametric case.

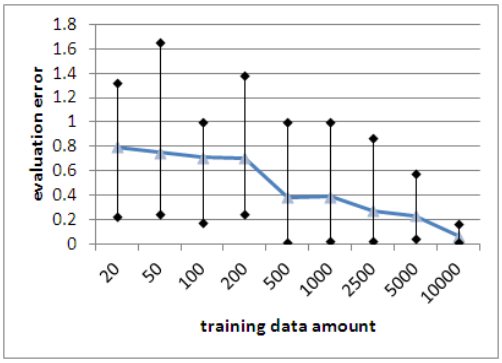


Figure 7: Error of evaluation probabilities depending on training data amount for non-parametric estimator

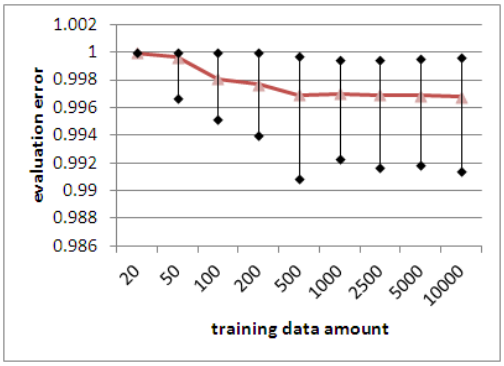


Figure 8: Error of evaluation probabilities depending on training data amount for parametric estimator

The error using the parametric estimator does not decrease significantly with increasing training data amount, while the non-parametric estimator performs significantly better. This is to be expected, since the original distributions were designed not to fit a known distribution function.

Maybe a more complicated parametric approach could explain the data better than the here estimated Erlang distribution for T1 and lognormal distribution for T2, but that was not within the scope of the experiment.

5.3. Experiment Discussion

The experiments show that a supervised training of Hidden non-Markovian Models is possible, and that the approach is capable of training parametric and non-parametric distributions. The quality of the trained models, in terms of difference in evaluation probability to a reference model, increases with increasing training data amount, when a certain minimal amount of data is available.

It becomes clear that parametric estimation yields better results even with little training data if the original data corresponds to a known distribution function. If this is not the case, a non-parametric estimator is a good choice because it always tends to the original data distribution given enough training examples.

6. CONCLUSION

This paper presented a supervised training approach for Hidden non-Markovian Models. Even though it is based on existing methods, only their combination leads to a first successful attempt to train Hidden non-Markovian Model. Experiments showed that this approach is able to minimize the gap between the source system and the trained model if the model structure is sufficient to map this system and enough training data is available.

The method to collect firing times of specific transitions is applicable to any augmented stochastic Petri net described as above. Parametric and non-parametric estimators are capable of approximating arbitrary data distributions. If the data fits to a parametric distribution the corresponding estimators reach that goal much faster than the non-parametric ones, although the latter approach is more general.

Disadvantages of the approach arise from the preconditions. An augmented Petri net fully specified in its structure is needed, otherwise the method to collect firing times cannot be applied. In addition to that this supervised approach also needs information on which transition has fired for each protocol entry. That means that the unobservable system part must be observable for the generation of training data. This might be a key problem for the applicability of the approach. However, in the case of gesture recognition within the approach of Bosse et. al (2011), it can be applied. For that purpose, users will need to mark their sequence of data generated from movements with the desired gesture phases.

In future work the approach should be applied to real gesture recognition with Hidden non-Markovian Models. This task involves scenarios with several users to collect training data from their specific gesticulation. Later on the users should carry out the defined gestures to see if the recognition rate increases through training.

A second opportunity of research is the applicability of the approach to so-called normal HnMMs in which not every transition does emit a symbol. That means the current state of the HnMM respectively the current marking of the Petri net when generating the firing times may be non-deterministic.

REFERENCES

Banks, J., Carson, J.S. II, Nelson, B.L., Nicol, D.M., 2001. Discrete-Event System Simulation. Upper Saddle River, NJ: Prentice-Hall

Bosse, S., Krull, C., Horton, G., 2011. MODELING OF GESTURES WITH DIFFERING EXECUTION SPEEDS: Are Hidden non-Markovian Models Applicable for Gesture Recognition, 10th International Conference on Modelling & Applied Simulation (MAS). September 2011, Rome, Italy.

Buchholz, R., Krull, C., Strigl, T., Horton, G., 2010. Using Hidden non-Markovian Models to Reconstruct System Behavior in Partially-Observable Systems, 3rd International Conference on Simulation Tools and Techniques, March 2010, Torremonlinos, Spain.

Buchholz, R., 2012. Conversive Hidden non-Markovian Models. Doctoral Thesis, Otto-von-Guericke-University Magdeburg, Germany.

Devroye, L, Lugosi, G., 2001. Combinatorial Methods in Density Estimation. New York: Springer-Verlag.

Eggermont, P.P.B., LaRiccia, V.N, 2001. Maximum Penalized Likelihood Estimation. New York: Springer-Verlag.

Fink, G.A., 2008. Markov Models for Pattern Recognition. Berlin: Springer-Verlag

Horton, G., 2002. A new Paradigm for the Numerical Simulation of Stochastic Petri Nets with General Firing Times, 14th European Simulation Symposium. October 2002, Dresden, Germany.

Krull, C., 2008. Discrete-Time Markov Chains: Advanced Applications in Simulation. Doctoral Thesis, Otto-von-Guericke-University Magdeburg, Germany.

Krull, C., Horton, G., 2009. HIDDEN NON-MARKOVIAN MODELS: FORMALIZATION AND SOLUTION APPROACHES, 6th Vienna International Conference on Mathematical Modeling. February 2009, Vienna, Austria.

Krull, C., Buchholz, R., Horton, G., 2010. MATCHING HIDDEN NON-MARKOVIAN MODELS: DIAGNOSING ILLNESSES BASED ON RECORDED SYMPTOMS, 24th European Simulation and Modeling Conference, October 2010, Hasselt, Belgium.

Mamitsuka, H., 1998. Predictiong Peptides that bind to MHC molecules using Supervised Learning of Hidden Markov Models, PROTEINS: Structure, Function and Genetics, Vol. 33, 460-474

Marsland, S., 2009. Machine Learning: An Algorithmic Perspective. Boca Raton, FL: CRC Press

Silverman, B.W., 1998. DENSITY ESTIMATION FOR STATISTICS AND DATA ANALYSIS, Monographs on Statistics and Applied Probabilty. London: Chapman and Hall.

Wickborn, F., Horton, G., Heller, S., Engelhard, F., 2005. A General-Purpose Proxel Simulator for an Industrial Software Tool, 18th Symposium of Simulation Techniques. September 2005, Erlangen, Germany

A CONSTRUCTION KIT OF FLEXIBLE IT-SERVICES FOR SUPPLY CHAIN PLANNING AND OPERATIONS

Sebastian Steinbuss[(a)], Katja Klingebiel[(b)], Gökhan Yüzgülec[(b)], Tobias Hegmanns[(b)],

[(a)]Fraunhofer Institute for Software and Systems Engineering
[(b)]Fraunhofer Institute for Material Flow and Logistics

[(a)]Sebastian.steinbuss@isst.fraunhofer.de, [(b)]{katja.klingebiel, goekhan.yuezguelec, tobias.hegmanns}@iml.fraunhofer.de,

ABSTRACT

Current logistic IT systems tend to be monolithic applications with low flexibility and low awareness of interoperability. Modern IT trends like Service Oriented Architectures and Cloud Computing are key factors to provide convenient access to customized IT-Services on demand and as a service. Hence, the idea of the concept "Logistics as a Service" is to provide domain specific IT services to support the planning and operations in supply chains which are supported by infrastructural technologies. In this paper, the general concepts behind "Logistics as a Service" as well as exemplary IT services are presented and proven to be applicable by demonstration of several use cases. "Logistics as a Service" is one leading topic in the EffizienzCluster LogistikRuhr (EfficiencyCluster LogisticsRuhr) and therefore part of the German High Tech Strategy.

Keywords: Logistics as a Service, Supply Chain Management, Cloud Computing, Service Oriented Architecture, Unified Service Description Language

1. INTRODUCTION

Complex supply networks are the business of Supply Chain Management (SCM). SCM deals with the integrated and process-oriented planning of operations in supply chains from the raw-material supplier to the customer (Kuhn and Hellingrath 2002). The corresponding task model (ten Hompel and Hellingrath 2007) describes the three underlying fundamental business areas of SCM (Klingebiel 2009, p. 40, Hegmanns 2010):

- Supply Chain Design (SCD) covers those strategic and long-term oriented tasks which design and configure the structure, processes and information systems of a logistic system.
- Supply Chain Planning (SCP) covers those long-, mid- and short-term planning activities concerned with anticipation of material requirements, required inventories and resource capacities in a given logistic network
- Supply Chain Execution (SCE) finally covers all operational tasks focussed on order management and order-related operations.

It is common knowledge that IT systems empower logistics managers and planners to deal with supply chain complexity and dynamics in supply chain planning and operations. Yet, available IT systems are just as diverse as their respective tasks: Advanced Planning and Scheduling Systems (APS) and Enterprise Resource Planning (ERP) systems focus on supporting interorganisational and intraorganisational planning tasks, while Warehouse Management Systems (WMS), Transport Management Systems (TMS) and Manufacturing Execution Systems (MES) manage supply chain operations (see for example ten Hompel 2012). Nevertheless, these systems are mostly designed monolithically, thus implying high effort in individual deployment and configuration (Klingebiel and Wagenitz 2012). In addition, the compatibility of these software solutions and the ability to be integrated easily into existing IT infrastructures is still inadequate (Delfmann and Jaekel 2012, page 14).

Against this background, there exists an increasing demand for adaptable and modular (IT) systems which can be quickly and cost-effectively harmonized with other supply chain systems. In this context, the term "Logistics as a service" stands for a new generation of service-based instruments for the management of supply chains which provide flexible decision support through appropriate IT functionalities (Leukel et al. 2011). Following this idea, innovative decentralised and interoperable IT services and infrastructural elements are being developed based on concepts like service-oriented architectures (SOA, Dostal et al. 2009) and cloud computing (Buyya 2009) within the projects of the EffizienzCluster LogistikRuhr (www.effizienzcluster.de).

This paper presents the first deliverables: A construction kit of flexible IT services for supply chain planning and operations. After discussion of the current state of the art in chapter 2, the architectural concept is presented in chapter 3. Chapter 4 describes and reviews the prototypical application of exemplary services in industrial practice. Chapter 5 concludes with a short summary and an outlook on further work.

2. STATE OF THE ART

Today's IT systems are designed by cost-efficiency-principles (Christopher & Beck 2004). Whereas these systems operate efficiently under stable environments, performance declines significantly under volatile and dynamic conditions (Tang and Tomlin 2008). Furthermore, a recent survey conducted by the German Logistics Association confirms that operational Supply Chain Management requires more IT-functions

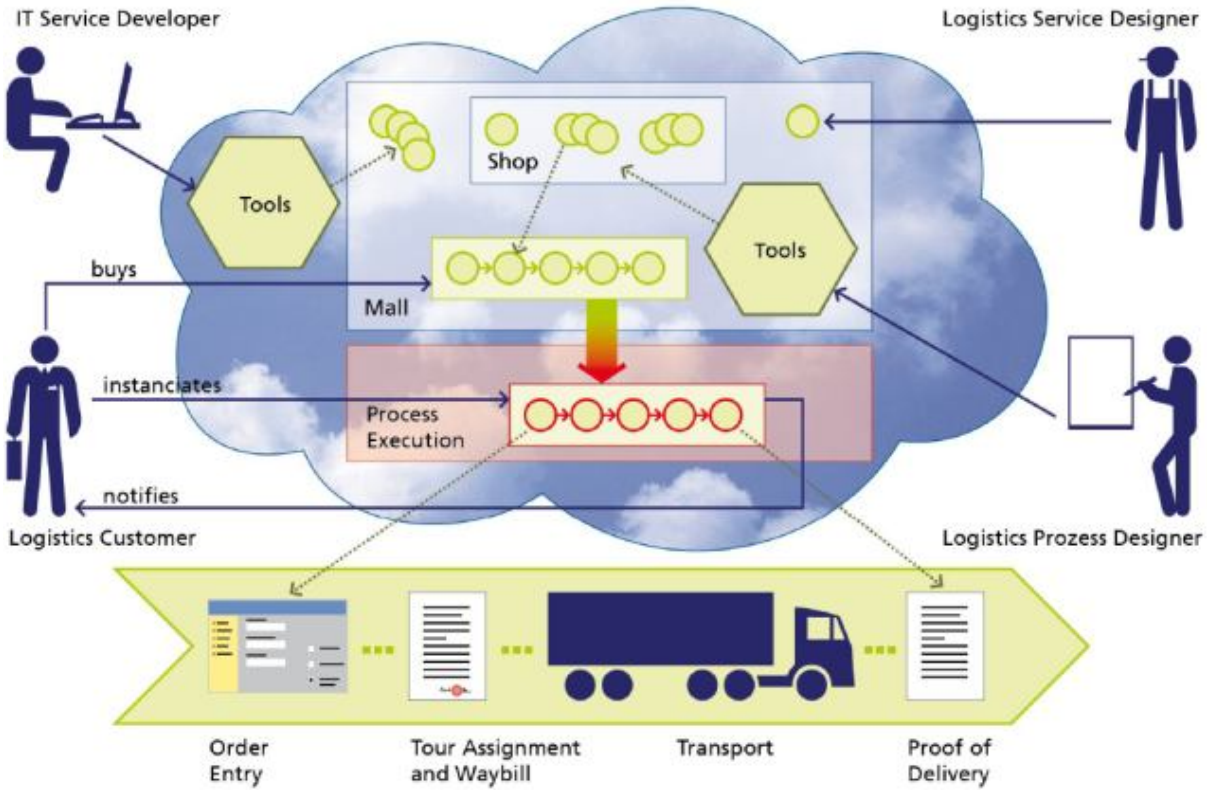


Figure 1. Logistics Mall business architecture

than currently available SCM-IT solutions offer (BVL 2012). Hence, increasing the interoperability of the systems and creating flexible IT architectures, which interlink functions of various systems and allow for an easy integration of new functionalities, seems to be a promising path to follow.

Technologies that capture process information from supply chain operations have improved significantly in the last decades. Advances in RFID and AutoID technologies have led to opportunities for new informational transparency in supply chains (Hellingrath and Alberti 2009; Hegmanns and Toth 2011). Nevertheless, the challenge for SCM-IT is to integrate, filter and draw conclusions from the flood of information available from both process event capturing and other Enterprise-IT systems.

Autonomous systems, SOA and cloud computing are on the rise and establish a new generation of decentralized and interoperable systems, so-called Logistics Assistance Systems (LAS). LAS supplement WMS-, TMS-, ERP- and SCM-systems with individual functionalities and combine the performance of modern decision support technologies with the expertise and judgement of the logistics planner (Klingebiel et al. 2011). However, the availability of individual LAS functionalities which allow an integrated logistics IT support by combinability and individual configurability is still non-existent (Klingebiel and Wagenitz 2012).

The technological scaffolding for scalability, flexibility and a strong alignment to business is service-oriented computing or service-oriented architectures with loose coupled services for the flexible configuration of the required set of services and a high cohesion to achieve interoperability. The Organization for the Advancement of Structured Information Standards (OASIS) defines the term SOA as 'A paradigm for organizing and utilizing distributed capabilities that may be under the control of different ownership domains. It provides uniform means to offer, discover, interact with and use capabilities to produce desired effects consistent with measurable preconditions and expectations.' (OASIS 2011). SOA technologies have reached a high maturity. Consistent orientation on SOA principles allows stepping from monolithic systems to configurable sets of standardized logistics functionalities, i.e. to achieve the development of interoperable logistics applications and provide them as a service, ideally via a cloud platform.

According to the Definition of the National Institute for Standards and Technologies (NIST), the term cloud computing is characterised as a model for enabling access to a shared pool of configurable resources in an ubiquitous, convenient and on-demand manner over a broadband network (Mell and Grance 2011). Mell and Grance distinguish the three cloud models: "Infrastructure as a Service (IaaS)", "Platform as a Service (PaaS)" and "Software as a Service (SaaS)". The basic principle can be identified in the allocation of a requested resource on demand and as a service. For example, the top layer of SaaS allows for applications to be provided via a web browser. This principle is transferred to the logistics domain by the concept of "Logistics as a Service" (LaaS).

The Fraunhofer Innovation Cluster "Logistics Mall – Cloud Computing for Logistics" is a project associated with the EffizienzCluster LogistikRuhr. With the idea of "logistics on demand" it shares the vision of Logistics as a Service. The developed Logistics Mall is a virtual cloud platform for logistics IT applications processes and services which can be offered, hired, and then run in the cloud environment (see also

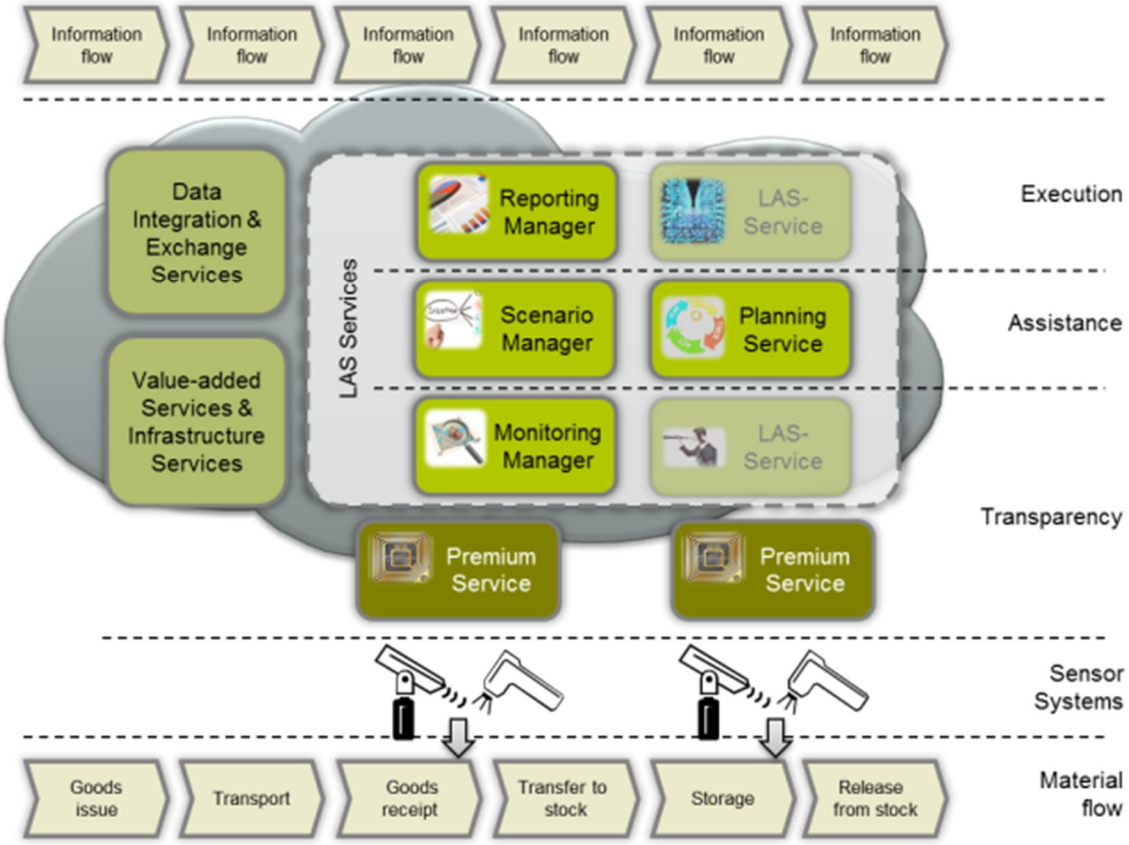


Figure 2. Architectural Framework for Logistics as a Service

http://www.logistics-mall.com). It consists of two main components:

Customers of the Logistics Mall benefit from cost and performance transparency by pay-per-use. Providers benefit from an infrastructure in which a multitude of cloud-based logistics-related services and software can be offered and even orchestrated using business process models (see figure 1). Current market studies prove that such a solution is appreciated by both logistics users and logistics application providers (Fraunhofer IML, 2010).

Today, applications are already made available via the Logistics Mall. These applications communicate through customer-driven data and interface converters. Application integration into the Logistics Mall requires the adoption of the Logistics Mall infrastructure (for details see Gsell and Nagel 2012). This migration may involve steps for:

- Interoperation with other applications (Logistics-as-a-Service, combinable with other services)
- Process-based coordination of applications

Also with the Logistics Mall the authors identified the trend from simple and isolated SaaS-offers to freely configurable and interoperable LaaS-applications. Nevertheless, this new kind of flexibility also leads to new challenges for the logistics domain:

- Standardized IT-functionalities demand the identification of reference processes for logistics and the deduction of standardized planning and operational processes
- These IT-functionalities must provide adequate methods for planning and execution
- The functionalities must then be encapsulated into services which require the provision of standardized interfaces and parameters to guarantee interoperability and a sustainable technological infrastructure to be deployed to.

In the following chapter we provide the first results related to these issues. The next chapter presents a framework for Logistics as a Service. Chapter 4 transfers the presented approach to the specific needs of a use case that has been conducted with an automotive supplier. Chapter 5 concludes with a summary and an outlook on further work.

3. AN ARCHITECTUAL FRAMEWORK FOR LOGISTICS AS A SERVICE

The development of all components is based on the principle of service-oriented architectures. This presupposes a flexible system architecture, which loosely couples professional services and functionalities in the form of autonomous services. This framework comprises four types of services based on open infrastructure services (see figure 2 and following paragraphs for details):

1. Services for data integration and interchange
2. Services for supply chain transparency (Sensor-based premium services)

3. Services for robust logistics planning and operations (LAS Services):
4. Value-added Services and Infrastructural Services

Open interfaces of all services allow docking single functionalities to comprehensive Logistics Assistance Systems (LAS), as well as these LAS to other internal IT systems or those of external network partners.

The interoperability of these systems is facilitated by the development of appropriate data structures in the form of semantic models, known as business objects (BO). BOs characterize relevant objects in a given domain by structured sets of attributes. They are used for the orchestration of applications and services and for the communication between them (Böhmer et al., 2012). A detailed discussion of Business Objects in the context of the Logistics Mall is given in Steinbuss and Weissenberg (2012). In the following paragraphs a detailed insight into the building blocks of the framework shall be given.

3.1. Infrastructure and platforms

Cloud-oriented service-marketplaces like the Logistics Mall (www.logistics-mall.de) target the dynamic provisioning of services or applications with a pay per use cost model. Especially for the German logistics market, which basically consists of small and medium sized enterprises (SME) with a small number of IT-staff, the cloud-based provisioning leads to broad access to professional IT-services (Holtkamp, Steinbuss, Gsell, Loeffeler and Springer 2010). These services or applications are delivered as a Software as a Service (SaaS) offering.

Generally a cloud-oriented service-marketplace consists of an underlying cloud infrastructure. This can be an Infrastructure as a Service (IaaS) or a Platform as a Service (PaaS) offering (Mell and Grance 2011; Stemmer, Holtkamp and Koenigsmann 2011). On top of a cloud-oriented service marketplace is a so-called trading service, which consists of a shop frontend for human users and an enhanced service registry for machines. A trading service can look up services by their functionality and by non-functional aspects, like a price-model, implemented security features (e.g. token-based authentication or encryption) or warranted service level agreements (SLA). In addition, value-added services like security-services, billing services or helpdesk services realize the non-functional aspects and offer other basic services for cross-cutting concerns, e.g. logging. A role model for cloud-oriented service-marketplaces differ the four roles Mall operator, Cloud operator, service provider and service consumer. A Cloud Operator is responsible for the provisioning of a service platform and other technical concerns. The Mall Operator, on the other hand, cares for business concerns and has a contractual relationship to service providers and service consumers.

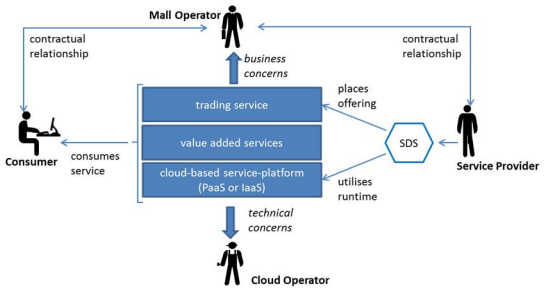


Figure 3. Roles and architecture of cloud-oriented service marketplaces

The project Service Design Studio (SDS) of the EffizienzCluster LogistikRuhr supplies a tool for service providers to place offerings on different marketplaces by utilizing a cloud runtime. Therefore, the functional service description, e.g. a WSDL, and the non-functional aspects are summarized in a semantic service description. A domain part contains a description of the service and its functionality in natural language. The SDS approach doesn't change the implementation of the service. The non-functional aspects encapsulate the service (see fig nn). The main focus of the SDS are the non-functional aspects: Security, Pricing and Service Level Agreements, which are identified as necessary aspects for the logistics domain (Iscan, Flake, Tacken, Ley and Schmülling 2012).

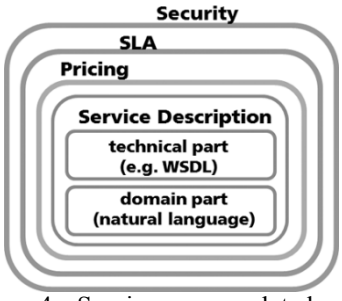


Figure 4. Service encapsulated with non-functional aspects as envelopes

The semantic service description and the executable part of the service get deployed to the marketplace, configure the runtime and get displayed on the trading service. A common format for a semantic service description is the Unified Service Description Language (USDL) (Barros and Oberle 2011). The Service Design Studio develops a logistic profile for the USDL.

A cloud-oriented service-marketplace offers value-added services to service providers. A basic set of value-added services consists of a billing service, security services for authentication, authorization, ciphering and data-integrity, helpdesk-services and a logging service for the traceability of business and technical transactions. A prototypic implementation of these value-added services for a cloud-oriented marketplace is presented by Iscan, Flake, Tacken, Ley and Schmuelling (2012). This infrastructure is a solid

platform that enables domain-specific building blocks to become a construction kit for flexible IT-services.

3.2. Domain-specific building blocks

The EffizienzCluster LogistikRuhr comprises projects which deal with the development of IT-services with regard to the different planning horizons of logistical tasks: strategic planning (Supply Chain Design), tactical planning (Supply Chain Planning) and operations (Supply Chain Execution).

Each of these projects is founded on industry cases of application, in order to validate the solutions in realistic scenarios. The services developed in the two projects, Supply Chain Planning and Supply Chain Execution, shall be presented in the following.

3.2.1. Supply Chain Planning

The project Supply Chain Planning deals with exemplary cases of application within the field oftactical planning. Supply chain planning activities anticipate, calculate and define demand requirements and resource capacities of the logistic network for the mid- to long-term future. Most of the factors influencing planning are uncertain. Therefore, IT-services for supply chain planning should offer the necessary functionalities for the configuration and evaluation of decision alternatives during the planning phase under uncertain knowledge about future demand requirements, market developments or other vague influence factors.

For this reason the building blocks for the described flexible construction kit for supply chain planning IT are defined as follows:

Building blocks for planning scenario configuration assist the planner in developing and configuring planning scenarios. Scenarios may result, for example, from different assumptions about future demand or service-level requirements, or may represent alternative plans, whose performance with respect to the business objectives have to be evaluated. Building blocks in this category are services enabling the user to design decision alternatives based on a data- and model-based representation of the real logistic network. The scope of the services comprises scenario design tools operated by the user based on his problem-specific knowledge and experience, but also more complex planning services which incorporate optimization, simulation or heuristic methods to derive possible solution alternatives to the planning task or sub-tasks.

Building blocks for plan evaluation allow the planner to analyse and study the effects of measures and plans on the logistic performance before deciding on plan realization. Ideally, the evaluation of the decision alternative covers the complete area of influence within the logistic network as well as the impact of dynamic effects. For this reason, simulation services are a substantial component of the planning services in this category. Simulation allows the analysis of complex interdependencies and dynamics of the logistic network and enables the planner to compare decisions in different scenario-based what-if analyses.

Building blocks for the short- to mid-term monitoring of planning performance allow continuous monitoring of the operations and the identification of necessary plan adaptations. These building blocks assist the planner in analysing the short- to mid-term development of the current operational situation given the defined plans and other decisions taken in prior planning steps. Building blocks in this category offer functionalities for identifying current or anticipated future performance gaps. In order to achieve this, information about the current state of operations must be retrieved from the network. This may require interfaces to various operational backbone IT-systems or event data repositories. To extrapolate the current operational situation into the short- to mid-term future, simulation techniques are again useful.

Building blocks for collaborative decision-making represent services for the coordination and harmonization of distributed planning processes between functional units of an organization or between companies. These services offer functionalities for the exchange of (local) planning results for negotiation and coordination within distributed decision-making processes.

3.2.2. Supply Chain Execution

Within the scope of the project Supply Chain Execution different service components are developed; services for the operational capturing (Premium Services) of material flow events and sensor data as well as control services (LAS Services) for the operational control on the materials flow level. The services are web-based and can be used by mobile devices or other systems on the network.

The use case taken from the project contract represents the quality control process in the manufacturing of pieces of furniture. Sensors are used to capture information about the quality of the boards used for cupboard doors. The captured data is provided in near real-time to the decision maker (for example, in production). If the actual quality deviates from the expected quality, the decision maker could use the control services as decision support to control and monitor the entire material flow. By means of the control services, the decision maker is given the opportunity to intervene in the manufacturing process to rework or exchange an inferior part. The following sections describe the design and prototypical implementation of the services for integration, interchange and LAS to explain what control task they support.

Premium Service (Decision preparation) is a collection of individual services, used for the identification and data capture of multiple sensors, that are aggregated by the Premium Service into, for example, a quality score.

Premium Services are responsible for the platform-independent communication with the sensors and provide the LAS Services with a standardized interface using web services. The quality parameters can be easily modified in the Premium Service. The capture of sensor data is performed completely transparent for the LAS Service.

The LAS Service "Monitoring Manager"(Decision preparation) performs the near real-time IT capture and monitoring of the quality information from the Premium Service. This information is transformed into business objects when it is retrieved by the control service and thereby linked with a specific item of wood or production order. Predefined plan values can be used to compare the captured and identified quality score of a board.

The LAS Service "Scenario Manager" (Decision making) provides, in the case of deviating quality in a test process, the decision maker with a list of all control options. Depending on the severity of the quality defect, it might be possible to forego any rework. It is also possible to rework the identified defect directly in the production line or, if it is a very severe quality deviation, take it out of production and rework it outside of the line. If rework is not possible, an alternative part can be allocated for use. The defective part can be assigned to another order and used in another spot. Once a scenario has been selected, a simulation of the selected control measure can be performed to see the future impact of the decision (see Simulation Manager).

The LAS Service "Simulation Manager" (Decision making) uses historical data or proposed decision alternatives from the Scenario Manager to simulate decisions and visualize their impact. A new production schedule is generated based on the selected decision as well as the current existing and forecasted orders. The current and future requirements for parts are presented in the form of Bills of Materials or parts lists. These can be used to simulate the future conditions in the supply chain.

The LAS Service "Optimization Manager" (Decision making) optimizes the restructuring of existing orders. In accordance with the plan sequence, predefined plan values for requested delivery dates or production runs can be adapted to and optimized for unforeseen changes (for example, a lengthy rework of a piece of furniture or a last-minute cancellation of an order that is already being processed).

The LAS Service "Execution Manager" (Decision implementation) facilitates the self-control of the supply chain by invoking one or more services in a specific sequence. Beginning with a quality test by the Premium Service, the Monitoring Manager determines if a quality defect is present. If this is the case, the Scenario Manager is automatically invoked and creates a list of all possible decisions based on predefined plan values. These control options are taken by the Simulation Manager and are simulated. The Reporting Manager checks the results of the simulations. The Execution Manager makes a decision from the generated results based on predefined and prioritized Key Performance Indicators and implements the selected control alternative through the connected management system. This service implements the self-control of the decision implementation.

The LAS Service "Reporting Manager" (Decision implementation and monitoring) can be used to retrieve order-related reports. First, the desired report type is selected (for example, throughput time of orders or stock overviews) and then other options can be specified such as item, orders, time period, and report display options. Restrictions and threshold values, for identifying areas, can be specified for some report types.

The LAS Service "Decision History Manager" (Decision implementation and monitoring) service documents the decisions that were made. All relevant information used to make a decision, including order-related data, is saved from the Scenario und Simulation Manager. The goal is to provide traceability for the decisions that were made because of defective quality for other instances. This information can also be used as a reference for future decisions. All decisions that represent an unscheduled intervention in the material flow can be documented. These documented interventions can be used as reference information and aid in future decision making.

LAS Service "Revision Manager" (Decision implementation and monitoring) links decisions that were made and documented with the Decision History Manager with incoming complaints. This type of information from downstream supply chain stages allows for reflections on the impact as well as the monitoring of decisions that were made.

LAS Service "Quality Assurance" (Decision implementation and monitoring) is the quality-related equivalent of the Reporting Manager. This means that the retrievable reports focus on the quality defects that occurred. The type, number, and length of the rework carried out in the past in a specific process step serve as a reference for future decision making about the control of the material flow when the same or a similar quality defect occurs.

4. USE CASE

The presented framework for a construction kit for flexible supply chain IT is developed and tested in applicational cases with small as well large industrial partners. As an example for the resulting IT tools, the case of Continental tire logistics is lined out in the following.

The use case comprises the production and distribution network for the Continental aftersales and replacement business for passenger and light truck tires. The logistic network contains 9 plants in Europe with 52 Mio. passengers car and light truck tires per year (European replacement market), 23 European markets with regional and central warehouses and more than 700 different articles

Supply chain planning is a challenging task in this case, since the tire business is affected by strong seasonal effects. During the year frequent plan adaptations are necessary to react to seasonal demand developments in the various markets. There is also a strong interdependency of the production planning decisions and the supply chain planning decisions especially concerning the inventory management in the distribution network. Lot-sizing decision and banking measures due to limited production capacities during the peak season influence inventory optimization and product availability in the markets.

For tactical supply chain planning in this use case in total four typical building blocks have been identified and developed. Together these building blocks form a supply chain planning IT-system to be used on-top of the operational backbone ERP systems at Continental. The interaction and information flow between these building blocks is depicted in Figure 3.

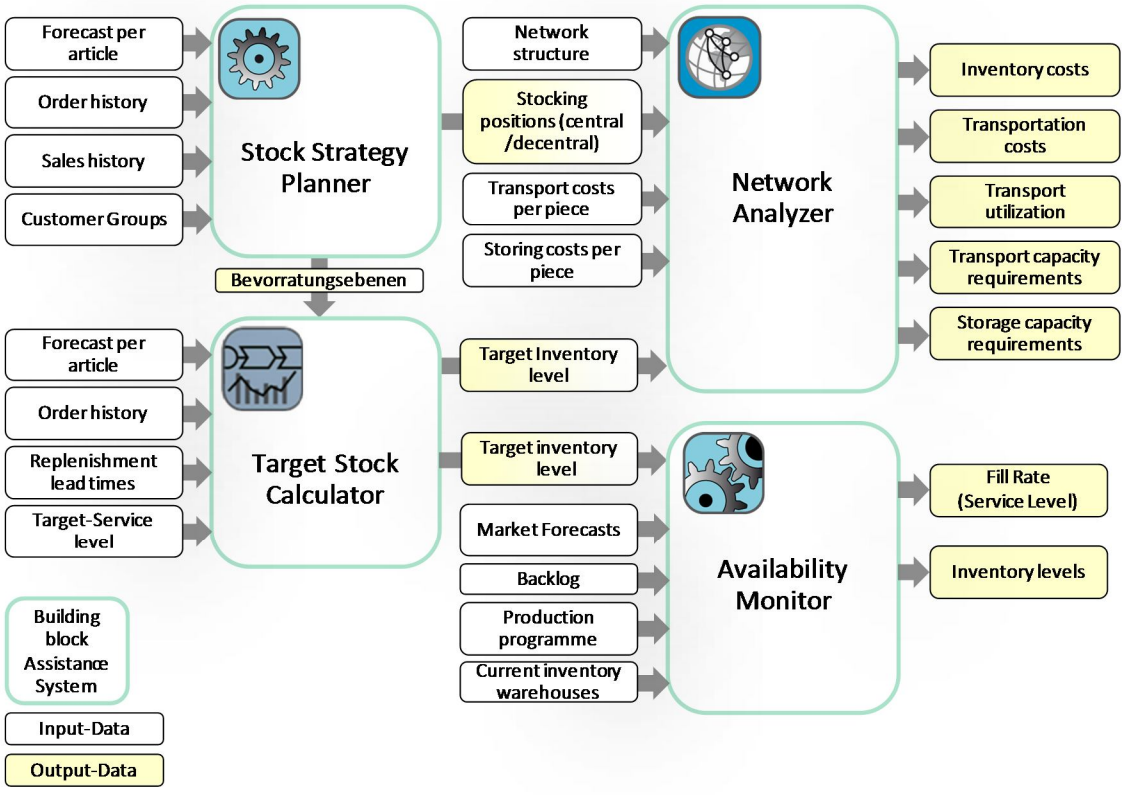


Figure 5. Information flow between building blocks

As a first step on the tactical level, it is necessary to decide on the stocking strategy for each article, i.e. which article is held at which stage of the distribution system (central warehouses vs. regional warehouses) at which point in time. Due to seasonal demand patterns, stocking strategy decisions can vary during peak-season and off-season. To assist with this planning task, the building block "Stock Strategy Planner" was developed. Using the building block "Stock Strategy Planner" the planner determines optimized stocking points. For each article these decisions are derived under consideration of article characteristics (e.g. value, product segment, customer expectations, service targets) and demand information using heuristic methods and clustering strategies offered by the tool. To define clusters, the tool also analyses sales data (historical data as well as forecasts) in order to determine decision-relevant measures such as volatility of demand and accuracy of forecasts.

The building block „Target Stock Calculator" is a tool the planner makes use of to determine relevant disposition parameters, e.g. target stock levels, safety stocks, days of inventory coverage, reorder points. Among others, the tool has adopted innovative methods for the determination of optimal reorder point strategies in multi-stage distribution. As data input the building block makes use of article sales data as well as historic and forecasted sales volumes of the markets.

Referring to the categorization given above, both building blocks, the „Stock Strategy Planner" and the „Target Stock Calculator", are examples of the category

of building blocks required to configure planning scenarios. The planner uses their respective functionalities to specify a planning scenario defining which article is stocked where in what quantity in the network.

The effect of these decisions on the logistic costs and performance of the distribution network can then be analysed using the building block "Network Analyzer". This building block offers a model-based representation of the real network as a basis for simulative evaluation of the network. The planning scenario defined by the results of the other two building blocks can be imported into the "Network Analyzer". A simulation component then evaluates the planning scenario under different demand assumptions.

Last, the building block „Availability Monitor" represents a tool for the rather short-term-oriented operative distribution planning. This building block evaluates the effect of production programme decisions on the demand fill rate of articles in the various markets. Again a simulation-based approach is used to determine the impact of planned production quantities per articles on inventory levels and service-levels per article, warehouse and market. By this, necessary adaptations to the production programme can be identified and their expected effect can be validated via simulation prior to realisation. For this analysis the building block integrates real orders from the current order book, forecasted orders and data from production planning system as well as the current As-Is inventory levels as reported from the warehouses.

In cooperation with Continental Tire Division the described building blocks for a supply chain planning system have been prototypically implemented and tested with real operative data. The building blocks and the planning system they constitute help connect and harmonize distinct activities in demand planning, production planning and inventory and distribution planning. All steps of the planning process receive assistance by model-based quantitative methods and make use of up-to-date supply chain and backbone ERP-information. Also, the time necessary to determine planning results is shortened: Simulation of one planning scenario comprising 9 production plants, several regional and central warehouses, more than 20 European markets and an order volume of 50 million pieces requires less than one hour.

5. OUTLOOK

In this paper the idea of "Logistics as a Service" as seen in the EffizienzCluster LogistikRuhr has been presented. After presentation of the state of the art in relevant research, a framework for Logistics as a Service has been developed. Domain-specific building blocks are the components of a construction kit of flexible IT Services for supply chain planning and operations. Exemplary services have been developed and demonstrated within this paper. Their practical application has been tested in realistic use cases, one of which is the described case of inventory management in the multi-stage production-distribution system of Continental tire logistics. An individual logistics assistant system has been assembled using building blocks from the general building block categories presented in chapter 3. Furthermore, the projects Service Design Studio and Logistics Mall contribute to this framework by providing the necessary sustainable technological cloud infrastructure and infrastructural services. One of the first significant results is the need for a holistic service description language. It could be identified that USDL fits the requirements of such a language, but is far too complex to be applied in industry scenarios. Therefore a logistics profile for USDL has been developed. Further research should analyse the suitability of USDL in other domains and develop similar profiles. Another approach is the decomposition of the USDL to a small core and optional, maybe domain-driven, modules. The research project "Linked USDL" (http://linked-usdl.org) starts with such an approach by application of linked data (Bizer, Heath and Berners-Lee2009).

To further support interoperability additional research is needed in the definition of logistics reference processes and standardized business objects. The presented projects cover specific supply chain management scenarios; nevertheless, other scenarios like warehouse logistics, transport logistics or the special requirements of urban regions are still open work. Here, reference processes need to be developed as the starting point. Only with these references processes, support by modern IT services, as depicted in this paper, may be given.

ACKNOWLEDGMENTS

The projects Service Design Studio, Supply Chain Execution and Supply Chain Planning are funded by the German Federal Ministry of Economics and Technology in the context of the High-Tech Strategy for Germany with support code 01IC10L23, 01IC10L02 and 01C10L01. The projects are part of the EffizienzCluster (www.effizienzcluster.de).

REFERENCES

Barros, A., Oberle, D., 2012. *Handbook of Service Description*. Berlin: Springer.

Bizer, C., Heath, T., Berners-Lee, T.2009. Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems (IJSWIS)*,Volume 5, Issue 3.

Böhmer, M., Daniluk, D., Schmidt, M., & Gsell, H., 2012. Business object model for realization of individual business processes in the logistics domain. In: Clausen, U., ten Hompel, M., Klumpp, M., eds. *Efficiency and Logistics*. Berlin: Springer Verlag.

BVL (Bundesvereinigung Logistik), 2012. *IT in der Logistik*. Hamburg: Deutscher Verkehrs-Verlag.

Christopher, M.; Peck, H., 2004. Building the resilient supply chain, *The International Journal of Logistics Management*, Volume 15.2: pp. 1-14.

Delfmann, W.; Jaekel, F., 2012. The Cloud – Logistics for the Future, in: Delfmann, W.; Wimmer, T. eds. *Coordinated Autonomous Systems*, Bobingen : DVV Media Group.

Fraunhofer IML (2010). Market Study Cloud Computing for Logistics. Retrieved from http://www.ccl.fraunhofer.de/presse-medien/publikationenICS

Gsell, H., Nagel, R., 2012. Application Integration in the Logistics Mall. *Proceedings of the SDPS 2012 - 17th International Conference on Transformative Science, Engineering, and Business Innovation*, June 10-14. 2012, Berlin, Germany.

Hegmanns, T.,2010. *Dezentrales Planungs- und Prozesskonzept für ein kollaboratives BKM in Produktionsnetzwerken*. Dortmund: Praxiswissen.

Hegmanns, T., Toth, M., 2011. Rfid-based Real-time Decision Support in Supply Chains. *Proceedings of the 10th International Conference on Modeling and Applied Simulation*, pp 303-308. 12. – 14. September 2011, Rome.

Hellingrath, B., Alberti, A., 2009. Aufbau und Einführung logistischer Assistenzsysteme auf Grundlage der RFID-Technologie. In: Wolf-Kluthausen, H., eds. *Jahrbuch der Logistik 2009*. Korschenbroich: free beratung Gesellschaft für Kommunikation im Marketing mbH.

Holtkamp, B., Steinbuss, S., Gsell, H., Loeffeler, T., Springer, U., 2010. Towards a Logistics Cloud. *Proceedings of Sixth International Conference on Semantics Knowledge and Grid (SKG)*, pp 305-308, November 1-3 2010, Ningbo China.

Iscan, H., Flake, S., Tacken, J., Ley, M., & Schmülling, C., 2012. Service Design Studio (SDS) – The execution environment of the Service Design Studio. *Proceedings of the SDPS 2012 - 17th International Conference on Transformative Science, Engineering, and Business Innovation*, June 10-14. 2012, Berlin, Germany.

Klingebiel, K., 2009. *Entwurf eines Referenzmodells für Built-to-Order-Konzepte in Logistiknetzwerken der Automobilindustrie*. Dortmund: Praxiswissen.

Klingebiel, K., Wagenitz, W., 2012. An Introduction to Logistics as a Service, In: Clausen, U., ten Hompel, M., Klumpp, M., eds. *Efficiency and Logistics*. Berlin: Springer Verlag.

Klingebiel K., Toth M. & Wagenitz, A., 2010. Logistische Assistenzsysteme. In Pradel, W. H., Süssenguth, W., Piontek, J., Schwolgin, A., eds. *Praxishandbuch Logistik*. Cologne: Fachverlag Deutscher Wirtschaftsdienst.

Kuhn, A., Hellingrath, B., 2002. *Supply-Chain-Management*. Berlin: Springer.

Mell, P., Grance, T., 2012. *The NIST Definition of Cloud Computing*. National Institute for Standards and Technologies Special Publication 800-145. Available from http://csrc.nist.gov [Accessed Apr. 2012]

OASIS, 2011. *Reference Architecture Foundation for Service Oriented Architecture Version 1.0.* OASIS. Available from http://docs.oasis-open.org [Accessed Apr. 2012]

Steinbuss, S., Weissenberg, N., 2012. Service Design and Process Design for the Logistics Mall Cloud. In X. Yang, L. Liu, eds. *Service-Oriented Methodology and Technologies for Cloud Computing*. IGIglobal.

Stemmer, M., Holtkamp, B., Koenigsmann, T., 2011. *Cloud-orientierte Service-Marktplätze – Integrationsplattformen für moderne Dienstleistungsplattformen und IT-Dienste*. Fraunhofer ISST White Paper. Available from: http://www.isst.fraunhofer.de [Accessed Apr. 2012]

Tang, C., Tomlin, B.: The power of flexibility for mitigating supply chain risks, International Journal of Production Economics, Volume 116, pp. 12 -27, 2008.

Ten Hompel, M., Hellingrath, B., 2007. IT & Forecasting in der Supply Chain. In: Wimmer, T., Bobel, T., Ulrich, H., eds. *24. Deutscher Logistik-Kongress und Eurolog. Effizienz, Verantwortung, Erfolg. Kongressband*. Hamburg: Dt. Verkehrs-Verl.

Ten Hompel, M., 2012. *IT in der Logistik*. Hamburg: DVV Media Group.

IMPROVEMENT OF THERMOCHEMICAL FINISHING PROCESSES: AN APPLICATION OF A BATCH TRACING SYSTEM

Robert Schoech[(a)], Ruth Fleisch[(b)], Christian Hillbrand[(c)]

[(a, b, c)]V-Research GmbH, Industrial Research and Development

[(a)]robert.schoech@v-research.at, [(b)]ruth.fleisch@v-research.at, [(c)]christian.hillbrand@v-research.at

ABSTRACT

Manufacturing processes in the area of thermochemical treatment of work pieces impose lots of challenges due to the rough environment. High temperatures and the use of chemicals are aggravating conditions of salt bath nitrocarburizing processes, which are composed of preheating, nitrocarburizing, oxidizing, and multilevel cleaning. These stages are passed through consecutively, for what customer orders are combined in batches. At present, the implementation of batch tracing mechanisms is difficult or impossible due to the harsh environment. As a consequence, wrong assignments of the finished pieces to the customers may occur while a possibility of documentation of the parameters concerning the manufacturing process is desirable. In this paper, we discuss batch tracing by means of intelligent RFID technology (radio frequency identification), which involves connections to sensors for measurements of ambient parameters. This allows the company to offer new supplementary services to the customers and to design the process in a more flexible way like prioritizing time-critical orders. In order to prove the technical feasibility of batch tracing in the field of thermochemical treatment two test cases are described: The first one is for gathering information about the functioning of the installed RFID system, the other one is the realization of the operational process in a simplified form.

Keywords: batch tracing, RFID, harsh environment, production system design, manufacturing process

1. INTRODUCTION

The number of product recalls has been increasing over the past decade. Despite increasing quality, there is no certainty that only top-quality products which are free from errors reach the market. The latest example includes the 2010 Toyota automotive recall. Besides the loss of reputation, the administrative and logistical effort is very high for a recall. The issue of reducing product recalls and minimizes the potential economic damage in case of a recall has been sufficiently discussed in the relevant literature (Engelhardt-Nowitzki and Lackner 2006, Kletti 2006, Potter 2011). Therefore tracking and tracing of products and batches in production control and logistics will be increasingly relevant for industry, final consumer and, last but not least, legislator. Several of the following reasons for the enhanced deployment of batch tracing thus apply (Engelhardt-Nowitzki and Lackner 2006):

- Besides the fulfillment of functional performance of the product, the request of customers for security and information about the origin of ingredients and products is an important part (customer pull effect).
- The industry benefit form increasing in efficiency and advancing of process reliability (industry pull).
- Innovative communication, information and location technology, such as Radio Frequency Identification (RFID) which are combined with Global Positioning System (GPS) or Wireless Local Area Network (WLAN) are standard today and generate a technology push effect.
- New laws and regulations provide document information “one step forward and one step back” in the logistic chain (regulatory push effect).

Norms such as Code of Federal Regulation 21CFR820 (Quality System Regulation 2006) or European Regulation 178/2002 (Regulation No 178 2002) give little scope especially to the food and drug industry. This regulation is a comprehensive system of traceability and allows precisely product recalls. Safety-related mechanical components are subject to similarly stringent requirements. From the manufacturer's product liability for defective products arises across all branches of industry (Regulation No 84/374/EEC 1985). In order to restrict the extent of liability for the individual companies in the value chain, the identification of batches is mandatory. In addition to labeling products, packaging and packaging units will be identifying which ensure a seamless traceability.

In some industries (e. g. food, drug, steel, and thermochemical sectors), the batch production is discontinuous which is characterized by a material flow which is interrupted in time. In this context a batch is an amount of pieces which form a whole and are processed together and therefore exhibit identical attributes with regard to the manufacturing process and the product

quality. Another feature of the batch production is limited capacity of the production system (e. g. thermochemical treatment processing system) and therefore only part of batches can be processed and removed afterwards. Due to the conditions of production the products produced in different batches vary and consequently variations in quality are obligatory (Engelhardt-Nowitzki and Lackner 2006, Günther and Tempelmeier 2000). For quality control purposes the operations of a detailed documentation of the production process are absolutely essential and are increasingly being leveraged in industry.

The ever-increasing demands of the internal material flow are the constant source within automated production control. In terms of material flow system has made great progress in recent years. The current trends are decentralized signal processing of sensors and actuators as well as reusable and exchangeable software components (Günthner and ten Hompel 2010). Intelligent infrastructure coordinates with transport units and allows to implementing a decentralized and variable batch tracing system. Particularly in combination with RFID, the design of the batch tracing system provides a way to satisfy the requirements for discontinuous production.

In this article we therefore propose the broad outlines of our research tasks, for the technical feasibility of an information system for the seamless traceability of work pieces through the stages of discontinuous production. The initial situation and the requirements of the problem are particularized for the real case of the integrants of thermochemical finishing processes. On the basis of the system requirements we discuss the feasibility check by the means of two test design patterns and describe the test results. We introduce the concept of network and system architecture for our batch tracing system. The paper ends with a disquisition of conclusions and outlook of the applicability of this system.

2. APPLICATION

We have attempted to apply the ideas characterized above to a real case from the industry. First the initial situation is described and the problem is presented.

2.1. Initial Situation and Presentation of the Problem

The company is a factory for the production of mechanical components including heat treatment. In the range of heat treatment the company deploys, among other things, the thermochemical treatment of the work pieces by salt bath nitrocarburizing by the TUFFTRIDE® process, which is used to improve the wear resistance, the fatigue strength, and the corrosion resistance of components made from steel, cast iron, and sintered iron materials. The use of chemicals and treatment temperatures of up to 630 °C pose a challenge to issues concerning this manufacturing process (Boßlet and Kreutz). The work pieces traverse various stages in a manually operated plant in sequence. Prior to that the formation of batches takes place by pooling customer orders. Thus the manufacturing process shows the characteristics of a discontinuous production.

At present, there exists a shortcoming in the process of the delivery of the work pieces, their thermochemical treatment, and their provision for supply to the customers: Although photographs on the order prints serve as assistance in recognizing, the refined pieces of a batch cannot be definitely related to the customers occasionally. The main problem is the missing identification of the charge carriers, which would allow an explicit assignment of the work pieces to the batches at the information level at the beginning of the process. Therefore there is a disruption in the information chain during the treatment of the pieces. As a consequence of the absence of batch tracing mechanisms there are wrong returns, which lead to customer dissatisfaction due to missed deadlines and additional transportation costs. Furthermore there exists no possibility of documentation of the parameters concerning the manufacturing process of the work pieces.

Through the implementation of batch tracing mechanisms the following objectives shall be accomplished in the medium term:

- In order to meet the increasing requirements of the customers, the company pursues the target of offering new services for quality management by documenting the manufacturing process of the work pieces. Essentially this is an automated analysis of the actual duration of the process and the temperatures in the individual stages of the salt bath nitrocarburizing plant for the work pieces. The compliance with regulations during the finishing process shall be attested by the help of greater transparency of the quality data. This important aspect of quality assurance shall reduce the product liability risk on the one hand and precisely determine possible causes of failure on the other hand.
- The identification and localization of the batches throughout the entire process shall help to design the process in a more flexible way like prioritizing time-critical orders.
- By assigning the refined pieces to the customers easily and more quickly and avoiding wrong returns with the aid of batch tracing, the non-productive time shall be reduced by 20 % and thereby the throughput shall be increased by 30 %.

The new supplemental services based on process documentation and prioritization of orders shall cause an enhancement of the market reach, by what a medium-term sales increase of more than 5 % is expected.

The idea is to realize the batch tracing system by means of intelligent RFID technology (radio frequency

identification), which involves connections to sensors for measurements of ambient parameters of the process. For this purpose a feasibility study is conducted as a first step, in order to check the technical practicability of using an RFID system for the automatic identification of charge carriers, because in the field of implementation of RFID technology there is no standard solution, but rather the solution has to be adapted to the specific situation, particularly in this case due to the aggravating conditions of salt bath nitrocarburizing processes. However, the study does not comprise the issue of gaining and handling the temperature data. We now want to present the study und its results.

2.2. Process Analysis

The process analysis begins with the ascertainment of the current process of thermochemical treatment including the receipt of the work pieces and the allocation for their returning. Subsequently, we look into the question of how to apply RFID technology as part of the batch tracing mechanisms in the present case, prior to expanding the sequence of the finishing process by the handling of the transponder solution.

The work pieces delivered for refinement are electronically recorded and photographed in the receiving department and then forwarded to the department of thermochemical treatment together with the printed order copy. Here the consolidation of the customer orders to batches is carried out consecutively: The pieces are put into baskets, which are arranged in a pile forming the charge carriers. The number of baskets per charge carrier can vary depending on the size of the work pieces. After the cleaning in a washing plant, the next step is the thermochemical treatment of the work pieces by salt bath nitrocarburizing by the TUFFTRIDE® process, which is composed of preheating (PH), nitrocarburizing (Tufftride), oxidizing (AB1), and multilevel cleaning. These stages are passed through successively (see Fig. 1). Nitrocarburizing takes place in a salt bath at 480 to 630 °C, the standard temperature is usually 580 °C (Boßlet and Kreutz). The treatment is done in a manually operated plant: A hand-operated crane transfers the charge carrier from stage to stage. The crane can be positioned with an accuracy of about 10 cm. Several cranes run along a closed crane rail, hence the sequence of the cranes does not change.

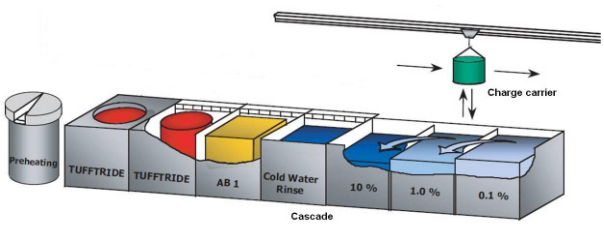


Figure 1: Tufftride®-Plant (Boßlet and Kreutz)

Two identical stages are available for nitrocarburizing however the batch carrier has to pass through only either stage. Since the stages are arranged one after another and the order of the cranes is invariant, it may occur that a charge carrier has to wait in front of the first nitrocarburizing stage, although the second one is vacant. In order to avoid such situations there is also the possibility of relocating from one to the other nitrocarburizing stage, because the procedure of nitrocarburizing can be interrupted.

After the thermochemical treatment lasting three to four hours, the process can be continued in three different ways:

- Ending of the process: The work pieces are provided for delivery to the customers. The assignment of the refined pieces to the customers is effected visually at what the photographs on the order prints serve as assistance in recognizing.
- Sandblasting and afterwards ending of the process.
- Sandblasting followed by assignment to a new batch and repeated thermochemical treatment.

Batch tracing shall enable the identification of the work pieces on the one hand and the possibility of documenting the duration of the treatment and the temperature in the different stages of the salt bath nitrocarburizing plant for the work pieces on the other hand. The identification of the work pieces is effected by the assignment to a batch, which is associated to a charge carrier afterwards. RFID technology is deployed for the automatic identification of the charge carriers. However, affixing the transponder directly on the charge carrier is not realizable due to the harsh conditions during the thermochemical treatment – but neither necessary, as the charge carrier can be unambiguously assigned to the crane, which is used for the transportation of the charge carrier. The crane is now equipped with a transponder. The fastening location of the transponder at the crane is situated up close to the crane rail and hence several meters away from the immediate environment of the thermochemical events. As a result, the requirements for the RFID system, especially for the transponders, reduce to an acceptable level. So the transponders are not attached to work pieces, packaging units, pallets, or charge carriers as is usual but to the means of transportation of the charge carriers. For every stage to be documented a reader is mounted on the crane rail. They are connected to a computer via Ethernet. The readers can be regarded as nodes of a network yet to be established where other parts are for example the sensors for measuring the parameters of the manufacturing process like temperatures or components which combine the RFID data, temperatures, and other data and edit them for further use.

In concrete terms, the handling of the RFID solution is integrated into the process of thermochemical treatment in the following way: The work pieces and the customer orders, respectively, are aggregated in batches and assigned to a charge carrier,

which is marked by a metal label with a unique number. When the charge carrier is hitched to the crane after the cleaning, the mating of them takes place by reading the ID of the transponder of the crane, manually entering the number of the charge carrier into a PC, and relating both IDs. The readers in the different stages record the transponder as long as it is located there. Thus the duration of stay of a crane in a stage is logged by use of RFID technology and one obtains the desired data for the particular orders such as the beginning and the length of stay via the assignments crane – charge carrier and charge carrier – order.

2.3. Technical Feasibility Check

In order to check the technical feasibility of the transponder solution, requirements on the RFID system are specified and appropriate components are chosen followed by the definition of the test cases and the analysis of the data gathered by the tests.

The requirements on the RFID system arisen from the process analysis are a reading range of 10 to 20 cm, passive energy supply of the transponders, that is transponders without a battery, and fixed read/write devices. Bulk reading is not necessary. As the RFID system has to run in an industrial environment, it should feature a high degree of robustness. Based on these specifications an HF system is selected meaning that the RFID system operates in the high frequency range (13.56 MHz). The components of the RFID system should conform to the norm ISO 15693 which comprises a standardization of RFID technology in the HF range. According to the determined requirements the RFID equipment is selected for the desired application of batch tracing in the salt bath nitrocarburizing plant and is used during the tests: discoid transponders of 8.5 cm in diameter, readers with integrated antenna, and communication modules coming with an Ethernet connection to link readers to a computer.

In order to examine the RFID system under conditions close to reality, the tests are performed at the company on site. Transponders are attached to two of the eight cranes used for the transport of the charge carriers with the work pieces to be refined during the thermochemical treatment. The two cranes succeed one another directly. During the experiment two stages shall be monitored and documented: The first one is the stage “Start”, where the charge carrier is hitched to the crane and the mating of both of them takes place, and the other one is “PH”, the stage for preheating. A software application has been developed for this feasibility study. With the aid of this prototype the data and the RFID components can be managed. This includes among other things the establishment of a connection to the read/write devices, the receiving of messages concerning transponder events like the entering and leaving of the antenna field, and the assignment of a charge carrier identification number to a transponder just being read.

Two test cases are defined for the experiment:

- *Scenario A:* This test case is for gathering information so as to check whether the installed RFID system operates without technical problems or whether the rough environment has negative effects on the individual components of the RFID system. For this purpose it is continually recorded when the transponders enter or leave the antenna fields. The capture of these transponder events happens automatically without manual inputs being necessary, because the assignment of the batches is not taken into consideration in this scenario.
- *Scenario B:* This test case is the realization of the operational process in a simplified form. After the charge carrier is hitched to the crane for the treatment in the plant, their mating takes place in the stage “Start” by entering the unique charge carrier number and a description of the batch into the software application and assigning them to the transponder of the crane. By recording the entering and leaving of the antenna fields one gains the requested data about arrival time and length of treatment of the work pieces by batch tracing mechanisms exemplarily for the preheating stage.

For the reason that the data are available for the analysis, a text file is generated for each of the two test cases, where the recorded transponder events are registered. As opposed to *scenario A*, in test case B the transponder events are only logged if the transponder ID is related to a charge carrier ID for the respective passage through the thermochemical treatment. One data set is essentially composed of the name of the stage (“Start”, “PH”), the name of the transponder (“Ida”, “Ada”), the date and time of the entry and the exit respectively, the length of stay of the transponder in the antenna field, and additionally in *scenario B* the charge carrier number and the short description of the batch. Besides, data are manually recorded in case B, so that it is possible to check the RFID data against them.

After having installed the RFID system, it has been running for 20 days and collecting data continuously. The analysis of the data concerning *scenario A* yields a positive result in this test case. This means that the components of the RFID system, particularly the transponders and readers, operate satisfactorily and that negative effects cannot be detected which may be caused by harsh ambient conditions like high temperatures and the use of chemicals or by other potential interfering influences such as the metal parts of the crane or the crane rail where the transponders and read/write devices are mounted. The data of *scenario A* do not show a breakdown of a transponder or a reader meaning that it does not occur that a transponder is not read anymore or a reader does not record transponders anymore from a certain point in time. A data set, which is missing in the consistent succession of the data, along

with a message about a thrown exception give an indication that, however, an interruption of the computer operation has happened at least once (see also *scenario B* below). Next we address the question of how many times a transponder enters and leaves the antenna field of a stage before the following transponder is read in this stage. Reasons for repeated entering and leaving are for example the positioning of the crane in order that the charge carrier is located exactly above the container for the treatment, into which the batch is lowered, or that the transponder is situated on the margin of the antenna field and enters and leaves it unintentionally. The transponder sojourns in the antenna field only once in a little more than half of cases and twice in approximately a quarter of cases. It is recorded three, four, …, or seven times in some cases and once even twenty times. The reading time is defined as the duration between the arrival of the transponder at and its departure from the antenna field. The acquired reading times of the events exhibit a few outliers: three in the stage "Start" and two in the preheating stage with values between 9 and 21 h. The explanation for these long terms is to be found in the circumstance that sometimes the batches are already positioned for later treatments in the stage "Start" or "PH" the previous day and thereby stay overnight or even, for example, from Sunday morning until Monday morning. The arithmetic mean of the reading times, calculated without the outliers, is 4 min for the remaining 167 events of the stage "Start" and 25 min for the remaining 118 events of the preheating stage. The reading times gathered in "Start" are distributed in the following way: 2 % of the events have values greater than 70 min (outliers), 10 % between 5 and 70 min, 28 % between 0.5 and 5 min, and 60 % less than 0.5 min. The duration of 46 % of the events is less than 3 s in fact (see Fig. 2). The reading times of the preheating stage show accumulations of the values in the ranges between 80 and 100 min (15 %), between 10 and 50 min (33 %), and less than 0.5 min (42 %). One third of the events last less than 3 s actually (see Fig. 3). The analysis of the reading times will serve as a basis for the determination of minimum reading times for each stage, which have to be reached in order that the event is registered, enabling to filter out the undesired events caused by positioning of the crane for example.

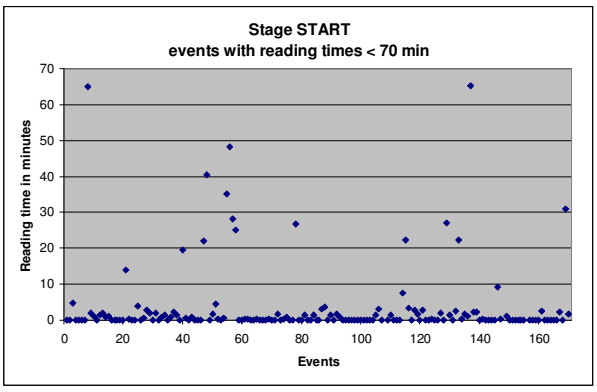


Figure 2: Reading Times for "Start"

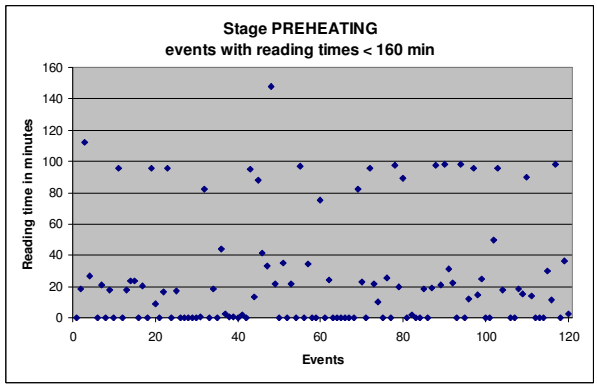


Figure 3: Reading Times for "PH"

In *scenario B*, the integration of the handling of the RFID system into the existing operational process, more precisely the entry of the charge carrier number together with a batch description into the computer and their assignment to a transponder, has been adopted by the employees of the company better than expected. The comparison of the manually recorded and the RFID data reveals that data sets are missing. The stay of the second transponder in the stage "PH" is not registered once and, at some other time, the data of the first transponder in the preheating stage and the ones of the second transponder in both of the monitored stages are lacking. In other words, the ending parts of two passages through the plant are not captured. However, the transponder events of the passages following the incidents are logged without any trouble-shooting being necessary or further problems then. This in combination with error messages of the software application implies that interruptions of the computer operation have occurred. For the analysis regarding the time data the manually recorded arrival times of the charge carriers are collated with the corresponding RFID data. The departure times are compared in an analogous manner. The greater of these two absolute differences is determined for each data set and is denoted by "D" hereafter. Only the data sets of the relevant preheating stage are taken into account and whose manually recorded data are complete additionally. The maximum value for "D" of the 40 data sets is 2.1 h. One quarter of the considered data has differences of less than one minute, another 30 % of the data between one and four minutes, and for the remaining 45 % the value "D" is greater than four minutes. The last-mentioned 18 data sets comprise two cases, where the manually recorded and the RFID length of stay are approximately the same but do not take place at one time, eleven situations the other way round, and five cases, where neither the lengths of stay nor the beginning or ending of the reading times of the manually collected and RFID data coincide. Several reasons for the deviations between the manually recorded and the RFID data are conceivable:

- The length of treatment of a batch manually logged does not equate to the duration of stay of the crane in the stage recorded by the RFID

system, because the charge carrier is just waiting in the stage and is not lowered into the container for the treatment. In order to acquire the precise instant of time when the batch plunges into and emerges from the container for treatment, supplementary sensor technology would be required. However this is not necessary according to the company, since the charge carrier is not situated in a relevant stage like the one for nitrocarburizing without being treated, but is transported to the next step of the thermochemical treatment immediately.

- The transponder is not read all the time although it is located in the antenna field. (In the cases, where the lengths of stay differ, the RFID reading time is longer than the manually recorded one eleven out of sixteen times and therefore this reason could come into effect only for the remaining five events.)
- Errors when logging the data manually, as this takes places during the daily business.
- Minimal differences can be caused by the fact that the manually gathered data are specified in minutes whereas the RFID data in seconds.

3. GENERAL DESIGN PRINCIPLES

The following chapter provides a general overview of the network concept and system architecture for our batch tracing approach.

3.1. Concept of the Network

Based on the proven technical feasibility of the RFID solution, the subsequent main issue of implementing a batch tracing system is the establishment of the network, which comprises not only RFID devices like transponders and readers but also sensors for measuring the ambient parameters of the manufacturing process (e.g. temperatures) as well as components which combine the RFID data, temperatures, and other data and edit them for further use. Other elements of the network manage all these parts. In order to link the nodes of the network, an appropriate communication medium is necessary. A usual standard for this in an industrial environment is the connection of sensors, actuators, and control devices by means of buses. With increasing frequency, Ethernet is used beside traditional bus systems like Profibus, although it is not real-time capable (Gevatter and Grünhaupt 2006). However, response times in the range of milliseconds are sufficient in many applications – as for instance in the present one. Hence we have already deployed Ethernet for the connections between the read/write devices and a computer in the feasibility study. The use of Ethernet technology in an industrial environment for automation and process control is termed by “Industrial Ethernet”. It includes the definition of cables, connectors, and the topology of the bus system (Schnell and Wiedemann 2006).

3.2. System architecture

The kernel of our batch tracing system is based on an existing generic application platform. The latter will be implemented adhering to the principles of service-oriented architecture (SOA) (Gioldasis et al.2003, 12f). The application platform provides base classes with generic services, which can be used to implement domain-specific batch tracing environments.

Service orientation models the manufacturing process into physical (e.g. readers or sensors) and virtual devices (e.g. stages or charge carriers). A physical device is specified by services which describe process-relevant information and is represented as atomic module and cannot be further broken down. A virtual device is a set of application functions for the manufacturing process. Physical devices are related to virtual devices and provide status information about the material flow. In this way the manufacturing process based on the “Industrial Ethernet” can offer his functionalities through services. The main intention has been to enable domain experts (i.e. shift manager or supply chain manager) for using batch information without the need for detailed manufacturing process expertise.

One of the next steps in the course of the implementation of batch tracing mechanisms will be the elaboration of this concept of the network and system architecture in more detail.

4. CONCLUSIONS AND OUTLOOK

Summarizing the results of the technical feasibility check, there are not detected any criteria against the application of RFID technology for the intended identification of charge carriers or rather of their means of transportation in harsh surroundings of thermochemical manufacturing processes. Though the question of the impact of incrustations, which are caused by the thermochemical processes and which are formed over time, has not been cleared up in the course of the three-week testing phase, the contaminations should not become a problem, since the selected components of the RFID system feature a high degree of robustness and are designed for the employment in industrial environments.

The technical feasibility proved by the conducted study serves as basis for the further implementation of batch tracing mechanisms in a salt bath nitrocarburizing plant. The next step for this purpose is the establishing of the RFID system for all stages to be monitored and all cranes and also the elaboration of an appropriate overall concept. An item to be approached for the draft is the definition of use cases in detail. The issue of the mating of the charge carrier and the crane, which enables the identification of the batches throughout the entire process, is necessary as a prerequisite. Besides the resultant possibility of documenting the beginning and length of stay of the treatment for the work pieces, another use case is the deposit of schedules for the treatment times of a batch in order that flares or audio signals can indicate when the treatment of the batch is

finished in one stage and the charge carrier can be transferred to the following stage. Additionally the overall concept includes the design of the information flow that is the interfaces to other systems, the software architecture of the control system, and the hardware topology. Apart from the implementation of an RFID system, a further objective to be accomplished is the gathering of ambient parameters of the process, like temperatures, measured by sensors of the plant and their linkage to RFID data in order to realize a batch tracing system for thermochemical finishing processes by means of intelligent RFID technology.

ACKNOWLEDGMENTS

This paper discusses the results and findings of a research project within the K-Project "Integrated Decision Support Systems for Industrial Processes (ProDSS)" which has been financed under the Austrian funding scheme COMET (COMpetence centers for Excellent Technologies).

REFERENCES

Boßlet, J. and Kreutz, M.. *Tufftride®-/QPQ®-Process. Technical Information.* Durferrit GmbH. Available from: http://www.durferrit.de/media/pdf/Tenifer_QPQ_eng.pdf [accessed 14 December 2011].

Engelhardt-Nowitzki, C. and Lackner, E., 2006. *Chargenverfolgung. Möglichkeiten, Grenzen und Anwendungsgebiete.* Deutscher Universitäts-Verlag Wiesbaden.

Gevatter, H.-J. and Grünhaupt, U., 2006. *Handbuch der Mess- und Automatisierungstechnik in der Produktion.* Springer-Verlag Berlin Heidelberg.

Gioldasis, N., Moumoutzis, N., Kazasis, F., Pappas, N., Christodoulakis, S., 2003. *Service Oriented Architecture for Managing Operational Strategies*. ICWSEurope, LNCS 2853, P. 11-23

Günther, H.-O. and Tempelmeier, H., 2000. *Produktion und Logistik.* Springer-Verlag Berlin Heidelberg.

Günthner, W. and ten Hompel, M., 2010. *Internet der Dinge in der Intralogistik.* Springer-Verlag Berlin Heidelberg.

Kletti, J., 2006. *MES Manufacturing Execution System – Moderne Informationstechnologie zur Prozessfähigkeit der Wertschöpfung.* Springer-Verlag Berlin Heidelberg.

Potter, A., 2011. *The big picture: key trends in international product recalls within the food industry.* Paper presented at the ASSET Conference 2011, Belfast, 21 – 24 March 2011.

Quality System Regulation, Code of Federal Regulations 21 CFR Part 820, Subpart F – Identification and Traceability 2006. Available from: http://www.medicom.ie/images/stories/downloads/QS%20Regualtion%2021%20CFR%20Part%20820.pdf [accessed 12 April 2012].

Regulation No 178/2002 of the European Parliament and of the Council, Article 18 Traceability 2002. Available from: http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2002R0178:20090807:EN:PDF [accessed 12 April 2012].

Regulation No 84/374/EEC of the European Parliament and of the Council 1985, liability of defective products 1985. Available from: http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:1985:210:0029:0033:EN:PDF [accessed 12 April 2012].

Schnell, G. and Wiedemann, B., 2006. *Bussysteme in der Automatisierungs- und Prozesstechnik. Grundlagen, Systeme und Trends der industriellen Kommunikation.* Vieweg Verlag Wiesbaden.

AUTHORS BIOGRAPHY

Robert Schoech is project manager at V-Research, an Austrian competence center for industrial research and development. Within the research area "Technical Logistics" the research center focuses on methods to support complex decisions in manufacturing and logistics processes. Robert Schoech's work focuses on simulation-based systems for industrial processes. His research interests also include track & trace solutions based on GPS, GSM, and RFID technology.

Ruth Fleisch is research associate at V-Research. After receiving her diploma in mathematics from the University of Innsbruck she joined V-Research in 2010. Her research focuses on discrete-event systems as well as RFID technology.

Christian Hillbrand is head of the research area „Technical Logistics" at V-Research and director of the research programme "ProDSS" (Integrated Decision Support Systems for Industrial Processes). He holds a PhD-degree in business information systems from the University of Vienna and works on decision-supporting frameworks for logistics and planning processes.

SIMULATION OF UNCERTAINTY IN RAINFALL-RUNOFF MODELS AND THEIR STATISTICAL EVALUATION IN THE FLOREON+ SYSTEM

Štěpán Kuchař[(a)], Tomáš Kocyan[(a)], Pavel Praks[(a)], Martina Litschmannová[(b)], Jan Martinovič[(a)], Vít Vondrák[(a)]

[(a)]VŠB - Technical University of Ostrava
IT4Innovations
17. listopadu 15/2172, 708 33 Ostrava, Czech Republic

[(b)]VŠB - Technical University of Ostrava
Faculty of Electrical Engineering and Computer Science
17. listopadu 15/2172, 708 33 Ostrava, Czech Republic

tomas.kocyan@vsb.cz, stepan.kuchar@vsb.cz, pavel.praks@gmail.com, martina.litschmannova@vsb.cz, jan.martinovic@vsb.cz, vit.vondrak@vsb.cz

ABSTRACT

Floods are the most frequent natural disasters affecting the Moravian-Silesian region. Therefore a system that could predict flood extents and help in the operative disaster management was requested. The FLOREON+ system was created to fulfil these requests. This article describes statistical evaluation of the rainfall-runoff models in the FLOREON+ system and modelling of uncertainty in the environment parameters of the model. The Monte-Carlo simulation method is used for estimating possible river discharge volumes based on the uncertainty of precipitation and meteorology forecast and provides several confidence intervals that can support the decisions in the operational disaster management. Experiments with other environment parameters and their influence on final river discharge volumes are also discussed.

Keywords: uncertainty modelling, Monte-Carlo simulation, rainfall-runoff modelling, river discharge volume statistical evaluation

1. INTRODUCTION

The main goal of the research project FLOREON+ (FLOod REcognition On the Net) is a development of prototypal open modular system of environmental risks modelling and simulation is based on modern internet technologies and platform independency. The final product of the project is going to be the system offering online communicational man-machine interface. The project results should help to simplify the process of crisis management and increase its operability and effectiveness. The main scopes of modelling and simulation are flood risk, transportation risk and water and air pollution risks. Another efficient utilization of the computing power could be computing the scenarios for the decision support. The prediction of land cover and land use changes (LULC) based on the thematic data collection (aerial photographs, satellite imagery) and application of the prediction tools bring attractive advantages to land use planning. Modelling the catchment response to severe flood events brings a possibility to improve the proposal of channel system set up and dimensioning in the scope of hydrology and water management.

2. RUNNING HYDROLOGICAL SIMULATIONS ON HPC IN THE FLOREON+ SYSTEM

HPC as a parallel environment is able to run many hydrological simulations at the same time. This allows the users to use the environment effectively and shortens waiting times for simulation results even during the high level of demand (e.g. during critical situations). Parallel computing is also very useful for model calibration in which many simulations with different calibration parameters can be run simultaneously and their results can be compared gradually.

However this comes with an implementation cost because used simulation models are not ready for such simultaneous launching. We had to solve this problem by creating multiple simulation environments integrated with preparation and finalization code. We named these functional environments Simulators and created one instance for each node and computation core that would be used to perform simulations.

Therefore when a user needs to run a simulation, he uses FLOREON+ system's Simulation Application to create new simulation and fill it with desired attributes based on the model he wants to use. The Simulation Application then calls the Run Model Web Service deployed on the HPC server and sends all given parameters. This web service utilizes the HPC environment to find a suitable Simulator instance in the pool of available instances (see Figure 1). The chosen Simulator prepares the required model and asks FLOREON+ Core Web Services for rainfall data, snow thickness, temperature and other data saved in the central FLOREON+ Database. These are used as input data to the model and the Simulator starts the simulation. Results of the simulation are sent to the

FLOREON+ Core Web Services to be saved in the FLOREON+ Database for future use. At the same time the resulting hydrographs are displayed to the user in the Simulation Application and the Simulator instance is returned to the pool of available instances.

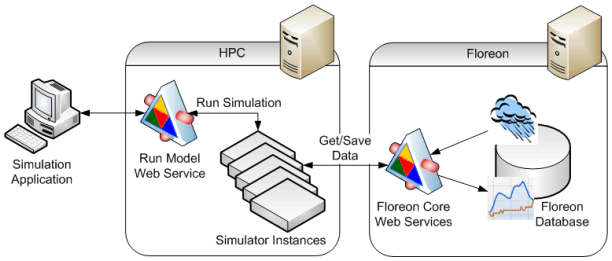


Figure 1: Running Hydrological Simulations on HPC in the FLOREON+ System

3. MODEL VERIFICATION IN THE FLOREON+ SYSTEM

It is necessary to run a lot of variants of the rainfall-runoff model to verify this model and their result are then used to run several variants of hydrodynamic simulations. The rainfall-runoff simulations take about 4 minutes to run on a one processor computer and hydrodynamic simulations take more than 1 hour. The whole simulation cascade can therefore take several days to complete, but this is not feasible in the FLOREON+ system that is intended for decision making support within operational disaster management. Since there is quite a big number of computation operations needed in order to compute the whole cascade of models considering the rainfall inputs, HPC capabilities offer a significant increase of computation speed, which is very important in operational practice, especially during the critical events.

Quality verification of the rainfall-runoff models is usually based on the comparison of the discharge volume forecast model and the actual discharge volume rates measured in these profiles. According to Refsgaard (1997), quality verification is the process of demonstrating that a given site-specific model is capable of making "sufficiently accurate" simulations. Model quality verification therefore involves running a model using measured input parameters or parameters determined during the calibration process and comparing it with real measured values. To verify the model both statistical and graphical techniques are used.

Graphical techniques provide a visual comparison of simulated and measured constituent data and a first overview of model performance (ASCE, 1993). Basic option for visual quality assessment of hydrological models is the hydrograph.

FLOREON+ web application offers visualization of hydrographs to users and then allows visual comparison of the real discharge volume with simulated hydrographs (see Figure 2). Additionally precipitations in analysed time period and flood levels are presented on the hydrograph. Displayed hydrographs then allow visual comparison of the real discharge volume with models that were created in the selected time. A plot of identity, which represents another possibility of visual quality assessment model, is also presented with the hydrograph. Plot of identity is a scatterplot of the simulated and measured data along with the identity line $y = x$. If the simulated and measured data are in basic agreement then the points in the scatterplot will line up closely to the identity line. Points lying below the line of identity indicate that the model underestimates the reality. Similarly, points lying above the line of identity indicate that the model overstates the reality (see Figure 3).

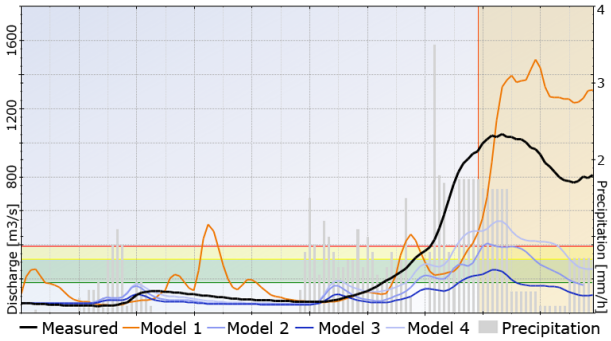


Figure 2: Hydrograph Example

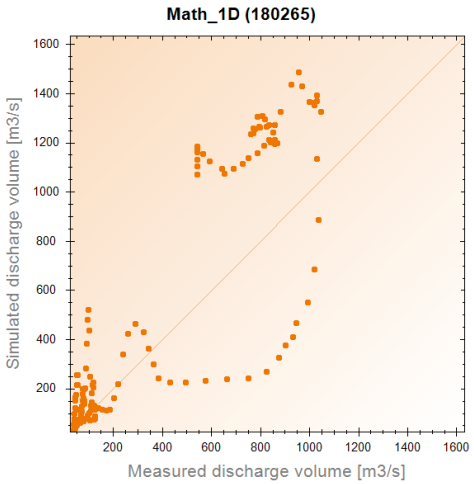


Figure 3: Plot of Identity Example

Table 1: Error Indices Example

ME [m3/s]	RMSE [m3/s]	MAE [m3/s]	MPE [%]	MAPE [%]	W [%]	MF [%]
111.8	274.0	185.6	47.8	65.0	33.6	41.4

Other methods used for model verification are the statistical methods that are based on the analysis of statistical indicators. Several error indicators are commonly used for quality model verification (see Table 1). The mean estimate error ME determines whether the model overestimates or underestimates. It is defined as

$$ME = \frac{1}{n}\sum_{i=1}^{n}\left(Q_i^{sim} - Q_i^{obs}\right), \quad (1)$$

where Q_i^{sim} is the *i*-th simulated value for the evaluated constituent, Q_i^{obs} is the *i*-th measured observation for the evaluated constituent, and *n* is the total number of observations. To assess the total error the following indicators can also be used:

- root mean square error

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Q_i^{sim} - Q_i^{obs}\right)^2}; \quad (2)$$

- mean absolute error

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Q_i^{sim} - Q_i^{obs}\right|; \quad (3)$$

- mean percentage error

$$MPE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Q_i^{sim} - Q_i^{obs}}{Q_i^{obs}}\right); \quad (4)$$

- mean absolute percentage error

$$MPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|Q_i^{sim} - Q_i^{obs}\right|}{Q_i^{obs}}; \quad (5)$$

- relative error in volume

$$VE = \frac{\sum_{i=1}^{n}\left(Q_i^{sim} - Q_i^{obs}\right)}{\sum_{i=1}^{n} Q_i^{obs}}; \quad (6)$$

- relative error for peak

$$MF = \frac{\max Q_i^{sim} - \max Q_i^{obs}}{\max Q_i^{obs}}. \quad (7)$$

These indicators are valuable because they indicate error in the units (*ME, MAE, RMSE*) or as a percentage (*MPE, MAPE, VE, MF*) of the constituent of interest, which aids in analysis of the results. Zero values of these indicators mean perfect fit. Singh et al. (2004) recommend to evaluate the most commonly used indicators, RMSE and MAE, as small enough, if they do not exceed half the standard deviation of the observations. It should be noted that the relative error in volume (VE) is defined in a similar manner as the percent bias (Gupta et al., 1999), steamflow percent volume error (Singh et al., 2004) and prediction error (Fernandez et al., 2005).

The current version of the FLOREON+ system represents the model output without uncertainty, see Figure 2. However, in order to represent the uncertainty of selected parameters in model scenarios, development of a software adapter for the FLOREON+ system has been started. The first results of Monte-Carlo simulations are evaluated by statistical methods and presented in the following subsections. This adapter will be able to take into account the uncertainty of selected model parameters within various prediction scenarios.

The Monte-Carlo method will also be used for the model calibration reduction, because only significant parameters will be identified by the simulation process.

4. UNCERTAINTY MODELLING USING THE MONTE CARLO SIMULATION METHOD

The Monte Carlo simulation method enables modelling of the probabilistic character of input uncertain parameters. A probabilistic distribution is used for modelling the stochastic character of the model inputs (Chudoba, et al., 2010).

By repeated realizations of the model over a random sample of input random parameters, statistic characteristics of the random output can be estimated (Atwood C. L., 1994). From the series of Monte Carlo simulation steps, it is also possible to establish an estimation of the hypothetic distribution function (Chudoba, et al., 2010).

4.1. Stochastic simulation of precipitation: a simulation approach

The stochastic modelling of the precipitation uncertainty is based on a perturbation of a precipitation matrix by the Monte Carlo simulation. The precipitation matrix *A*, which represents the precipitation forecast, has been randomly perturbed in the *k*-th Monte Carlo simulation by a constant random factor ξ_k. The aim of this uncertainty modelling is to simulate up to a 10% change of the precipitation according to the formula:

$$Apert(i,j) = A(i,j) + A(i,j)\xi_k \quad (8)$$

where the perturbation parameter ξ_k has been modelled by the uniform probabilistic distribution between [-0.1; 0.1]. These values represent the simulated 10 % uncertainty of the assumed precipitation forecast. The precipitation matrix *A* has thus been changed from the user-defined time-index, which represents time where these uncertainties are considered in the FLOREON+ model.

4.2. Scenario 1: Deterministic Cn curve parameter with a small value

In this scenario, the Cn curve parameter has been elected in by a small deterministic value (Cn = 55).
Based on the simulation procedure described in Section 4.1, a set of *k*=100 Monte Carlo simulation steps has been generated. The results of the Monte Carlo simulation model are presented in Figure 4, in which the horizontal axis represents time and the vertical axis represents the discharge volume of the simulated river at the selected station in m^3 / s. For the sake of clarity, Figure 4 also presents a 10% and 90% percentiles of the simulated discharge volumes.

In Figure 4 a small variability of the observed flow model can be seen that indicates a low sensitivity of the model to the random change of the precipitation matrix when the Cn curve parameter has a small deterministic value.

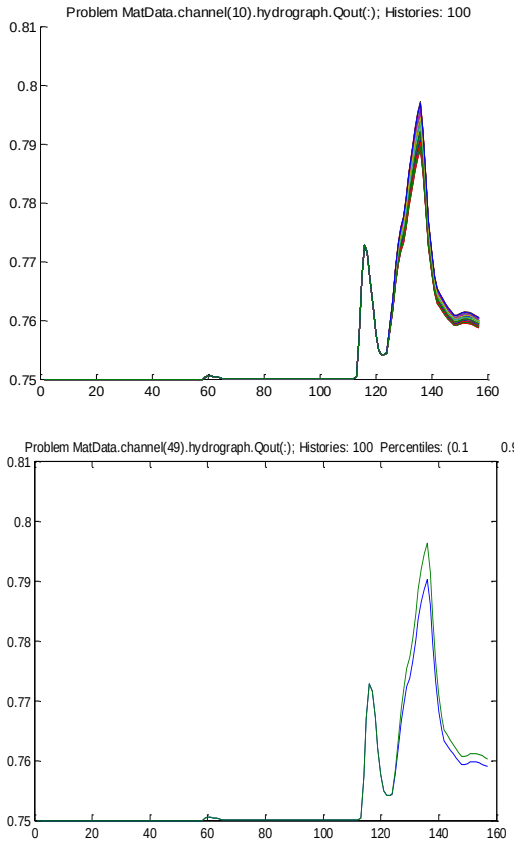


Figure 4: Discharge Volume Predictions under Scenario 1 (up); 10% and 90% Percentiles (down).

4.3. Scenario 2: Deterministic Cn curve parameter with a large value

Scenario 2 represents a similar variant to Scenario 1, except that the Cn curve parameter has a large value: Cn=95. The predicted discharge volume in the river again shows a little variability of the model, which indicates a low sensitivity of the model to the random change of precipitation. In contrary to Scenario 1, the predicted discharge volume is almost an order of magnitude larger (see Figure 4). Similarly to Scenario 1, in Figure 5 a small variability of the observed flow model can be seen that indicates a low sensitivity of the model to the random change of the precipitation matrix when the Cn curve parameter has a large deterministic value.

4.4. Scenario 3: Stochastic Cn curve parameter

Scenario 3 represents a case, when the precipitation matrix is also perturbed according to the stochastic model, as described in Section 4.1. In contrary to previous scenarios, the Cn curve parameter has been assumed to be random in the interval [CnLow, CnUpp], where CnLow = 55 and CnUpp = 95. The randomness of the Cn curve parameter has also been modelled by the uniform probabilistic distribution.

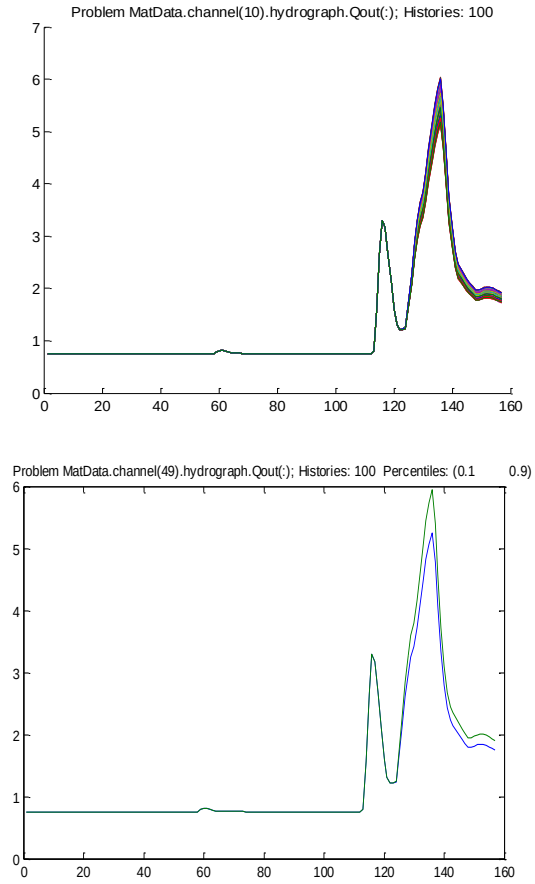


Figure 5: Discharge Volume Predictions under Scenario 2 (up); 10% and 90% Percentiles (down).

The Monte Carlo simulation model assumes that there is no dependency between the both perturbed parameters. Figure 6 shows results of the Monte Carlo simulation steps. There is a large uncertainty of the discharge volume predictions corresponding to time 120 – 140, see the y-axis, where the discharge volume oscillates between values 1-6 m^3/s, approximately. Thus, it is obvious that the randomness of input parameters of the prediction model cannot be simply replaced by the deterministic values in this scenario. Here presented Monte Carlo simulation analyses therefore appear to be meaningful for understanding of major uncertainties in the prediction model.

5. INTEGRATION TO THE FLOREON+ SYSTEM

The experiments in the previous section proved that the uncertainty of the input parameters has a great impact on the precision of predicted simulations. Therefore it is feasible to integrate these models to the FLOREON+ system to enhance the information provided to the decision-making process. The HPC environment described in section 2 is ideal for running the Monte Carlo method for simulating the uncertainty of input parameters because it consists of many similar and independent simulations that can be executed concurrently.

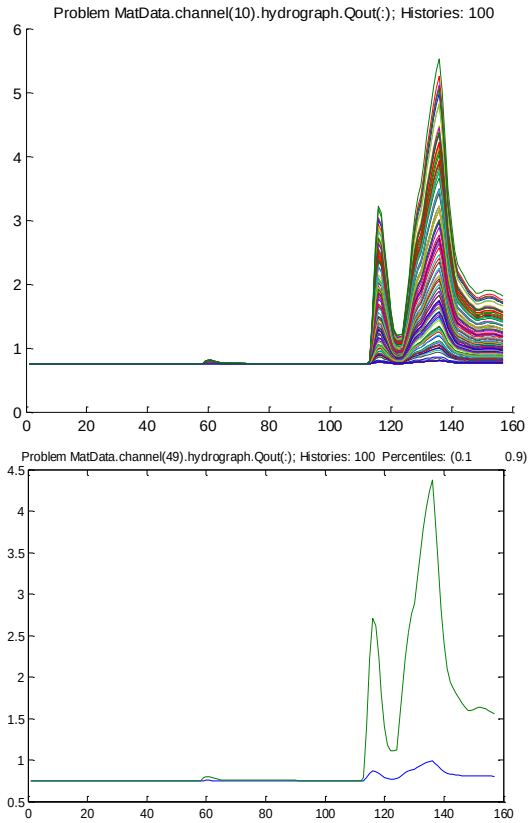


Figure 6: Discharge Volume Predictions under Scenario 2 (up); 10% and 90% Percentiles (down).

First the uncertainty values of chosen parameters had to be defined to launch these simulations automatically on the HPC cluster. Both the Cn and precipitation parameters were considered important and their possible variances set to specific probability distributions. The Cn curve parameter follows the normal distribution with small variance because this parameter is already pre-calibrated. The mean value of the normal distribution is set to the pre-calibrated value and the standard deviation is set to:

$$3\sigma = 0.05\ Cn \tag{9}$$

Precipitation uncertainty follows the uniform distribution where ξ_k is defined by the interval [-0.1; 0.1] but only precipitation forecast values are perturbed in this way. Only the prediction part of the simulations is therefore affected by the precipitation uncertainty.

After all the Monte Carlo simulations are finished on the HPC cluster, their results are collected and only their significant values are stored in the database. These significant values were defined as 5%, 15%, 25%, 75%, 85% and 95% percentiles of the simulated discharge volume for each time step to create three confidence intervals – 90% (between 5% and 95% percentiles), 70% (between 15% and 85% percentiles) and 50% (between 25% and 75% percentiles). These confidence intervals are then provided by the FLOREON+ system using the hydrographs that are shown in Figure 7.

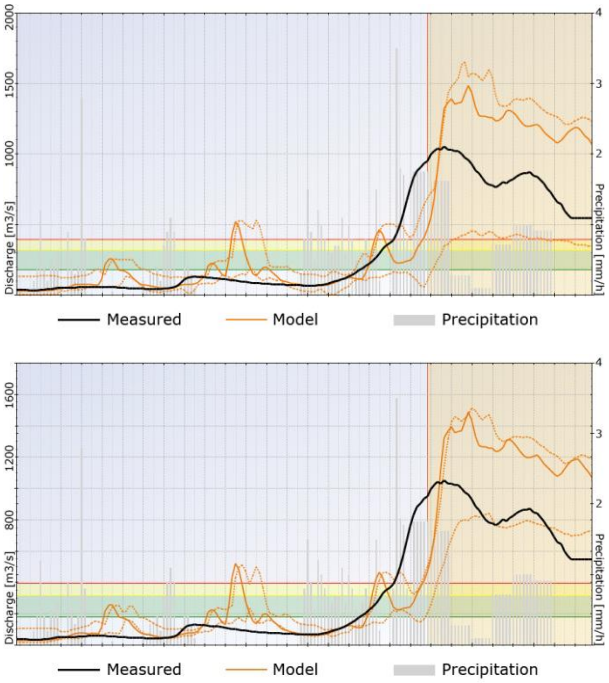


Figure 7: Discharge Volume Predictions with 90% (up) and 70% (down) Confidence Intervals.

According to Baillie & Bollerslev (1992), McNees & Fine (1996) and Christoffersen (1998), the standard evaluation of prediction interval proceeds by simply comparing the nominal coverage probability to the empirical (conditional) coverage probability (see Figure 7). The empirical coverage probability was calculated as the ratio between the number of observations that fall in the calculated prediction intervals and the total number of observations in analysed time period. Then, if the ratio overcomes the percentage of confidence interval, the interval is evaluated as successful. Numerical evaluation of selected episode with three confidence intervals can be found in Table 2.

Table 2: Evaluation of confidence intervals

Confidence interval	Values inside	Values outside	Interval Success	Is Successful
90%	45	3	93.75%	YES
70%	26	22	54.17%	NO
50%	4	44	8.33%	NO

6. CONCLUSION

This paper described the way to verify the quality of the model and the influence of the uncertainty in the input parameters to the discharge volume prediction models. The hydrographs were used for the visual quality assessment of hydrological models. A plot of identity and several error indices assessing the degree of consensus model with reality were also specified and Monte Carlo method was used to model the uncertainty of selected input parameters (precipitation and Cn curve parameters). These models were then integrated to the

FLOREON+ system and an example of their statistical and visual verification was presented.

The FLOREON+ system focuses on obtaining and analysing relevant data in real time. Prediction algorithms are then applied on the data to supply the information to the support decision-making processes in crisis management. These decisions can be supported by the predicted discharge volume on measuring stations or prediction and visualization of the flood lakes on the landscape.

Information about the quality of model predictions and uncertainties are provided in understandable form to anyone who needs to find out about the actual flood situation, whether it is an ordinary citizen, the mayor of the municipality responsible for crisis management, or expert in the field. This information can help them to understand and evaluate the situation and react to the situation appropriately.

ACKNOWLEDGMENT

This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070).

REFERENCES

ASCE Task Committee on Definition of Criteria for Evaluation of Watershed Models of the Watershed Management Committee, Irrigation and Drainage Division (1993). "Criteria for Evaluation of Watershed Models." *J. Irrig. Drain Eng.*, 119(3), 429–442.

Atwood C. L. (1994). "Hits per Trial: Basic Analysis of Binomial Data." Department of Energy; DOI: 10.2172/10191270

Baillie, Richard T. & Bollerslev, Tim, (1992). "Prediction in dynamic models with time-dependent conditional variances", *Journal of Econometrics*, Elsevier, vol. 52(1-2), 91-113.

Christoffersen, P. F. (1998). "Evaluating interval forecasts", *International Economic Review* 39(4), p. 841–862.

Chudoba J., Praks P., Labeau P. E.: Modelling of transport of radioactive substances from underground storage for risk assessment - a case study. Reliability, Risk and Safety – Ale, Papazoglou & Zio (eds). © 2010 Taylor & Francis Group, London, ISBN 978-0-415-60427-7. Pg. 1019-1025.

Fernandez, G.P., G.M. Chescheir, D.M. Amatya and R.W. Skaggs. 2005. Development and Testing of Watershed Scale Models for Poorly Drained Soils. Trans of ASAE. 48(2): 639-652.

Fishman G. Monte Carlo: Concepts, Algorithms, and Applications. Springer Series in Operations Research and Financial Engineering. 1st ed. 1996. 698 p.

Gupta, H., Sorooshian, S., and Yapo, P. (1999). "Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration." *J. Hydrol. Eng.*, 4(2), 135–143.

Gwo J.P., Yeh G.T., 2004. High-performance simulation of surface-subsurface coupled flow a reactive transport at watershed scale. *1st International Conference on Computational Methods* (ICCM04). Singapore, Singapore.

Martinovič J., Štolfa S., Kožusznik J., Unucka J., Vondrák I., 2008. FLOREON – the system for an emergent flood prediction. *ECEC-FUBUTEC-EUROMEDIA*. Porto, Portugal.

McNees, S.K. and Fine, L.K., (1996), "Forecast Uncertainty: Can it be Measured?", Presented at the Conference on Expectations in Economics, Federal Reserve Bank of Philadelphia, October, 1996.

Refsgaard JC (1997) Parametrisation, calibration and validation of distributed hydrological models. Journal of Hydrology, 198, 69-97.

Singh, J., H. V. Knapp, and M. Demissie. (2004). Hydrologic modeling of the Iroquois River watershed using HSPF and SWAT. ISWS CR 2004-08. Champaign, Ill.: Illinois State Water Survey. Available at: www.sws.uiuc.edu/pubdoc/CR

AUTONOMOUS LOGISTIC PROCESSES OF BIKE COURIER SERVICES USING MULTIAGENT-BASED SIMULATION

Max Gath[(a)], Thomas Wagner[(b)], Otthein Herzog[(c)]

Center for Computing and Communication Technologies (TZI)
University of Bremen, Am Fallturm 1, 28359 Bremen

[(a)]mgath@tzi.de, [(b)]twagner@tzi.de [(c)]herzog@tzi.de

ABSTRACT

We present an agent-based approach for solving *pickup and delivery problems* (PDP) in dynamic environments and show its application by bike couriers in urban districts. To avoid social conflicts, the focus is on the computation of fair allocations of orders to bike couriers who are paid by commission fees. We realize autonomous logistic processes and present a mulitagent system with specially adapted negotiation mechanisms. In addition, we apply the mulitagent-based simulation platform PlaSMA for modeling, simulating, and evaluating different transport scenarios within the transport infrastructure of the city of Bremen, Germany. The customer orders are based on real data provided by a German bike courier company. The results show that the approach indeed improves the fairness of distributions significantly. Finally, we present a new business model for a commercial online dispatching platform enabled by our approach.

Keywords: Multiagent-based Simulation, Scheduling, Pickup and Delivery Problem, Courier and Express Services

1. INTRODUCTION

Transporting goods in urban districts by bike couriers provides many advantages. It contributes to a reduction of traffic and CO_2 emissions, increases the service quality through short transit times, and satisfies individual demands of customers. However, determining optimal routes to satisfy transport requests in real-world operations involves a set of practical complications. For instance, to avoid social conflicts and increase the service quality of bike couriers there is an essential demand for a "fair" distribution of orders to couriers who are paid by commission fees. In fact, in high-frequency districts many customers can be served by a commonly favored tour. Nevertheless, unpopular tours have to be done as well.

This paper addresses the so-called *Pickup and Delivery Problem* (PDP) with special requirements that are introduced in Section 2. In Section 3, we provide a short introduction to autonomous logistic processes and refer to related work. In Section 4, we present a novel agent-based approach for solving PDPs in dynamic real-world environments and show its application to bike couriers in urban districts. The participating objects and actors in our logistic scenarios are represented by intelligent software agents that act autonomously. Through specially adapted interaction protocols, agents exchange tasks with each other and improve solutions continuously. We show that the mechanism provides "fair" distributions without reducing the logistic service quality.

For the purpose of evaluation, we use the multiagent-based simulation platform PlaSMA (Warden et al. 2010) for modeling and simulating different transport scenarios within the transport infrastructure of Bremen, Germany. PlaSMA is described in Section 5. In Section 6, we specify the experimental setup and compare the formally specified fairness of the routes to the results of a robust distance minimizing algorithm. The customer orders are based on real data provided by a German bike courier company. The results show that the approach improves the fairness of distributions significantly. Finally, we argue that the developed approach enables innovative business models for commercial online dispatching platforms that can be used by bike couriers and retailers to offer reliable deliveries as well as sustainable transportation.

2. THE DYNAMIC PICKUP AND DELIVERY PROBLEM OF BIKE COURIERS

The well-known *pickup and delivery problem* (PDP) (Berbeglia, Cordeau, and Laporte 2010; Parragh, Doerner, and Hartl 2008) is concerned with determining a set of minimum cost routes for a fleet of vehicles to satisfy customer requests for transporting objects from an origin to a destination. Certain scenarios involve additional constraints and customer demands that have to be considered, e.g., time windows (Dumas, Desrosiers, and Soumis 1991).

The investigated problem is a variant of the PDP. In logistic processes of courier, express and parcel services (CEP) most requests are received on their service day and have to be operated upon within minutes. The dynamics and the complexity increase significantly with the consecutive appearance of new requests as well as the incidence of unexpected events like changing traffic conditions.

Not only the needs of the customer and of the logistic company have to be considered, but also the demands and varying capabilities of the bike couriers themselves, who are usually paid by commission fees and have an individual physical fitness. Additionally, the pickup and delivery stops are located in various districts and thereby valued differently. In inner city sectors more stops can be accomplished by short tours, whereas in outskirts only one delivery can be served by driving a long distance. Nevertheless, all incoming requests have to be serviced. Thus, there is the essential demand for the generation of "fair" distributions without privileging or overstraining any courier with respect to his fitness.

3. AUTONOMOUS LOGISTIC PROCESSES

In the last decades, numerous efficient heuristics and meta-heuristics have been developed for the transportation domain like ant systems, tabu-search, simulated annealing and genetic algorithms, just to name a few e.g., by Bräysy and Gendreau (2005a, 2005b) and Parragh et al. (2008). However, central planning and controlling in dynamic and complex logistic processes is increasingly difficult due to the requirements of flexibility and adaptability of changing environmental influences (Scholz-Reiter et al. 2004). Often one faces scenarios where it is not possible to acquire all relevant information for the decision making process centrally.

In autonomous logistic processes the decision making is shifted from central, hierarchical planning and controlling systems to decentralized, heterarchical systems (Scholz-Reiter et al. 2004). In agent-based processes autonomously acting software agents represent logistic objects, such as shipments, trucks, and containers (Fischer, Müller, and Pischel 1995; Schuldt 2011). They have the ability to interact with other agents by means of negotiation and communication mechanisms. The general problem is split into smaller problems that agents solve locally concurrent within short time windows to optimize the behavior of the overall system. In cooperating systems the agent's goal is to pursue a globally optimized behavior and achieve common goals whereas in competitive systems each agent acts selfish to reach its own objectives.

The advantages of applying multiagent systems are high flexibility, adaptability, scalability, and robustness of decentralized systems through problem decomposition and proactive, reactive and adaptive behavior of intelligent agents (Brooks 1986; Rao and Georgeff 1995). The potential of autonomous logistic processes is even bigger in open, unpredictable, dynamic and complex environments (Böse and Windt 2007; Kuske, Luderer, and Tönnies 2010; Schuldt 2011; Wessels, Jedermann, and Lang 2010; Windt and Hülsmann 2007). A comprehensive survey and the state of the art in research on autonomous logistic processes are provided by Hülsmann, Scholz-Reiter, and Windt (2011), by Hülsmann and Windt (2007) and by Schuldt (2011).

Examples of multiagent systems for resource allocation, scheduling, optimization, and controlling within industrial applications are provided by Himoff, Rzevski, and Skobelev (2006), by Neagu, Dorer, Greenwood, and Calisti (2006) and by Skobelev (2011).

4. EXTENDING INTERACTION PROTOCOLS FOR EVENHANDED DISPATCHING

Similar to other autonomous logistic processes (Schuldt 2011), autonomously acting, intelligent agents represent bike couriers and orders. Orders must be classified as popular or unpopular. However, the subjective value of an order depends on the courier (e.g., on the position, the currently accepted requests, and the individual fitness). Nevertheless, the probability to combine orders successfully in one tour is increasing in high-frequency areas. Therefore, the trading area is clustered in high-frequency and low-frequency areas using historical data. We apply a density based clustering (Ester, Kriegel, Sander, and Xu 1996) to determine interesting clusters of different sizes, diameters and shapes. The pickup and delivery locations of a shipment define whether it belongs to commonly popular orders or to unpopular orders.

Whenever a new request has to be acted upon, an agent is created that represents the given shipment. The goal of the agent is to find a proper transport service provider for carrying the shipment from its origin to the destination with respect to the given time windows.

The agent starts negotiating with other agents representing bike couriers. These agents are utility based agents that act selfish to optimize their individual utility function. Beside the monetary costs the agent considers its fitness value for computing the utility value of orders.

Definition 1 (The Orders Revenue). *Let* A *denote the set of agents,* Ω *the set of orders,* $\omega_{\text{Value}} \in \mathbb{R}^+$ *the monetary value of order* $\omega \in \Omega$ *and* $\alpha_{cost}: 2^{\Omega} \rightarrow \mathbb{R}^+$ *a mapping that determines the costs of agent* $\alpha \in A$ *of picking up and delivering orders* $S \in 2^{\Omega}$ *by solving a TSPTW. The mapping* $r: A \times 2^{\Omega} \rightarrow \mathbb{R}$ *computes the expected monetary revenue of orders S for Agent α by*

$$\mathrm{r}(\alpha, \mathrm{S}) = \sum_{\omega \in S}(\omega_{\text{Value}}) - \alpha_{\text{cost}}(\mathrm{S}) . \quad (1)$$

Definition 2 (The Fitness of a Bike Courier). *Let* $\alpha_{fitness} \in \mathbb{R}^+$ *denote the constant physical fitness value of the courier represented by agent* α *and* $dist: 2^{\Omega} \rightarrow \mathbb{R}^+$ *a mapping determineing the total distance that a courier has to drive for shipping orders S. The mapping* $f: A \times 2^{\Omega} \rightarrow \mathbb{R}^+$ *yields the fitness of the courier after delivering orders S. It is defined as*

$$f\big(\alpha, dist(\mathrm{S})\big) = 1 + \frac{a_{fitness}}{dist + a_{fitness}} \quad (2)$$

Definition 3 (The Bike Courier's Utility of Orders). *The mapping* $u: A \times 2^{\Omega} \rightarrow \mathbb{R}$ *determines the utility of orders S for agent* α*. The mapping u is defined as*

$$u(\alpha, S) = r(\alpha, S) \cdot f(\alpha, S) \quad (3)$$

where $r: A \times 2^{\Omega} \rightarrow \mathbb{R}$ *and* $f: A \times 2^{\Omega} \rightarrow \mathbb{R}^{+}$ *are defined in Definition 1 and Definition 2, respectively.*

In transport logistics the costs of an order are commonly based on the additional distance that has to be driven by the courier for picking up and delivering the auctioned shipment. To compute the distance, the agent has to solve the Traveling Salesman Problem with Time Windows (TSPTW) (Dumas, Desrosiers, Gelinas, and Solomon 1995). We adapted the well established Solomon I1 insertion heuristic (Solomon 1987) to solve the TSPTW while supplementary considering the essential sequence constraints. The algorithm is not optimal, but is often applied because it obtains quite good solutions in a short time. Figure 1 shows the main procedure.

Input: The pickup stop, the delivery stop, the current path
Output: A path containing the pickup and the delivery stop, if no path exists it returns null

Procedure *InsertionHeuristic*(pickup,delivery,path)
begin
1: path ← *Insert*(pickup, path)
2: path ← *Insert*(delivery, path)
3: **if** (*isEmpty*(path)) **then**
4: path ← *Insert*(delivery, path)
5: path ← *Insert*(pickup, path)
6: **end if**
7: **return** path
end

Figure 1: The Procedure of the Insertion Heuristic.

Figure 2 shows the insertion procedure. The new stop is inserted between all other stops in the current path successively. The subroutine *append*(x,y,z) appends stop x in path y before stop z and returns a new list. Within the subroutine *checkTimeAndSeqConstraints*() we check for all orders, if the delivery stop is behind the pickup stop as well as all time window constraints. The *cost*() function returns the distance of the path by summing up the weights of all edges.

Input: The stop to insert, the current path
Output: A path containing the new stop, if no path exists it returns an empty path

Procedure *Insert*(stop, path)
begin
1: newPath ← *empty*
2: **for all** (stop) s ∈ path **do**
3: tempPath ← *append*(stop,path,s)
4: **if**(*checkTimeAndSeqConstraints*(tempPath) ∧ (*isEmpty*(newPath) ∨ *cost*(newPath) > *cost* (tempPath))) **then**
5: newPath ← tempPath
6: **end if**
7: **end for**
8: **return** newPath
end

Figure 2: The Insert Procedure.

Figure 3 shows the improvement procedure. Each stop of the current path is removed with the subroutine *remove*() and reinserted using the insert procedure. As a result, we have an anytime algorithm that finds better solutions the more time it keeps running. It returns a valid solution if it is interrupted. If no further improvement is possible, the current path is returned. In the worst case, the improvement algorithm creates all permutations of the stops (O(n!)). In order to reduce the worst case complexity, we integrated the *abort*() subroutine to stop the procedure after a defined number of cycles and to return the best valid solution.

Input: The current path
Output: An optimized path

Procedure *ImprovementHeuristic*(path)
begin
1: tempPath ← *empty*
2: **while** *cost*(path) > *cost*(tempPath) || *abort*() **do**
3: **if** (*isEmpty*(tempPath)) **then**
4: tempPath ← path
5: **end if**
6: path ← tempPath
7: **for all** (stop) s ∈ tempPath **do**
8: *remove*(s,tempPath)
9: *insert*(s, tempPath)
10: **end for**
11: **end while**
12: **return** path
end

Figure 3: The Improvement Procedure.

To ensure that autonomous selfish agents generate fair distributions, we developed a stable interaction protocol.

Definition 4 (The Fairness of a distribution). *The surjective mapping* $d: \Omega \rightarrow A \cup \{\lambda\}$ *maps an order* $\omega \in \Omega$ *to an agent* $\alpha \in A$ *or to the symbol* λ*, if* ω *is not allocated to any agent. It creates a partition of all elements* $\omega \in \Omega$*. Let dom(d) denote the domain of d. The inverse mapping*

$$d^{-1}(d, \alpha) = \{\omega \mid \omega \in dom(d) \wedge d(\omega) = \alpha\} \quad (4)$$

defines all orders $\omega \in dom(d)$ *of an agent* α*. The standard deviation* σ *of each participant's individual utility*

$$\sigma_{\alpha \in A} u(d^{-1}(d, \alpha), \alpha) \quad (5)$$

determines the fairness of a distribution. The fairest distribution minimizes the standard deviation σ with

$$\min_d \ \sigma_{\alpha \in A} u(d^{-1}(d, \alpha), \alpha). \quad (6)$$

For sure, it is not a valid solution if all orders ω are allocated to λ, although Equation 5 and Equation 6 are zero.

We extended the Vickrey auction (Vickrey 1961) by introducing a currency for bidding on shipments. Every agent α gets a fixed start-up budget β_α^0 that he may use for bidding at time 0. The function $u^t(\alpha, \omega)$ defines the utility value of order ω by agent α at time t similarly to Definition 3 with respect to already accepted orders. The bid is computed by the following equation:

$$b(\alpha, \omega, t) = \begin{cases} u^t(\alpha, \omega), & \text{if } u^t(\alpha, \omega) \leq \beta_\alpha^t \\ \beta_\alpha^t, & \text{else} \end{cases}. \quad (7)$$

Consequently, the agent can at most bid the maximum amount of its current balance.

The winning agent α_{winner} is determined by the shipment agent with

$$\alpha_{winner} = \arg \max_{\alpha \in A} \beta(\alpha, \omega, t). \quad (8)$$

While using the Vickrey auction the winning agent has to pay the second highest bid:

$$p = \max_{\alpha \in A \setminus \{\alpha_{winner}\}} \beta(\alpha, \omega, t). \quad (9)$$

Bidding the true valuation is the dominant strategy in private value Vickrey auctions for every participant (Shoham and Leyton-Brown 2009, S.319). Thus, the negotiation process is stable and no agent has any incentive to reveal false valuations and manipulate the negotiation process. Consequently, it is not possible for an agent to manipulate the outcome of a negotiation for its own purposes. Additionally, it must be guaranteed that agents can also increase their individual credit balance. To subsidize the transport of unpopular service requests, they get a higher amount of money for handling unpopular orders than for popular ones:

$$\text{reward}(\omega) = \begin{cases} c_1, & \text{if } \omega \text{ is popular} \\ c_2, & \text{else} \end{cases} \text{ with } c_1 > c_2 \quad (10)$$

The reward motivates the couriers to accept also unpopular orders to increase their credit balance. In addition, it prevents that one agent purchase only popular orders by auction if the budget reduces to zero.

Theorem 1 (Upper bounds of c_1 and c_2). The highest utility value of an order is the upper bound for c_1 and c_2.

Proof. "A solution x is Pareto efficient – i.e., Pareto optimal – if there is no other solution x' such that at least one agent is better off in x' than in x and no agent is worse off in x' than in x." (Sandholm 1999, S.202). An advantage of using the Vickrey auction is that every order is auctioned "Pareto efficient to the bidder that values it the most" (Sandholm 1999, S.213). If c_1 and c_2 are higher than the maximum utility value of an order, Equation 7 will not intervene the result. The mechanism is transformed into the standard Vickrey auction. Thus, the agents positioned in high-frequency areas will get the most popular orders because their additional costs are lower than the cost of other agents. Consequently, a Pareto efficient and cost minimized solution is constituted, which maximizes the social welfare (the total sum of all agent's utility). □

Theorem 2 (Lower bounds of c_1 and c_2). The lowest utility value of an order is the lower bound for c_1 and c_2. Otherwise, the negotiation maximizes neither the social welfare (the total sum of all agent's utility) nor it will lead to fair allocations.

Proof. The order is not sold to the bidder, who values it at most, if Equation 7 reduces the bid because of an insufficient budget. Consequently, the social welfare is not optimized. If c_1 and c_2 are always lower than the minimum computed utility value, the budget of all agents reduces monotonously and converges to zero. As a result, the allocation will not lead to fair allocations, but the orders are sold successively without considering the potential for allocating efficient solutions. □

It addition, the initial budget β_α^0 has to be sufficient low that it impacts the allocation. To support waiting couriers who are working but not processing any order, they are rewarded by receiving a small amount *w* for each waiting unit.

However, the theorems hold only for the auctioned item in the actual situation. If another shipment is auctioned afterwards, the situation can change significantly and earlier profitably valued orders can be rendered worse. In addition, the complexity is aggregated by the high degree of dynamics that result from unexpected events like changing traffic conditions and delays. To adapt tours and timetables in respect to the current situation, we apply the Vickrey auction iteratively to exchange tasks between the agents. In this case the auctioneer is also a service provider and can optionally refuse a deal. To avoid this problem, the auctioneer only accepts a transaction if he receives the profit that he would gain by serving the request by itself from the winner of the auction.

5. EVALUATION USING MULITAGENT-BASED SIMULATION

Multiagent-based simulation (MABS) combines concepts of mulitagent systems and qualitative simulation. Applying MABS for evaluating multiagent systems before their deployment in real applications is an accurate cost and time reducing method (Schuldt, Gehrke, and Werner 2008). This holds especially for

scenarios with run-time agent interactions that cannot be predicted in advance (Jennings 2001).

PlaSMA (Warden et al. 2010) is a simulation middleware that extends the well-known JADE framework (Bellifemine, Caire, and Greenwood 2007) which implements the standards for agent interaction and communication defined by the *IEEE Foundation for Intelligent Physical Agents* (FIPA). PlaSMA provides a discrete time simulation that ensures a correct conservative synchronization with time model adequacy, causality and reproducibility (Schuldt et al. 2008). The GUI provides information about time progression and the current simulated processes, e.g., how many active agents are registered at the platform as well as the positions of the physical objects.

The physical world within the simulation environment is modeled as directed graph. Nodes represent traffic junctions or logistic sources and sinks while edges represent different types of ways, e.g., roads, motorways, trails, and waterways. They have additional parameters that determine the maximum allowed velocity and the traffic density on an edge. If the edge is denoted with a high traffic density, the maximum possible velocity is reduced respectively. To simulate dynamics of the environment and the appearance of unexpected events (e.g., accidents and changing traffic conditions), the values can vary during simulation runs.

In order to model sound panning and controlling processes in the logistic domain, we extended PlaSMA to import transport infrastructures from OpenStreetMap (see: www.openstreetmap.org) (OSM) databases. This enables the integration of high detailed graphs with up to 300,000 edges and 150,000 nodes. After clipping a user defined map section and choosing relevant types of edges (e.g., roads, waterways, highways and railways) several preprocessing procedures are started to reduce the complexity of the overall graph without effects on the granularity of the infrastructure model. For example, redundant nodes as well as nodes that are only important to mark the course of the roads are deleted. The result is a directed graph which includes information about the real world speed limits, the distance as well as the type of an edge.

Particularly within large infrastructures determining the shortest path between nodes is an essential, costly, and time-consuming procedure within the decision making process of an agent (see Section 4). Consequently, we implemented Dijkstra's single-source shortest paths search (Dijkstra 1959) that is realized by a memory-efficient joint representation of graph and radix heap nodes (Ahuja, Mehlhorn, Orlin, and Tarjan 1990). Therefore, we can guarantee that the search is optimal and has linear time complexity.

To induce meaningful results, the simulation platform supports batch runs for the evaluation of a set of scenarios as well as multiple runs of a certain scenario with varying random seeds for the reproducible generation of random variables. In each run individual performance indicators are measured and saved in a database. Thus, they can be verified with an arbitrary spreadsheet program.

Further detailed information about the PlaSMA Simulation framework is provided, e.g., by Gehrke and Ober-Blöbaum (2007) and by Warden et al. (2010) as well as on the corresponding website http://plasma.informatik.uni-bremen.de.

6. RESULTS

In several preliminary scenarios we determined plausible values for an appropriate start-up budget and the amount of money an agent receives for transporting popular or unpopular orders. Thereby, we respected the upper and lower bounds deduced from Theorem 1 and Theorem 2. Afterwards, the evenhanded mechanism was evaluated by the comparison with a distance minimizing algorithm.

6.1. Experimental Setup

In this investigation we modeled the road network of the City of Bremen, Germany (see Figure 4). The whole modeled transport infrastructure of Bremen contains 2,103 Nodes and 4,448 Edges.

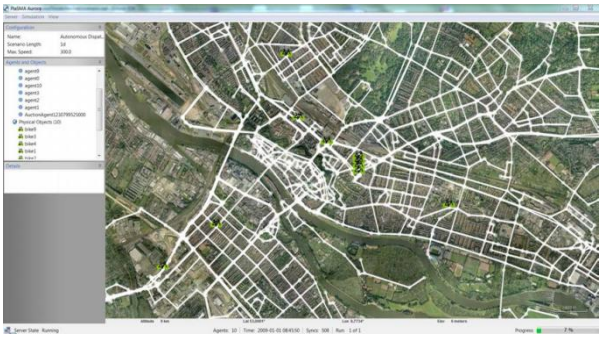


Figure 4: The Picture Shows a Part of the Transport Infrastructure.

The costs of handling a specific task are equal to the additional distance in kilometers, which must be driven by the courier. Accordingly, we assume that the revenue for delivering an order is twice the distance in kilometers between its origin and destination. This assumption reflects real pricing systems. The maximum possible velocity of the couriers is set to 19 kilometers per hour and the fitness value $a_{fitness}$ to 30 for all agents that represent bike couriers. With respect to Definition 2, it implicates the realistic assumption that the fitness of each courier is half as much after driving 30 kilometers.

To investigate the fairness of the mechanism, we compared the computed allocations with solutions that were computed by a distance minimizing algorithm. Therefore, we implemented a distance minimizing algorithm that is based on our approach. However, this requires totally cooperating agents that accept every order even if it is less profitable for themselves. A currency is not needed. As a result, the orders are sold by the use of the standard Vickrey auction to the bike courier, who minimizes the total costs, which are equal

to the driven distances. Thus, the first shipments will probably be served by one courier, because the additional costs are likely lower than the costs of the other couriers that are waiting at the depot. The negotiation protocol for changing orders during operations is adapted accordingly. In order to implement reactive behavior in both mechanisms, each agent starts a negotiation with another arbitrary bike agent consecutively in every minute.

The customer orders in the simulated scenarios are based on real customer orders of a representative week provided by a German bike courier company. The data include the origin, the destination, the time windows, and the dispatching time. We simulated a whole week with more than 1000 orders that have to be dispatched. The number of orders varied between 155 and 273 each day.

6.2. Parameter Optimization

We started simulating several preliminary scenarios with varying values for c_1, c_2, and the start-up budget β_a^0. We observed that the parameters c_1 and c_2 adjust the influence of the mechanism. If higher values are used, the influence of the mechanism is lower. As a result, the total distance of all couriers is decreasing, while it does not lead to fair allocations. Choosing low values we observed the same effect which is described in Theorem 2. The tours are neither fair nor do they have a short distance. The same effect could be observed by varying the start budget β_a^0, respectively. The best results within the infrastructure of Bremen were obtained with the values $c_1 = 6$, $c_1 = 5$, and $\beta_a^0 = 10$. With these values the mechanism computed fair allocations with an adequate total distance. The parameters scale with the distances an agent must drive. Consequently, they are dependent on the infrastructure of the catchment area.

6.3. Comparison of fairness

To evaluate the approach, we compared the generated distribution of both algorithms. We simulated different scenarios with five to ten couriers working concurrently.

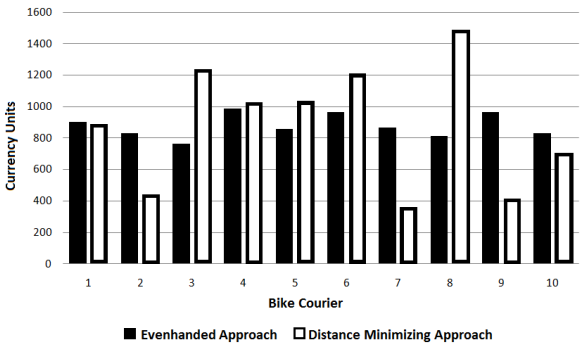


Figure 5: The Black Bars Show the Revenue of 10 Bike Couriers After One Week using the Evenhanded Dispatching Approach, while the White Bars Depicts the Revenues using the Distance Minimizing Algorithm.

Figure 5 shows the revenue of the bike couriers after a whole week. Respectively, Figure 6 depicts the covered distances. While the maximum pay gap in distance minimized distributions is 75.7%, the highest difference in evenhanded allocations is only 22.6%. Additionally, the maximum difference in covered kilometers varies up to 72.0% using the distance minimizing algorithm and up to 30.1% using the evenhanded approach. Indeed, after one week the pay gaps using distance minimizing algorithms are significantly higher than using the evenhanded approach.

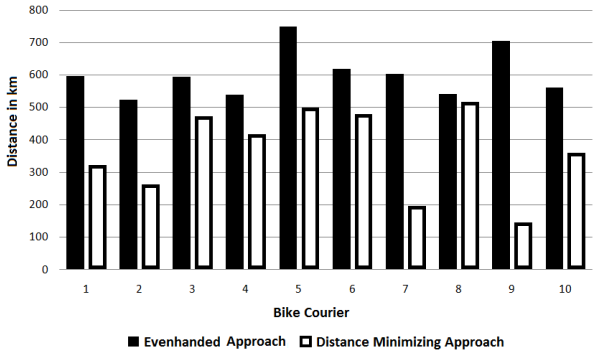


Figure 6: The Black Bars Show the Covered Distances of 10 Bike Couriers After One Week using the Evenhanded Dispatching Approach, while the White Bars Depict the Covered Distances using the Distance Minimizing Algorithm.

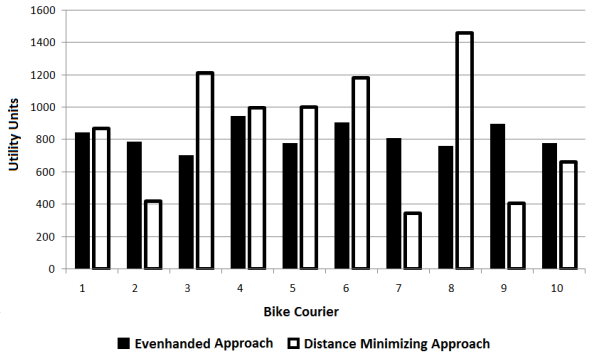


Figure 7: The Black Bars Show the Bike Courier's Utility of Orders of 10 Bike Couriers After One Week using the Evenhanded Dispatching Approach, while the White Bars Depicts the Utility Values using the Distance Minimizing Algorithm.

Figure 7 shows the bike courier's utility of all orders defined by Definition 3 of the whole week. As the fitness value affects the utility value, we computed the utility value for each day and summed up the results. The utility values are used to determine the fairness of the distribution with respect to Definition 4. While the average standard deviation σ of each participant's individual utility value is 1153.82 utility units using the

distance minimizing algorithm, it is only 226.69 units in the evenhanded allocation. The obtained results reveal deviations higher than 80.36% concerning the fairness of the allocations.

However, Figure 6 indicates that the total distance increases from 3672,78 km to 6038,24 km and every courier has to drive more kilometers than in the distance minimized distribution. Therefore, we investigated separate days individually. Figure 8 shows that the evenhanded approach is even significant fairer after one day in contrast to the distance minimizing algorithm. In addition, it shows that courier No. 8 has to drive nearly half of the distance of the entire week on one day.

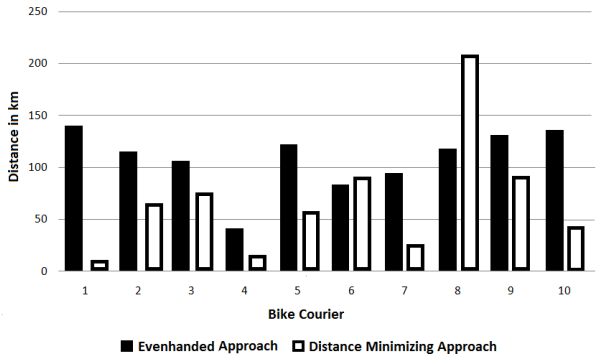


Figure 8: The Black Bars Show the Driven Distances of 10 Bike Couriers After One Day using the Evenhanded Dispatching Approach, while the White Bars Depict the Driven Distances using the Distance Minimizing Algorithm.

Other results from scenarios with less bike couriers (not shown in diagrams) pinpoint that the reduction of bike couriers lead to a decreasing number of orders operated within the guaranteed time windows.

6.4. Discussion

Several simulated scenarios based on real customer orders revealed that distance minimizing algorithms are not applicable for the dispatching of bike couriers who are paid by commission fees on a daily basis. The results show that the utility, the driven distances as well as the revenue of the couriers are unequally distributed and varying significantly using the distance minimizing algorithm. In contrast, our approach creates fair distributions if we compare the effort as well as the individual revenue of all couriers. In addition, the fairness as specified in Definition 4 indicates that the evenhanded approach computes fair distributions.

However, the driven distances of the whole week are higher for all couriers. Nevertheless, the results show that it is reasonable to cover larger distances in the entire week, if the total distance is equally distributed over the days. The courier does not prefer driving the most kilometers on one day while waiting for requests on all other days. He would pass over his physical limits on one day. Moreover, the intention of a courier is to carry shipments and not to wait at the central depot and earn even less money. The reduction of operating couriers is not reasonable because customer requests cannot be determined in advance and must be serviced within short time windows. In real scenarios, cost-intensive external carriers are instructed to deliver the orders in time, if not enough couriers are available.

Autonomous logistic processes based on mulitagent technology ensure a high flexibility and adaptivity in dynamic environments. In addition, it allows the modeling of a system with heterogeneous agents having individual capabilities. For instance, the maximum possible velocity and the fitness of represented couriers may differ.

In summary, the evenhanded approach enables a fair allocation of orders to bike couriers by a multiagent system with autonomous selfish acting agents with a reasonable increase of the driven distances.

7. CONCLUSION AND OUTLOOK

In this paper we presented a new approach to generate fair distributions for the dynamic *pickup and delivery problem* (PDP) and showed its application to bike couriers.

The mulitagent-based simulation platform PlaMSA served to evaluate the approach within the transport infrastructure of Bremen. We dispatched real customer requests provided by a German bike courier company.

The results reveal that applying distance minimizing algorithms is not pertinent, whereas our fairness approach motivates couriers and distributes the effort as well as the revenue equally to the bike couriers with respect to their fitness.

Automated evenhanded dispatching enables new business models for commercial internet dispatching platforms of bike couriers implementing autonomous logistic processes. Couriers can register themselves with the agent system by their mobile application for receiving orders. Afterwards, the mulitagent system computes evenhanded allocations of orders to bike couriers and suggests routes while taking into accout the dynamics of the environment. Therefore, the approach enables large scale online dispatching platforms which can be used by couriers and retailers to offer reliable deliveries as well as sustainable transportation.

With respect to new transport visions and the increasing usage of e-bikes for the transport of heavy goods the relevance of automated dispatching platforms for logistic transport service providers paid by commission fees is increasing.

To decrease the total effort and distribute the revenue equally, also profit sharing methods for freight carriers may be considered (e.g., Krajewska, Kopfer, Laporte, Ropke, and Zaccour 2007). However, profit sharing concepts changes the pricing system and couriers are not paid by commission fees anymore. Consequently, this effects the agent's strategies as well as their incentive to deliver shipments.

Further investigations will focus on the integration of traffic simulation within the PlaSMA simulation framework as well as on modeling and simulation of

unexpected events to evaluate the evenhanded mechanism in dynamic environments. Another point of interest is how a heterogeneous fleet affects the fairness of a distribution. We seek to carry over techniques well-proven in autonomous logistic processes such as team formation and cooperation of selfish acting agents.

In order to develop new potential for horizontal cooperation, extra change depots in inner city districts where couriers can exchange shipments are of special interest.

ACKNOWLEDGMENTS

The presented research was partially funded by the German Research Foundation (DFG) within the Collaborative Research Centre 637 "Autonomous Cooperating Logistic Processes: A Paradigm Shift and its Limitations" (SFB 637) at the University of Bremen, Germany.

REFERENCES

Ahuja, R. K., Mehlhorn, K., Orlin, J. B., & Tarjan, R. E. (1990). Faster Algorithms for the Shortest Path Problem. *Journal of the ACM*, *37*(2), 213-223.

Bellifemine, F., Caire, G., & Greenwood, D. (2007). *Developing Multi-Agent Systems with JADE*. Chichester, UK: John Wiley & Sons.

Berbeglia, G., Cordeau, J.-F., & Laporte, G. (2010). Dynamic Pickup and Delivery Problems. *European Journal of Operational Research*, *202*(1), 8-15.

Brooks, R. (1986). A Robust Layered Control System for a Mobile Robot. *IEEE Journal of Robotics and Automation*, *2*(1), 14-23.

Bräysy, O., & Gendreau, M. (2005a). Vehicle Routing Problem with Time Windows, Part I: Route Construction and Local Search Algorithms. *Transportation Science*, *39*(1), 104-118.

Bräysy, O., & Gendreau, M. (2005b). Vehicle Routing Problem with Time Windows, Part II: Metaheuristics. *Transportation Science*, *39*(1), 119-139.

Böse, F., & Windt, K. (2007). Catalogue of Criteria for Autonomous Control in Logistics. In M. Hülsmann & K. Windt (Eds.), *Understanding Autonomous Cooperation and Control in Logistics – The Impact of Autonomy on Management, Information, Communication and Material Flow* (pp. 57-72). Berlin: Springer-Verlag.

Dijkstra, E. W. (1959). A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik*, *1*(1), 269 - 271.

Dumas, Y., Desrosiers, J., Gelinas, E., & Solomon, M. (1995). An Optimal Algorithm for the Traveling Salesman Problem with Time Windows. *Operations Research*, *43*, 367-371.

Dumas, Yvan, Desrosiers, J., & Soumis, F. (1991). The Pickup and Delivery Problem with Time Windows. *European Journal of Operational Research*, *54*(1), 7-22.

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proceedings of the 2nd International Conference on Knowledge Discovery and Data mining* (pp. 226-231).

Fischer, K., Müller, J. P., & Pischel, M. (1995). Cooperative Transportation Scheduling: An Application Domain for DAI. *Journal of Applied Artificial Intelligence*, *10*, 1-33.

Gehrke, J. D., & Ober-Blöbaum, C. (2007). Multiagent-Based Logistics Simulation with PlaSMA. In M. Koschke, R.; Herzog, O.; Rödiger, K.-H.; Ronthaler (Ed.), *Informatik 2007 - Informatik trifft Logistik* (pp. 416-419).

Himoff, J., Rzevski, G., & Skobelev, P. (2006). Multi-Agent Logistics i-Scheduler for Road Transportation. *Proceedings of 5-th International Conference on Autonomous Agents and Multi Agent Systems (AAMAS 2006)* (pp. 1514-1521).

Hülsmann, M., Scholz-Reiter, B., & Windt, K. (Eds.). (2011). *Autonomous Cooperation and Control in Logistics: Contributions and Limitations - Theoretical and Practical Perspectives*. Berlin: Springer-Verlag.

Hülsmann, M., & Windt, K. (Eds.). (2007). *Unterstanding Autonomous Cooperation and Control in Logistics*. Berlin: Springer-Verlag.

Jennings, N. R. (2001). An Agent-Based Approach for Building Complex Software Systems. *Communications of the ACM*, *44*(4), 35-41.

Krajewska, M. a, Kopfer, H., Laporte, G., Ropke, S., & Zaccour, G. (2007). Horizontal Cooperation among Freight Carriers: Request Allocation and Profit Sharing. *Journal of the Operational Research Society*, *59*(11), 1483-1491.

Kuske, S., Luderer, M., & Tönnies, H. (2010). Autonomous Units for Solving the Traveling Salesperson Problem Based on Ant Colony Optimization. In H.-J. Kreowski, B. Scholz-Reiter, & K.-D. Thoben (Eds.), *2nd International Conference on Dynamics in Logistics (LDIC 2009)* (pp. 289 - 298). Springe-Verlag.

Neagu, N., Dorer, K., Greenwood, D., & Calisti, M. (2006). LS/ATN: Reporting on a Successful Agent-Based Solution for Transport Logistics Optimization. *Distributed Intelligent Systems: Collective Intelligence and Its Applications, 2006. IEEE Workshop* (pp. 213-218).

Parragh, S. N., Doerner, K. F., & Hartl, R. F. (2008). A Survey on Pickup and Delivery Problems Part II: Transportation between Pickup and Delivery Locations. *Journal für Betriebswirtschaft*, *58*(2), 81-117.

Rao, A. S., & Georgeff, M. P. (1995). BDI Agents: From Theory to Practice. *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95)* (pp. 312-319).

Sandholm, T. W. (1999). Distributed Rational Decision Making. In G. Weiss (Ed.), *Multiagent Systems: A*

Modern Approach to Distributed Artificial Intelligence. The MIT Press.

Scholz-Reiter, B., Windt, K., Kolditz, J., Böse, F., Hildebrandt, T., Philipp, T., & Höhns, H. (2004). New Concepts of Modelling and Evaluating Autonomous Logistic Processes. In G. Chryssolouris & M. D (Eds.), *IFAC Manufacturing, Modelling, Management and Control*. Athens, Greece: Elsevier.

Schuldt, A. (2011). *Multiagent Coordination Enabling Autonomous Logistics*. Heidelberg, Germany: Springer-Verlag.

Schuldt, A., Gehrke, J. D., & Werner, S. (2008). Designing a Simulation Middleware for FIPA Multiagent Systems. *2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* (pp. 109-113). IEEE Computer Society Press.

Shoham, Y., & Leyton-Brown, K. (2009). *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press.

Skobelev, P. (2011). Multi-Agent Systems for Real Time Resource Allocation , Scheduling , Optimization and Controlling : Industrial Applications. In V. Mařík, P. Vrba, & P. Leitão (Eds.), *Holonic and Mulit-Agent Systems for Manufacturing: Volume 6867 of Lecture Notes in Computer Science* (pp. 1-14).

Solomon, M. (1987). Algorithms for the Vehicle Routing and Scheduling Problems with Time Window Constraints. *Operations Research, 35*, 254-265.

Vickrey, W. (1961). Counterspeculation, Auctions, and Competitive Sealed Tenders. *Journal of Finance, 16*(1), 8-37. JSTOR.

Warden, T., Porzel, R., Gehrke, J. D., Herzog, O., Langer, H., & Malaka, R. (2010). Towards Ontology-based Multiagent Simulations: The PlaSMA Approach. In A. Bargiela, S. Azam Ali, D. Crowley, & E. J. . Kerckhoffs (Eds.), *24th European Conference on Modelling and Simulation. European Council for Modelling and Simulation* (pp. 50 - 56).

Wessels, A., Jedermann, R., & Lang, W. (2010). Transport Supervision of Perishable Goods by Embedded Context Aware Objects. *WSEAS Transactions On Circuits And Systems, 9*(5), 295-304.

Windt, K., & Hülsmann, M. (2007). Changing Paradigms in Logistics – Understanding the Shift from Conventional Control to Autonomous Cooperation and Control. In M. Hülsmann & K. Windt (Eds.), *Understanding Autonomous Cooperation and Control in Logistics – The Impact of Autonomy on Management, Information, Communication and Material Flow* (pp. 4-16). Berlin: Springer-Verlag.

Max Gath received his diploma in Computer Science from the University of Bremen (Universität Bremen) in October 2011. Subsequently, he joined the Artificial Intelligence group at the Center for Computing and Communication Technologies (TZI). He is associated with the Collaborative Research Centre 637 Autonomous Logistic Processes - A Paradigm-Shift and its Limitations (CRC 637) and working on the transfer project “Autonomous Groupage Traffic - Agent-Based Autonomous Control in Groupage Traffic”.

Thomas Wagner is a senior researcher at the Artificial Intelligence Research Group of the University of Bremen. He studied Artificial Intelligence and Computerlinguistics at the University of Osnabrück. He holds a PhD from the University of Bremen on "*Qualitative Navigation in Unstructured Environments"* and worked on different projects with varying topics like knowledge based configuration, qualitative spatial reasoning (e.g., in the *Robocup*), knowledge representation and machine learning.

Otthein Herzog started out to study Computer Science at the Universitaet Karlsruhe in summer 1969 and received his diploma in Applied Mathematics and Informatics from the Universitaet Bonn in 1972. He received his PhD from the Informatics Department of the Universitaet Dortmund in 1976. From 1977 to 1993 he worked with IBM where he held several technical and managerial positions in international software product development and research. His projects included mainframe operating systems, quality assurance (DOS/VSE and Unix, 1977-1984), communications software, full-text information retrieval systems (IBM SearchManager), CIM repository systems, and a system for environmental impact analysis (1991-1993). From 1985 to 1991 he directed the Institute for Knowledge-based Systems in the Scientific Center of IBM Germany, where he headed numerous AI research projects, the most important one being LILOG - Natural Language Processing and Text Understanding. From March 2000 to February 2002, Dr. Herzog was on leave from the university and held the CTO position at Lenze AG where he was responsible for hardware and software development. From 1993 to 2009, Dr. Herzog held the chair on Artificial Intelligence in the Department of Mathematics and Computer Science at the Universität Bremen and headed the Artificial Intelligence Research Group. He continues to supervise PhD students, to direct research projects, and to acquire new projects.

MOSIPS AGENT-BASED MODEL FOR PREDICTING AND SIMULATING THE IMPACT OF PUBLIC POLICIES ON SMES

Federico Pablo-Martí, María Teresa Gallo, Antonio García-Tabuenca, Juan Luis Santos and Tomás Mancha

Institute of Economic and Social Analysis – IAES. University of Alcala (Madrid, Spain)

federico.pablo@uah.es, maria.gallo@uah.es, antonio.gtabuenca@uah.es, juanluis.santos@iaes.es, tomas.mancha@uah.es

ABSTRACT

This paper presents MOSIPS (MOdel of Simulation of Impacts of Public Policies on SMEs), a multi-agent model of an open economy. It is part of a user-friendly object-oriented interactive intelligent policy simulation system allowing forecasting and visualizing the socio-economic potential impact of public policies for supporting SMEs. MOSIPS model specifies the characteristics of every agent, and its particular feature is that it locates agents in a spatial raster, using this information to precisely determine the interaction network. It represents the dynamic behaviour of people and enterprises, analysing their decisions and interactions in the social networks. It can be used to model macro-economic features of a system and permits focusing on a specific part of the economy, at sector and spatial level.

Keywords: agent-based model, prediction and simulation, policy evaluation

1. INTRODUCTION

Recently, agent-based models (ABM) have increased their importance in economics. In particular, the last financial crisis was not predicted by standard macroeconomic models. Due to several of their assumptions, they were not able to represent that significant deviation from the equilibrium growth path predicted. In contrast, if the approach is bottom-up, starting with the specification of the agents involved in the economy, it appears and emergent behaviour of the system which cannot be explained from the behaviour of the representative agent. This allows the appearance of bubbles, followed by a sharp reduction in prices and a lowering in expectations. Multi-agent models have been used to study economic systems in several ways: we can find examples of conceptual works on agent-based economic models, such as Tesfatsion and Judd (2006); a variety of agent-based models focused on a part of the economy, for example, leverage effects in financial markets (Farmer and Foley 2009). Multi-agent models of the economy as a whole are infrequent, examples are the models by Gintis (2006), Dosi et al. (2008), Madel et al. (2009) and the EURACE model (Dawid et al., 2011).

MOSIPS model includes a number of features of the previous referred models, but it represents the economy making the emphasis in Small and Medium Enterprises (SMEs) and the factors faced in their creation and growth. Thus, entrepreneurship and access to finance appears as two major issues. These enterprises choose their location not only optimizing the place, but taking into account the residence of the owner. This characteristic makes necessary to locate the entrepreneurs. Moreover, the demand SMEs face is determined by their location and their size, which conditions their visibility. SMEs' suppliers, workers and consumers tend to be near their location. Then, it is crucial to locate every agent in its real place to allow determining realistic interaction networks and the correct performance of every firm. For our purposes, in order to truthfully represent a local, regional or national economy, both firms and people should be placed with their individual characteristics. In addition, public administration, financial sector and the external sector are represented as well, as they interact with SMEs establishing policies, giving access to finance, competing with them or allow selling part of their production abroad.

What is described here is the abstract model. The description follows partially the agent-based model documentation guidelines developed at the 100th Dahlem Conference "New Approaches in Economics after the Financial Crisis" (Wolf et.al. 2011). As specified by these guidelines, the first part provides an overview (Section 2) and the second one explains general concepts underlying the model's design (Section 3). The fourth part provides the specification of data needing and its treatment, as it requires a much more complex process than the majority of macroeconomic models. Then, there is a description about the introduction of policies in the model. Finally, a brief conclusion looks at the further developments of the model and summarizes its major characteristics.

2. OVERVIEW

The rationale behind MOSIPS model (Figure 1) is closely related to the purpose of the project for which is designed: to develop a policy simulation system allowing forecasting and visualizing the socio-economic potential impact of public policies for supporting SMEs.

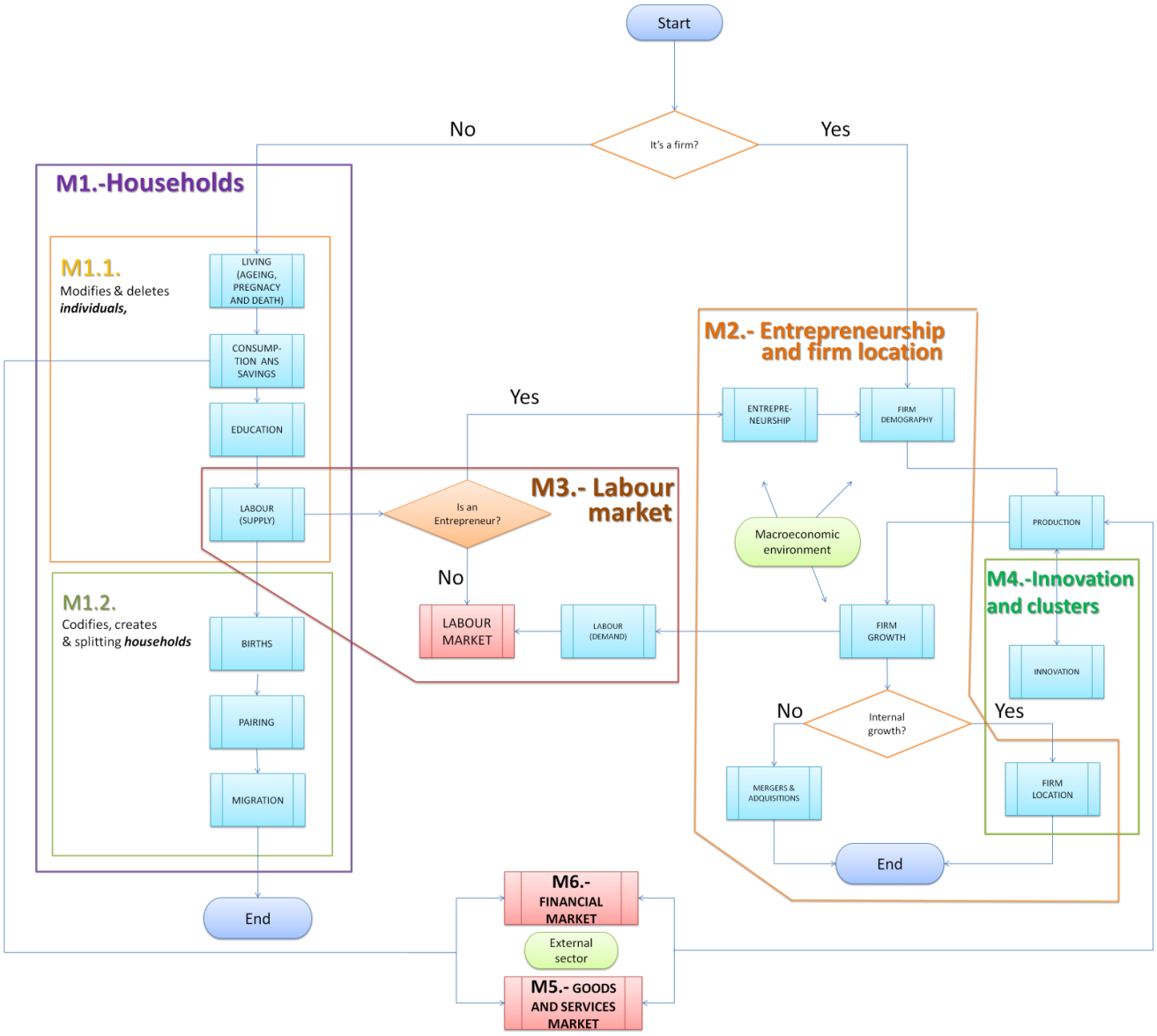


Figure 1: Diagram of modules included in MOSIPS model

2.1. Rationale

What is the object under consideration? What is the intended usage of the model? Which issues can be investigated?

MOSIPS model represents the dynamics of behaviour and decisions of agents, and their interactions. It forecasts the evolution of an economic system over a time horizon of one quarter to several years. It is based on a multi-agent approach at the micro-economic level. It can be used to model macro-economic features of a system and allow focusing in a specific part of the economy, at sector and spatial level, evaluating the effects of a policy over the firms and the individuals, depending on their initial characteristics. The concept of industrial district arises as a leading figure in which there are combined concepts such as localization and innovation, key for the performance of the different actors that are part of the socio-economic reality.

MOSIPS model provides the framework to test the accuracy of micro-foundations specified outside the scope of the representative agent paradigm reproducing a virtual reality to evaluate the effects of economic policy. The obtained results have a range of error due to the randomness of individual processes and the building of the database. This approach can be seen as an extension to Arrow-Debreu general equilibrium theory (Arrow and Debreu 1954), as the unique result computed by standard models is one of the possible outcomes: the optimal trajectory excluding part of the heterogeneity of the agents, and not having into account spatial issues with a sufficient degree of accuracy.

The object of study is a local economic system, disaggregated by branches, which represents the economic system under consideration. Then, the effects of a policy are studied both at sector and at spatial level.

These effects also can be observed according to other characteristics such as the size of the firms, their innovative behaviour or their financial situation. The effects of policies can also affect to the population, and they can also be studied taking into account their location or the individual characteristics

2.2. Agents

What kinds of agents are considered in the model? Is there a refined taxonomy of agents? Are there agent groupings which are considered relevant?

The model includes two basic types of agents: individuals-families and entrepreneurs-firms. These agents, by means of commercial or social mechanisms, interact among themselves and with other entities within their environment but which are external to the agents' identity and decisions.

MOSIPS concerns particularly these two types of agents and their scope from the moment that a part of individuals chooses to become entrepreneurs with the intention to start up a firm. While such a choice is clearly influenced by the decisions they have made previously (Stam et al. 2008, Nielsen and Sarasvathy 2011, Aldrich and Cliff 2003) (e.g., their type of education, family influence, or acceptance of a firm as inheritance), as well as by the environment created by other entities or markets (Ardagna and Lusardi 2010) (i.e., government regulations or rate of interest required on a loan to an entrepreneur), the analysis underscores the behaviours and decisions of 'entrepreneurs-firms' agents.

Partnerships among firms are also possible, so that sometimes groups or conglomerates of firms aiming at achieving common goals (for innovation or export activities for example) can be formed. Therefore, in the proposed model, groups of agents are or may be relevant (Roessl 2005; Street and Cameron 2007).

In addition, each 'firm' agent can be considered from the perspective of the 'individuals' agents comprising such firm, that is, from the different degrees of responsibility and the ability to make decisions that individuals belonging to a firm have. In this way, one can include workers, technicians, managers, directors, and owners. According to this view, which rests upon the theories on human resources and knowledge management as well as the agency problem or theory, each individual who is part of the firm takes initiatives based on simple, or sometimes complex, management options, which may even be opportunistic or contradictory to the objectives of the firm or the interests of the owner (Brunet and Alarcón 2004). These behaviours, dealt with individually (each member of the firm is an individual), would lead to a deepening or specialisation of the proposed model, which eventually would bring new ideas and approaches on firm development and growth in the territory analysed (with the information and data warehouse used), as well as any possible imbalances.

Every agent has the characteristics pointed by Wooldridge and Jennings (1995). They decide their characteristics autonomously trying to maximize their expected profits/utility. The communication among agents takes place by market prices and social networks. Agents react to the changes in their environment, but sometimes anticipate these changes in order to define their decisions, showing a proactive behaviour.

2.3. Other entities

What are the other entities which are time-evolving but not decision-making

In addition to these two basic types of agents, there appear other complementary entities for their activities involved to a higher or lesser extent in the modelling process but which are pivotal in the composition of agents. These entities do not make decisions directly in the process, but the evolution of their behaviours in time clearly impacts on the creation of the expectations and decisions of firms (and individuals). Specifically, these entities are the public sector, the financial system, and the local environment.

2.4. Boundaries

What are additional inputs to the model at runtime? Which outside influences on the model are hence represented?

The model faces a major constraint in its development: the existence and behaviour of the agents which comprise the external sector. Agents and institutions that compose it (companies, public administrations, households...) incorporate permanent or sequential new inputs to the model in its running time, interacting with businesses in the analysed territory. For example, in trade relations between import and export companies, decisions are produced on both sides, in the country and abroad. Often, in business practice are taken decisions on capital investment out of the country. However, such decisions are not fully considered in the model. In turn, these decisions are based on the behaviour of other agents or external entities that have their own behaviours and different rules (for example, labour legislation, taxation of companies, the price of industrial land, or the difference productivity in the tradable good or service).

These external effects modify the behaviour of agents in the area to be analysed and, therefore, may be measured at least indirectly or by a method of approximation. Then, all the consumers of the exported goods and services, and firms that produce the imports are considered as several entities (one for each economic region), and their behaviour is only predicted in an aggregate way making use of several macroeconomic indicators. For the purposes of international trade, the model uses international exchange rates, which express fairly accurately the

strengths and weaknesses of different economies, or what is the same level of competitiveness of domestic and foreign companies. Likewise, in the case of investment in the country by foreign companies, there are several difficulties in the determination of the behaviour of foreign agents not individually and explicitly modelled.

2.5. Relations

What kinds of relationships structure the agents' interactions? To which extent do these represent institutions?

The types of relationships that structure the interactions among agents are of a diverse nature. These are developed based on the various activities conducted by entrepreneurs and firms in their processes of recruitment and procurement of inputs, human resources management, production, innovation and technology management, product development, financial management, and marketing and sales strategy. Also influential are the possible strategies for growth and territorial expansion of investments (within or outside the territory). Individuals on themself or grouped in households have relations with other individuals and with firms resulting from their labour, consumption and investment activities.

In general, the market itself sets a 'virtual' kind of network relationships among firms in their pursuit of needs/opportunities for personnel, intermediate inputs, production equipment and technology, and sales niches, which are provided by the different types of markets (labour, goods and services…) (Coviello and Munro 1995; Slotte-Kock and Coviello 2010). But the general market also establishes the relationship of rivalry and competition between them. In sum, these are network relations that provide information for cooperation or competition as appropriate (Gulati 1999, Meyer et al. 2004).

Contractual relationships on the labour market are structured between individuals and firms. They are based on search processes, where firms in a context of imperfect information choose the best candidates and individuals offer their work to firms that provide the most attractive terms. In general, these relationships are normative since they are based on labour legislation.

2.6. Activities

What kind of actions and interactions are the agents engaged into?

Individuals are born into households where they grow, consume, pursue an education, and, eventually, die. Households may change their location and increase or decrease the number of their members. Upon reaching working age, individuals choose in each period whether to join the workforce and, if so, become employees or entrepreneurs. Job seekers offer their work to firms in their environment recruiting workers, which take the decision regarding who to hire. Individuals who choose to become entrepreneurs create firms and choose the location.

Firms produce and sell their products on the market. They also choose the cheapest and most reliable suppliers. Additionally, they change in terms of size through internal growth or by acquiring other firms, provided they have adequate funding. They can apply for funding from the financial system, based on their repayment capacity and on the financial market conditions. Further, firms decide the level of their commitment to innovation, both in terms of processes and products. They modify their workforce by hiring or laying off workers according to their labour skills or to production needs. Firms disappear if they go bankrupt or if the businessperson so chooses.

The public sector incorporates its rules and policies by modifying the attributes and behaviours of agents. Banks procure funds from agents with a funding capacity, and they provide funds to those who require them. The financial resources available for firms and households can differ from the level of savings of agents due to the financial flows with other countries and the circumstances of the financial system. The local environment includes aspects that affect the agents from a territorial perspective such as closeness from infrastructures, the existence of firm clusters, or congestion problems.

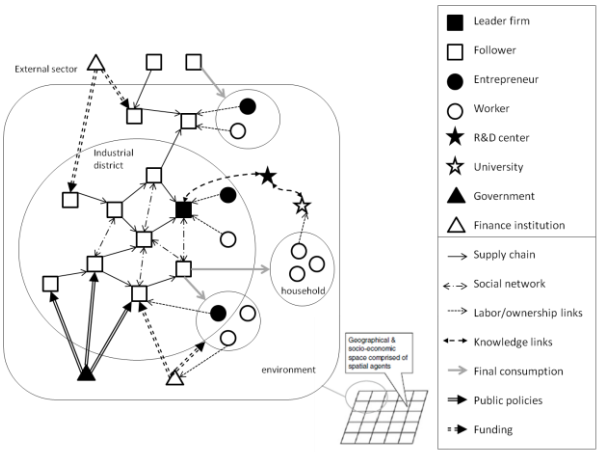


Figure 2. Agents and environment incorporated in MOSIPS. Based on Albino et.al (2007).

The kind of relations that structure the interactions between agents have different nature depending on the agents that form part of them, and the activity or characteristic that links them. These relations are developed based on the various activities undertaken by companies in their recruitment and procurement of inputs, human resources management, production, innovation and technology management, product development, financial management and business strategy sales. Individuals establish their relations with firms from their consumption and saving decisions, employment and ownership. Social relations with other

individuals take place through companies (co-owners, work colleagues, etc.) or with some of their relatives and neighbours.

In the modelling process, the properties of the protocols that govern the interaction between individuals and companies are based on relations of production and consumption, employment and lending, adopting a microeconomic perspective. In this sense, the price of competitors' products in relation to themselves constitutes one of the main signals received by the agents. They act taking the suitable decisions for the acquisition of inputs, recruitment of factors, production and sales. These relations take place in markets.

With respect to these factors and product markets, the model provides an approximation to local environments, on one side. But on the other side it also assumes the existence of other broader environments, at national or international level. That is, the analysis of interaction of agents is focused on their interest in a defined territorial space, such as a region. Relationships with agents from other areas are analysed in a more simplified approach. For example, the majority of goods are bought by large retailers in international markets, and then they sell them to small retailers. Final customers do not have access to a high number of sellers due to informational costs. Then, the appropriate scale of the first market is international, small retailers only have access to the regional market and final costumers tend to purchase goods at local level.

The same assumptions are made in the relation between firms and workers, who are unemployed or employed in other companies and they want to change their jobs. It could be considered that both companies and workers face the regional employment supply and demand, respectively. In most of the approaches, it is assumed that all the agents act against the market, the aggregate behaviour of the rest of agents, looking to optimize their interest. However, in most of the cases, every agent creates its own behaviour associated to the decisions of its neighbours. It arises from the information and expectations generated by the rest, weighted depending on their spatial and relational proximity. For example, a company located in a municipality is able to produce and sell its production with a slightly different price from a competitor of a neighbouring municipality, while in other part of the region prices can be lower. Then, agent actions and decisions are highly affected by the behaviours of agents in the proximity, but it also depends on the aggregate behaviour, emerged from the decisions of every agent.

Thus, all companies are somehow interconnected, but these links are stronger in environments which are closer. In any case, those behaviours associated with the environment may also depend on the sector, the concentration of supply and demand or the degree of public promotion of a product (e.g., which is derived from the impact of advertising).

Individuals face the same interaction protocols and information flows, but applied to their decisions. They obtain most of the information from firms which they are linked, but also from the aggregate behaviour (e.g. the unemployment rate, GDP growth, price index). Individuals also condition their decisions taking into account the performance of other agents who are linked with. Then, a potential entrepreneur will decide to create his own enterprise with a higher probability if both their acquaintances and the information she has about the general performance of the economy is promising for her success

3. DESIGN CONCEPTS

This section of the paper presents some of the highlights of the general approach of the model

3.1. Time, activity patterns and activation schemes

What is the basic sequence of events in the model? Are activities by agents triggered by a central clock or by actions respectively messages sent by other agents?

Time is modelled discretely. Each period consists of several steps. The length of the period determines the temporal resolution of the model and is determined largely by the characteristics and temporal reference of the data used. The model can consider any time interval without affecting its characteristics. However, the quarter has been taken as the primary reference since it is considered that most of the decisions of the agents have a maturation period around this length.

Actions are triggered instantly at the time when the 'central clock' determines each period. They, however, do not need to be carried out in each period. The user can choose a different periodicity for some of them.

By observing this basic temporal sequence of events, the model is fed with information and data proposed in the system architecture for the years 2007 to 2011. This means that the simulation system starts from 2007 and forecasts of the modelling can be developed from 2012.

3.2. Interaction protocols and information flows

What are the general properties of the protocols governing the interaction between agents? How is determined which agents can interact with each other

Matching interactions and business activities are bilateral. These are gravitational interactions where intensity depends on "visibility", which, for an agent, means the expected relevance of its interaction with the counterpart.

In the case of matching of individuals, each individual selects a group of people with whom s/he

interacts and who s/he subjectively evaluates based on its attributes. In the case of firms, sellers offer their product to the market and buyers choose their supplier from a group of sellers who are selected according to their closeness and to the size of their firms.

Each firm demands workers featuring certain characteristics. Among the firms seeking a worker's profile, workers choose the most "attractive" ones in terms of salary and distance. Matching occurs when the best possible combination for both parties is achieved.

3.3. Forecasting

Are agents in the model forward looking or purely backward looking? If agents are forward looking, what is the basic approach to modelling forecasting behaviour?

Agents base their forecasts on their past experience, within a context of incomplete information. Households determine their levels of consumption and savings from their income experience in prior periods following a scheme inspired by the life cycle and permanent income hypotheses. Firms make their decisions based on the experience gained with clients and competitors.

3.4. Behavioural assumptions and decision making

Based on which general concepts is decision making behaviour of the different types of agents modelled?

Agents have bounded rationality and act in an environment of imperfect information. Interactions take place predominantly in the close environment of the agents. The chances of interaction among agents depend on their "visibility", understood as indicated in section 3.2. A possibility to model it is the following:

$$V_i(t) = \left(\sum_{0}^{j} (1-\delta_3)^j \, (size_i(t-j) + adv_i(t-j)) \right) + \left(\sum_{1}^{k} size_k(t) \, SP_{i,k}^2 \right)$$

The visibility of each firm is determined by its current and past size, and its advertisement effort. Agglomeration increases the value of this variable, adding the size of the k nearest firms, weighted by the spatial proximity.

Firms select their suppliers and their workers. Consumers choose the firm in which they work and their consumptions elections. In all the situations the process is the same. Firstly, the agent evaluates its performance. For this purpose, it examines not only its results, but also receives information about other agents. It can be biased and incomplete. If the performance is sufficiently good, the agent does not search for new agents (suppliers, workers, etc.), whereas if its performance is not sufficiently good, according to the expectations of the agent, or worse than the average perceived performance, the firm will look for better agents to work with (Figure 3).

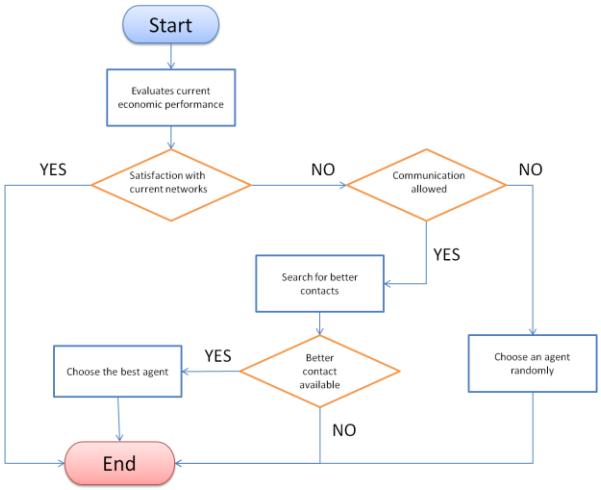


Figure 3: Social links formation and updating process. Based on Giardini et.al. (2008)

3.5. Learning

Are decision rules of agents changed over time? If yes, which types of algorithms are used to do this?

The structural characteristics of households and firms evolve through learning. This is done by imitation and mutation procedures. The structural characteristics of individuals are modified in each period either due to random factors or by imitation of the behaviour of the agents regarded as displaying more appropriateness (like benchmarking).Thus, by means of a selection evolution process, agents get adapted to the circumstances through learning, as it happens, for instance, in the case of the initial reservation wage, or the choice of distance when seeking suppliers.

These learning processes are represented by the changing of several structural characteristics (consumption patterns, mark-ups or technology) of households and firms.

Agents tend to imitate other ones of their same kind which are in their network and show a better behaviour. Not always the adoption of a new behaviour is positive, every so often the better result obtained could have happened by chance. Thus, imitation has not always the desired effects. Additionally, previous behaviour does not warranty current or future success because of the unforeseen or unexpected sharp changes in the economy. A good example is found in the real estate sector during bubbles. It shows high profits during a certain amount of time, but the last starters only achieve losses.

3.6. Population Demography

Can agents drop out of the population and new agents enter the population during a simulation run? If yes, how are exit and entry triggered?

In the model, entries and exits of households and firms occur in each period.

Within households, there are births and deaths. Births depend on the location and personal circumstances of the mother, while deaths hinge on the individual's age as well as other factors. Individuals can change their location when appropriate in a cost-benefit scheme.

Some individuals are entrepreneurs, and they emerge as such when, from a subjective point of view, it is convenient for them to be entrepreneurs. Similarly, they stop being entrepreneurs when it no longer suits them. Entrepreneurs with one or more firms are businessmen.

The birth and death of firms follow the decisions taken by businessmen, who are dependent on economic or personal factors. They also make decisions about the location of their firms.

3.7. Level of randomness

How do random events and random attributes affect the model?

There are two main sources of randomness. The major one is generated in the creation of agents. There is available a certain information, but it is incomplete, and agents are created following random rules, accordingly and conditioned to their known characteristics. For example, an employed person who works in a given sector is matched to a firm of that sector with an establishment near his residence, but there is a possibility of make incorrectly this process. Then, if the actual firm has a good performance and the matched firm goes bankrupt, this individual will lose his job, while in the reality, he continues working. The counterpart is an actual unemployed will continue be working in the modelled reality. As we can see, part of these random possibilities cancel at aggregate level, but can have huge effects at the micro level, making unfeasible to obtain accurate micro information.

Another significant source of randomness is the creation of networks, as this is the part of the model for which there is a lower level of data available. Networks are extremely important for learning processes and changing behaviours. Thus, the model increases its randomness along periods of simulation, leading to stochastic dynamics for a large number of variables (household characteristics, prices, amounts, innovation, location, etc.).

3.8. Miscellaneous

Any important aspects of the used modelling approach that do not yet any of the items above should be explained here

The MOSIPS model is inspired by the circular pattern of income where the financial field is explicitly integrated. This is crucial in the current crisis process given the serious financial constraints of firms.

Additionally, MOSIPS provides a highly precise spatial outlook since agents are located individually using GIS techniques, which makes it possible to observe the impact of policies at a micro-spatial level.

4. DATA BASES FOR THE MODEL

In order to represent the society it is necessary to build two databases, one for individuals and families and other for firms and establishments. They are complemented with the macroeconomic environment, and public, financial and foreign sector, taken as a whole.

The variables for the macroeconomic environment and public policies, allow the agents information creating their expectations and it is added to the information they have from themselves and other individuals and firms whom they have connections.

We make use of two techniques that allow achieve a great grade of accuracy in the process: statistical matching and downscaling. Thus, the model has a high degree of scalability, and allows the user focusing from individual effects up to total variation or putting the focus in the agents of a specific area or a group of SMEs with similar characteristics.

4.1. Statistical Matching

Statistical matching is used to fusion information from different microdata sources (Diorazio 2006)

In the model the main one for individuals is the Population Census, which informs about the number of the total population and the main individual characteristics: age, gender, location at regional level, family composition, level of studies, etc. It is complemented with the Labour Force Survey to become acquainted labour status of individuals. The EU Labour Force Survey is a large household sample survey providing quarterly results on labour participation of people aged 15 and over as well as on persons outside the labour force. All definitions apply to persons aged 15 years and over living in private households. The Global Entrepreneurship Monitor explores the role of entrepreneurship in national economic growth, unveiling detailed national features and characteristics associated with entrepreneurial activity. It is used to value the chances of developing new enterprises. Mobility surveys make possible to know patterns of mobility of individuals, and allow creating demand functions spatially defined (Schenk et.al. 2007).Tax Income Panel is the best indicator of income and Household Budget Survey permits identify the consumption in every sector and the savings as the remaining part of the income. It allows us to ascertain the consumption expenditure of households residing, as well as the distribution of said expenditure among the different consumption divisions. Other sources are used to acquire information not present in the previous presented sources. The main one is mortgage duration

and amount, present in the Annual Report of Property Registers.

Enterprises data base is built in a similar way, starting from the Business Directory, which contains information about the number of enterprises, the sector they belong and their size. However, is a poor source with respect to the Population Census and only includes aggregate information. Then, it must be fulfilled with the microdata exhaustive information enclosed in other statistical source called Amadeus.

This data base should be enlarged including characteristics from the Survey on Access to Finance and Technical Innovation Panel. The Survey on Access to Finance covers micro, small, medium-sized and large firms and it provides evidence on the financing conditions faced by SMEs compared with those of large firms every six months. In addition to a breakdown into firm size classes, it provides evidence across branches of economic activity, euro area countries, firm age, financial autonomy of the firms, and ownership of the firms. The Technological Innovation Panel is a statistical instrument for studying the innovation activities of firms over time. It takes into account the heterogeneity in the firms' decisions (such as different shares of intramural R&D and external R&D in total innovation expenditures) or in the effects (such as the different impacts on productivity).

The statistical matching process is taken in several stages. For individuals is necessary to start from the Population Census, and then add labour information though a microsimulation model that is built in order to obtain accurate data. Then, we are sure that every individual is coherent with himself across the time periods. For example, a civil servant cannot be fired. Then, it is possible to link information about income and consumption at household level, due to the available microdata information. Finally, entrepreneurial activity is compute for every adult. Enterprises data base is constructed following the same scheme, starting with the census (business directory) and adding information about their characteristics (number of workers, financial statements...) and access to finance, which is a major issue for the growth and overall performance of SMEs.

4.2. Downscalling

In order to accurately define the placement of agents, a raster of locations must be included, with information of land price, demography, uses of land and transport networks. This raster is built making use of downscaling techniques. From information at local level or other (e.g. per Ha or per Km2) we arrive to the level of each cell included in the raster. (Gallego 2010)

After making the statistical matching process, both firms and households are located into the raster. Finally, networks are built between individuals, firms and among individuals and firms. These relations are related to the information provided by markets of factors, goods and services.

5. PUBLIC POLICIES

The model permits the study of almost any policy that the public administration implements which has effects on SMEs, not only the policies designed specifically to encourage business activity. All the principles included in the Small Business Act (European Commission, 2008) have been taken into account and turned into a concrete set of policy domains. Then, the model allows the inclusion of any policy related to entrepreneurial activity, infrastructure, innovation, internal managing, inter-firm relations, labour market, funding, relations with the administration, environment and macro-economic environment such as changes in taxation or trade restrictions.

These policy domains can affect the agents in three different ways, as it is shown in Figure 4. They can change their cost function, their characteristics (data) or their behaviour. They affect the performance, development and decay of every industrial district present in the economy, as well as the processes undertaken between the components of the district such as cooperation, competition, innovation and knowledge dissemination.

Figure 4: Illustration of policy domains included in MOSIPS model.

Policies can be integrated in the model as appears in the above figure as it is shown in the following example.

It is possible to think about two kinds of subsidies that affect innovation: a grant intended to increase the potential of SMEs in order to acquire high-tech machinery and another one that is granted for training their employees in a R&D course. In the first case, it affects the cost function by reducing the costs of producing a certain amount of product. Instead, the second measure affects the behaviour of SMEs, making them more able to produce and adopt innovations, but in the next period they will produce with the same cost function and data (employees, machinery…) as if the policy had not taken place. However, in the following periods, it would behave in a more pro-innovative way, adapting and developing new procedures that will reduce costs.

6. CONCLUSIONS AND OUTLOOK

We have discussed the characteristics of agent-based models. Agents, other entities, relationships, activities and decision rules within the model MOSIPS have been described. This allows to predict and to simulate public policies on SMEs.

The opportunity to deal with real data makes the model a powerful tool to predict the real functioning of the economy at individual level. Consequently, it is necessary to develop complex databases that include the agents heterogeneity.

ACKNOWLEDGMENTS

The work presented in this document has been conducted in the context of the EU Framework Programme project with Grant Agreement 288833 MOSIPS (Modeling and Simulation of the Impact of public Policies on SMEs). MOSIPS is a 36 months project started on September 1st, 2011.

REFERENCES

Albino, V., Carbonara, N. and Giannoccaro, I. 2007. Supply chain cooperation in industrial districts: A simulation analysis. *European Journal of Operational Research*, 177, 261-280.

Aldrich, H. E. and Cliff, J. E. 2003. The pervasive effects of family on entrepreneurship: toward a family embeddedness perspective, *Journal of Business Venturing*, 18(5), 573-596.

Ardagna, S. and Lusardi, A. 2010. Explaining International Differences in Entrepreneurship: The Role of Individual Characteristics and Regulatory Constrains". In Lerner, J. and Schoar, A., eds. *International Differences in Entrepreneurship*, University of Chicago Press, 17- 62.

Arrow, K.J. and Debreu, G. 1954. The existence of an equilibrium for a competitive economy. *Econometrica*, 22, 265-290.

Brunet, I. and Alarcón, A. 2004. Teorías sobre la figura del emprendedor, *Papers: revista de sociología*, 73, 81- 103.

Coviello, N. and Munro, H. 1995. Growing the entrepreneurial firm: networking for international market development, *European Journal of Marketing*, 29(7), 49-61.

Dawid, H., Gemkow, S., Harting, P., van der Hoog, S., and Neugart, M. 2011. The Eurace@Unibi Model: An Agent-Based Macroeconomic Model for Economic Policy Analysis. Available from www. wiwi.uni-bielefeld.de/vpl1/projects/eurace/eurace-unibi.html.

DíOrazio, M., DiZio, M. and Scanu, M. 2006. *Statistical Matching: Theory and Practice*, Wiley, New York.

Dosi, G., Fagiolo, G., and Roventini, A. 2008. Schumpeter Meeting Keynes: A Policy-Friendly Model of Endogenous Growth and Business Cycles. Laboratory of Economics and Management, SantAnna School of Advanced Studies, LEM Working Paper Series, 21.

European Commission. 2008. Think small first: A Small Business Act for Europe, DG Enterprise, Brussels, Belgium.

Farmer, J. and Foley, D. 2009. The economy needs agent-based modelling. *Nature*, 460, 685-686.

Gallego, F.J. 2010. A population density grid of the European Union, *Population and Environment*, 31, 460–473.

Giardini, F., Di Tosto G. and Conte, R. 2008. A model for simulating reputation dynamics in industrial districts. *Simulation Modelling Practice and Theory*, 16, 231-41.

Gintis, H. 2006. The emergence of a price system from decentralized bilateral exchange. *B. E. Journal of Theoretical Economics*, 6, 1302-1322.

Gulati, R. 1999, Network location and learning: The influence of network resources and firm capabilities, *Strategic Management Journal*, 20(5), 397-420.

Mandel, A., Fürst, S., Lass, W., Meissner, F., and Jaeger, C. C. 2009. *Lagom generiC: an agent-based model of growing economies*. ECF Working Paper, 1. Available from www.europeanclimate-forum.net/index.php?id=ecfworkingpapers.

Meyer, M., Aderhold, J. and Duschek, S. 2004. Organizing social complexity in production networks, *Journal of Academy of Business and Economics*, 3(1), 1.

Nielsen, K., and Sarasvathy, S. 2011. Who Re-enters Entrepreneurship? And Who Ought to? An Empirical Study of Success After Failure, in Dime-Druid Academy, *Winter Conference in Aalborg*, 20-22 Jan, Aalborg, Denmark.

Roessl, D. 2005. Family Businesses and Interfirm Cooperation. *Family Business Review*, 18, 203-214.

Schenk T., Löffler G., Rauh J. 2007. Agent-based simulation of consumer behaviour in grocery shopping on a regional level. *Journal of Bussiness Research*, 60, 894-903.

Slotte-Kock, S. and Coviello, N. 2010. Entrepreneurship Research on Network Processes: A Review and Ways Forward. Entrepreneurship Theory and Practice, 34(1), 31-57.

Stam, E., Audretsch, D. and Meijard, J. 2008. Renascent entrepreneurship. *Journal of Evolutionary Economics*, 18 (3-4), 493-507.

Street, C. T. and Cameron, A.-F. 2007. External Relationships and the Small Business: A Review of Small Business Alliance and Network Research, *Journal of Small Business Management*, 45(2), 239- 266.

Tesfatsion, L. and Judd, K. 2006. Agent-Based Computational Economics. 2 in *Handbook of Computational Economics*. Elsevier, North-Holland.

Wolf, S., Bouchaud, J-P., Cecconi, F., Cincotti, S., Dawid, D., Gintis, H., Hoog, S. Jaeger, C.C.,

Kovalevsky, D.V., Mandel, A., Paroussos, L. (2011). Describing economic agent-based models, Dahlem ABM documentation guidelines. Proceedings of the 100th Dahlem Conference, New Approaches in Economics after the Financial Crisis, 2010 Aug 28-31.

Wooldridge, M. and Jennings, N.R. 1995. Intelligent agents: theory and practice. *Knowledge Engineering Review*, 10, 115-152.

AUTHORS BIOGRAPHY

Federico Pablo-Marti received the M.Sc. in industrial economics from University Carlos III in 1995 and the Ph.D degree in economics and business administration from University of Alcala in 2000. Previously he received the diploma in energy negotiate from Spanish Energy Institute–ICADE and worked as analyst in the Ministry of Industry of Spain, and as Reader in European University of Madrid, up to 1994. Currently, he is Associate Professor in University of Alcala with the Economic Structure, Statistics and International Economic Organization Department, University of Alcala, Alcala de Henares, Spain. He is also one of the researchers in the Institute of Economic and Social Analysis, in the Department of Sector, Financial and SME Studies. He has participated in several European Commission financed projects. His research interests include agent-based simulation, entrepreneurship, geographic information systems, firm demography and new approaches on the study of international trade.

María Teresa Gallo received the degree in economics from National University of Piura, Peru, and PhD in economics from University of Alcala. Senior Lecturer of Applied Economics in University of Alcala, She has developed his professional activity in universities and research centers integrating multidisciplinary teams of work in Peru and Spain. She has also participated in projects of development aid with incidence in the private sector, community organizations and public administration. Currently she is a researcher in the Department of Territorial and Urban Studies, in the Institute of Economic and Social Analysis. Her basic research lines focus on the regional economy, primarily in the analysis of growth and regional disparities and the urban and rural development. She has been involved in several investigations as the economic and social cohesion policy; location of enterprises and the effects of diffusion in the Community on Madrid; Polycentric model and dispersion, among others.

Antonio García-Tabuenca received the Ph.D in economics from the University of Alcala in 1997 and he is Associate Professor since 2000 with the Economic Structure, Statistics and International Economic Organization Department, University of Alcala, Alcala de Henares, Spain. After finishing his degrees in economics and in law in Deusto University in 1974 and 1973 respectively, he served in the public administration in several positions such as Associate Director in Madrid Development Institute from 1988 to 1994, Director in SME Institute, in the Ministry of Industry, President of the State Society for the Development of Industrial Design, President of the National Innovation Company, Vice-President of the European Association of Regional Development Agencies, among other significant occupations. Currently he is the Associate Director in the Institute of Economic and Social Analysis and Director in the Department of Sector, Financial and SME Studies. His research interests are firm funding and innovation, and entrepreneurial activity.

Juan Luis Santos has received the M.Sc. in Applied Economic Analysis from University Complutense of Madrid and University of Alcala. Currently is working in his Ph.D. thesis. He is one of the researcher assistants in the Department of Territorial and Urban Studies, in the Institute of Economic and Social Analysis. His research interests are simulation models, agent-based modeling, social networks and industrial districts.

Tomás Mancha received the Ph.D in economics from the University of Malaga in 1982. He worked as a researcher in London School of Economics in 1985-86 and in Dormunt University in 1990. He is Professor since 1992 in Alcala University, where he has occupied various academic positions. He is the current Associate Director of the Department of Applied Economics, University of Alcala, Alcalá de Henares, Spain. He is visiting Professor in several universities such as Santa Fe and Salta Universities in Argentina, and Magallanes University in Chile. He has worked has expert and assessor for Interamerican Development Bank, Spanish Government, European Commission and many regional and local public administrations. He is the current Director in the Institute of Economic and Social Analysis. His research interests are economic policy, regional economics, European integration and political-economic cycles

MONTE CARLO SIMULATION - THE BANK ACCOUNT SELECTION IN THE CZECH REPUBLIC ACCORDING TO THE BANK CHARGES

Martina Kuncova[(a)], Lenka Lizalova[(b)]

[(a)]College of Polytechnics Jihlava, Dpt. of Economic Studies, Tolsteho 16, 58601 Jihlava, Czech Republic
[(b)]College of Polytechnics Jihlava, Dpt. of Economic Studies, Tolsteho 16, 58601 Jihlava, Czech Republic

[(a)]kuncova@vspj.cz , [(b)]lizalova@vspj.cz,

ABSTRACT
Simulation methods belong to the suitable instruments that can be used in the real world situations to better understand the reality or to make a responsible decision. Monte Carlo simulation is a method for iteratively evaluating a deterministic model using sets of random numbers as inputs. This method is often used when the model is complex, nonlinear, or involves more than just a couple of uncertain parameters. As in the Czech Republic the situation with the selection of the appropriate bank account is complicated (because of the non-transparent bank charges), we have created a simulation model to find the best account for 3 different types of retail clients. We compare our results with the solution obtained from MS Excel, Crystal Ball and the multi-criteria evaluation of alternatives model.

Keywords: Monte Carlo Simulation, Bank Account, Bank Charges

1. INTRODUCTION

Simulation nowadays means a technique for imitation of some real situations, processes or activities that already exist in reality or that are in preparation – just to create a computer model (Banks 1998). The reasons for this are various: to study the system and see how it works, to find where the problems come from, to compare more model variants and select the most suitable one, to show the eventual real effects of alternative conditions and courses of action, etc. Simulation is used in many contexts, including the modeling of natural or human systems in order to gain insight into their functioning (manufacturing, automobile industry, logistics, military, healthcare, etc.), simulation of technology for performance optimization, safety engineering, testing, training and education.

As simulation models use principles taken from mathematics and statistics, they are sometimes added to the problematic of the operational research or management science (Turban Meredith 1994) where different models are constructed to find the optimal solution or the optimal choice. Simulation itself usually has not the main aim to find the best alternative but it might help in this process.

In economy we must face a lot of decisions that have to be made, and pay a lot of money afterwards often without knowing whether we have done right or wrong. When everything is given, the solution or decision can be based on the common sense or on the solution of some mathematical model (optimization on or as a result of the decision-making model). But the problem is that a lot of things not only in economy are not certain – especially when we think about money spent for something. People are usually able to describe the expenses as "something between 8 and 12 thousand crowns" or "15 thousand crowns at a medium". Although it seems to be vague, inaccurate and insufficient, with some knowledge of statistical distributions we are able to use given information and even make a decision or recommendation via Monte Carlo simulation model.

Monte Carlo simulation can be used in any situation where we would like to calculate some outputs that are dependent on random inputs. This is typical situation for various decision-making processes. Simulation in finance and banking is not still widely spread (especially in the Czech Republic in real-life situations). As we face an increasing problem with high bank charges, we have decided to use Monte Carlo simulation to find out the cheapest bank according to the charges and random amount of demanded services. We describe the situation in the banking sector, the simulation model and we compare the results with the results taken from the static decision-making model when the WSA method is used.

2. CZECH BANKING SECTOR

The banking sector of the Czech Republic is formed by 44 subjects (CNB 2012), where 17 are the real banks (4 big, 4 middle, 9 small ones), the rest includes foreign bank branches and building societies. The structure of the market is nearly steady but during the last two years the situation has changes because 3 new banks came to the Czech market – Fio Bank, Air Bank and Zuno Bank AG. These small banks (the first of this type was mBank) have different strategies especially in small or zero fees and they are aimed at the usage of the internet banking. The main part of the sector is still hold by the group of four big banks (KB, CSOB, CS, UniCredit)

whose share of the actives of the whole banking sector was nearly 58% in 2010. The announcement of the Czech National Bank (2012) says that nearly 74% of the non-interest profit of the Czech banking sector is made by the profit taken from the bank charges and bonuses. As you can see on the Figure 1, the profit from these fees is still rising (between the years 2009 and 2010 the increase was more than 5%).

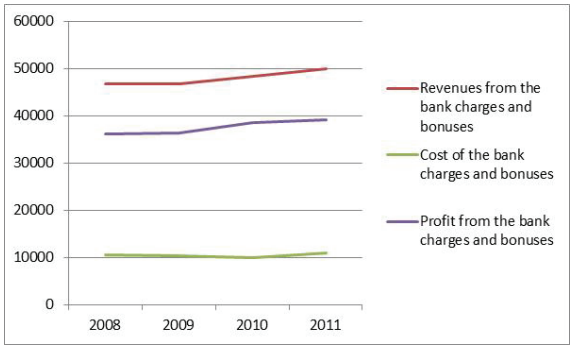


Figure 1: Revenues and Cost of the Bank Charges and Bonuses (www.bankovnipoplatky.com)

The growth rate of the profits is slowing down – there might be different reasons for this trend but one of them is the increasing competition (coming of the new banks) and also easier change of the bank by clients. The report of the Capgemini company (2011) created in cooperation with ING and EFMA shows the main changes and trends in retail banking sector in all over the world – and it shows that fees are the third most important reason for leaving a bank – 50% of customers change a bank because of the fees (Figure 2).

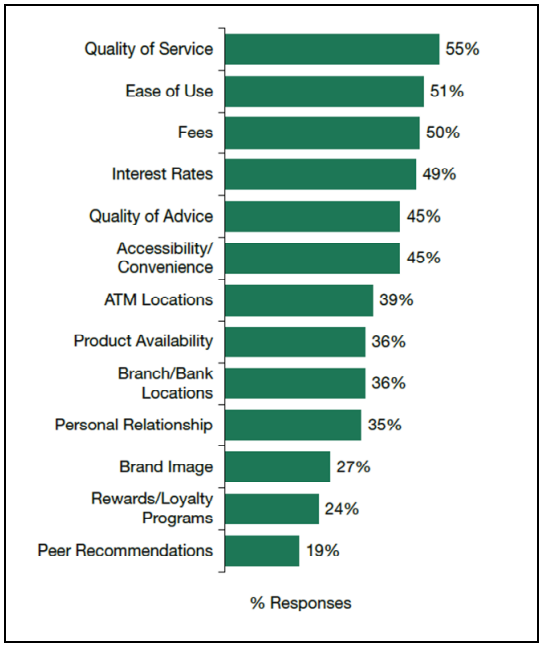


Figure 2: Factors That Affect Why Customers Leave a Bank Revenues and Cost of the Bank Charges and Bonuses (www.bankovnipoplatky.com)

From the previous information it is clear that the bank charges (or fees) are very important not only in the Czech Republic but all over the world. But there is still a lot of banks that things fees are not so high to persuade the client to leave. So our simulation model shows how different the real situation can be. One of the reasons why we use Monte Carlo simulation is the fact that it is not possible to state that each month the behavior of the person connected with the bank account transactions is the same. That is why we suppose this kind of calculations are more precise than the results you can have from various calculators aimed at the Czech bank account fees (www.bankovnipoplatky.com; www.penize.cz; web pages of the Czech banks).

3. MONTE CARLO SIMULATION

Monte Carlo simulation (or technique) is closed to statistics as it is a repeated process of random sampling from the selected probability distributions that represent the real-life processes (Turban, Meredith 1994). On the basis of the existed information we should select the type of probability distribution (that corresponds with our expectations about the values of the variable, and we are able to define all the parameters for).

The problem of some economic models is the lack of the information – especially in the retail sector sometimes only managers themselves know how the process works, what the typical number of customers during a period is etc. In this kind of situations we cannot use basic statistical or mathematical models as we do not have the strict or real data. That is why Monte Carlo simulation can help as it uses random variables from different distributions. The most typical and frequent distribution types are normal, triangular, uniform (discrete uniform), Poisson, lognormal and exponential ones. Mathematical specification of these variables and the calculations derived from them might be complicated (especially when non trivial distribution is chosen). But via the simulation Monte Carlo and via MS Excel and its add-ins (for example Crystal Ball) it is possible to analyze the problem and find a solution or a recommendation for each specified situation (Kuncova 2006).

3.1. MS Excel Usage for Simulation

MS Excel spreadsheet can be a good environment where to start Monte Carlo simulation, since almost nearly all people working with computer know how to work with it, although Excel is not the best place to run a scientific simulation, especially discrete or continuous one. Excel contains a pseudo random number generator that was tested for sufficiency in 1991 by Law and Kelton (Law 2000). The function is invoked using the Excel function =RAND(). It generates a uniformly distributed pseudo random number between 0 and 1. Its values can be easily updated by pressing the Calculation Key F9 (every press means new simulation experiment) or we can have more simulation experiments using data tables. Via this generator it is possible to generate

random variables having any other distribution – see Table 1.

If the distributions described in the Table 1 are sufficient to describe all uncertain variables that we have, it is possible to use an Excel sheet to solve the problem – only define the interconnections between the variables and specify the decision function. If other distributions are necessary it is better to use some add-in application such as Crystal Ball, @Risk, Lumenaut, Simtools, Formlist, MonteCarlito, Simulacion 4.0, SimulAr, Risk Analyzer, etc. For our analysis we use Excel and Crystal Ball.

Table 1: The Excel expressions for generation of random variables from given distribution (Kuncova 2006)

Distrib.	Parameters	Excel Expression
Uniform	a,b	=a+RAND()*(b-a)
Normal	μ,σ	=NORMINV(RAND(), μ, σ)
Lognormal	μ,σ	=LOGINV(RAND(), μ, σ)
Exponential	1/λ	=(-1/ λ)*LN(RAND())

3.2. Crystal Ball

Crystal Ball is one of the MS Excel add-in applications for the Monte Carlo simulation models. "Oracle Crystal Ball solutions begin with Oracle Crystal Ball, the base package for predictive modeling, Monte Carlo simulation and forecasting. Oracle Crystal Ball Enterprise Performance Management builds on that set of tools by adding the power of Oracle Crystal Ball to your Oracle Enterprise Performance Management (EPM) and Oracle Business Intelligence (BI) applications" (www.oracle.com). The advantage of this software is the usage of Excel tables, so it is possible to use models created before but change the distribution for random inputs generation. Then usually 1000 trials are run and afterwards the programme gives all statistics (and histogram) of the selected decision cell. Figure 3 shows all the possible statistical distributions that can be chosen.

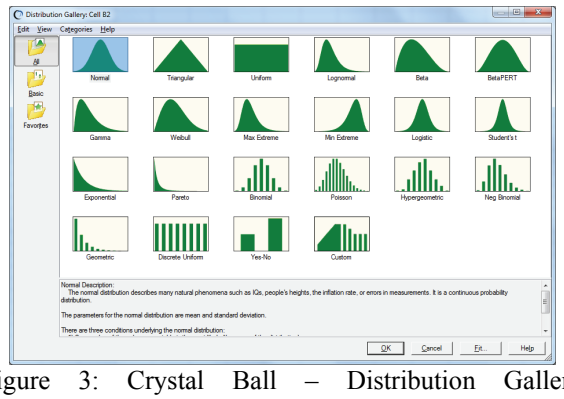


Figure 3: Crystal Ball – Distribution Gallery (www.oracle.com)

4. MATERIALS AND METHODS

Before we start the analysis we have to select the alternatives (bank accounts), the criteria and the distributions for the random variables generation. As we would like to compare the results with the order of the accounts created by the WSA method, we shortly describe also principles of this method.

4.1. Bank Accounts

Each from the 17 banks offers different types of accounts. As we are interested in the retail market (the accounts for the physicians) we have selected following 12 banks that offer this type of account:

- Komercni banka (KB) (www.sazebnik-kb.cz)
- CSOB (www.csob.cz)
- Ceska sporitelna (CS) (www.csas.cz)
- mBank (www.mbank.cz)
- Fio Bank (www.fio.cz)
- Raiffeisenbank (www.rb.cz)
- City Bank (www.citybank.cz)
- Airbank (www.airbank.cz)
- GE Money Bank (www.gemoney.cz)
- Volksbank (www.volksbank.cz)
- UniCredit Bank (www.unicreditbank.cz)
- LBBW Bank (www.lbbw.cz)

According to the up-to-date scale of charges we have found out all the fees at the web pages of the banks.

4.2. Criteria Selection

The main aim is the comparison of the bank accounts according to the various bank charges. We have found 20 cases when bank wants a fee for something – usually for the cash withdrawal, incoming or outgoing payments (it differs if it is to or from the own bank or another bank). As the cash deposit made at the desk is free of charge for all selected bank accounts, we have following 19 criteria where it is necessary to pay a fee:

1. monthly account charges
2. electronic debit card
3. cash withdrawal from ATM of own bank
4. cash withdrawal from ATM of another bank
5. incoming payment (from another bank)
6. incoming payment (from own bank)
7. outgoing payment to the own bank (at the desk)
8. outgoing payment to the own bank (via internet)
9. outgoing payment to another bank (at the desk)
10. outgoing payment to another bank (via internet)
11. permanent order for payment to the own bank (at the desk)
12. permanent order for payment to the own bank (via internet)
13. permanent order for payment to another bank (at the desk)

14. permanent order for payment to another bank (via internet)
15. collection (own bank, at the desk)
16. collection (own bank, via internet)
17. collection (another bank, at the desk)
18. collection (another bank, via internet)
19. cash withdrawal (at the desk)

4.3. Decision-Making and WSA Method

Multi-criteria evaluation of alternatives belongs to the category of discrete multi-criteria decision-making models where all the alternatives and criteria are known. To solve this kind of model it is necessary to know the preferences of the decision maker. These preferences can be described by aspiration levels (or requirements), criteria order or by the weight of the criteria (Evans 1984). WSA (Weighted Sum Approach) method belongs to the group where the importance of the criteria is given by their weights. This method sorts the alternatives based on the values of their utility functions which in this case are assumed to be linear. Higher value of utility means better alternative. The calculations are easy and can be made is MS Excel or in some add-in applications like Sanna (Jablonsky 2009).

4.4. Random Variables for 3 Client Types

On the various web pages you can find a lot of calculators that should help you to find the best account for your needs (usually with the lowest fees). Sometimes not only fees are important so it is not comparable with our case – but the problem of the calculators is that they need exact data (such as number of transactions, number of ingoing and outgoing payments etc.). But these numbers are not fixed in real-life situations, they usually differs from month to month. That is why we think that Monte Carlo simulation is better than the calculators.

Inspired by Hedvicakova and Soukal (2011), we have created 3 types of the clients for whom we are looking for the best account with the minimum bank charges. Each client has its own frequency of withdraws and its communication channel. These clients are:

- Active Client
- Branch Office Client
- Average Internet Client

The differences can be seen from the Table 2 where number of each activity is specified.

An active client uses all the paid services more frequently than the others. Branch office client usually prefers the cash withdraws at the desk (in the branch office) than from ATM. Average internet client takes advantage of the internet banking. As the description of the clients' behavior is given by two numbers (minimum and maximum) we started the random generation of the numbers of each activity with the uniform distribution.

Table 2: The types of the clients and monthly numbers of each activity

Number of the activites per month	Active Client	Brach Office Client	Avg.Internet.client
monthly account charges	1	1	1
electronic debit card	1	1	1
cash withdrawal from ATM of own bank	2 to 6	1 to 4	2 to 5
cash withdrawal from ATM of another bank	0 to 3	0 to 1	0 to 2
incoming payment (from another bank)	2 to 6	1 to 3	1 to 4
incoming payment (from own bank)	1 to 4	0 to 2	0 to 2
outgoing payment to the own bank (at the desk)	0 to 1	1 to 3	0
outgoing payment to the own bank (via internet)	2 to 7	0 to 2	1 to 3
outgoing payment to another bank (at the desk)	0 to 1	1 to 4	0
outgoing payment to another bank (via internet)	3 to 7	0 to 2	2 to 6
permanent order for payment to the own bank (at the desk)	0 to 1	1 to 3	0
permanent order for payment to the own bank (via internet)	2 to 4	0 to 1	0 to 2
permanent order for payment to another bank (at the desk)	0 to 1	1 to 4	0
permanent order for payment to another bank (via internet)	3 to 6	0 to 1	2 to 5
Collection (own bank, at the desk)	0 to 1	0 to 2	0
Collection (own bank, via internet)	1 to 3	0	0 to 1
Collection (another bank, at the desk)	0 to 1	0 to 3	0
Collection (another bank, via internet)	1 to 3	0	0 to 3
cash withdrawal (at the desk)	0 to 1	0 to 2	0

5. RESULTS

First we summarize the task: to compare 12 selected accounts according to the 19 criteria with respect to three different types of clients. Table 3 summarizes the bank charges for each of the selected criteria (mentioned above).

Table 3: Bank charges for the account and criterion

crit.	KB	CSOB	CS	mBank	Fio	Raiffeisenbank	City Bank	Airbank	GE Money Bank	Volksbank	UniCredit Bank	LBBW Bank
1	50	30	69	0	0	75	169	0	59	47	50	100
2	200	540	200	0	0	540	0	0	708	492	200	600
3	5	6	6	9	0	10	0	0	15	4	5	6,5
4	35	35	40	35	35	40	30	25	40	35	30	6,5
5	5	6	7	0	0	5	0	5	5	0	6	2
6	5	6	5	0	0	7	0	0	5	0	6	0
7	29	40	0	0	0	48	49	0	50	0	45	35
8	6	6	5	0	0	4	0	0	6	0	0	5
9	29	40	0	0	0	50	0	5	50	50	45	40
10	6	6	7	0	0	6	0	5	6	5	0	5
11	39	6	5	0	0	40	0	0	0	0	40	5
12	6	3	5	0	0	9	0	0	0	0	0	0
13	39	6	7	0	0	40	0	5	0	50	40	5
14	6	6	7	0	0	9	0	5	0	6	0	5
15	39	6	5	0	0	7	0	0	0	0	40	5
16	0	6	5	0	0	7	0	0	0	0	0	5
17	39	6	7	0	0	7	0	5	0	50	40	5
18	0	6	7	0	0	7	0	5	0	6	0	5
19	100	60	0	35	30	60	49	0	60	60	55	55

5.1. Results from Monte Carlo Simulation in MS Excel

We have generated the numbers of the transactions for each client according to the Table 2 (from the uniform distribution). We compare the results of 24 experiments that show us the amount paid by the client for 2 years. Table 4 shows the results for the active client - minimum, maximum and average fee per month, sum per 2 years, order of the bank accounts and difference from the best account. You can see that the best current account offers Fio bank and the customer should pay about 1530 CZK (Czech koruna) per two year at average. On the other hand in GE Money Bank it should be more than 25 thousand CZK, so the difference between the best and the worst two-year's fee is more than 23 thousand CZK which is nearly the average monthly salary in the Czech Republic.

The same situation we have for the branch office client and average internet client – all the results with the average monthly fee are on the Figure 4.

Table 4: Simulation results for the active client

Bank account	MIN	MAX	AVG	SUM	order	dif.
KB	471	751	604	14486	7	12866
CSOB	761	997	884	21210	9	19590
CS	412	624	527	12647	6	11027
mBank	18	194	102	2454	2	834
Fio bank	0	135	68	1620	1	0
Raiffeisenbank	876	1132	1017	24411	11	22791
City Bank	199	357	267	6410	4	4790
Airbank	65	180	119	2865	3	1245
GE Money bank	916	1196	1055	25327	12	23707
Volksbank	639	856	766	18386	8	16766
UniCredit Bank	384	599	500	11995	5	10375
LBBW Bank	814	977	908	21793	10	20173

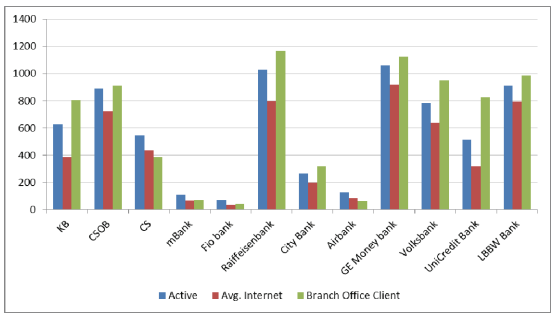


Figure 4: The average monthly fees for all current accounts and three different client types

5.2. Results from Monte Carlo Simulation in Crystal Ball

For Monte Carlo simulation in Crystal Ball it is not necessary to change much in the Excel file – only instead of random generation in Excel we have to define assumption for each cell where random variables should appear. As the next step the cells that contain our results (monthly fees) should be defined as forecast. Afterwards Crystal Ball runs 1000 simulation experiment and gives us all the statistics and histograms for all 12 accounts. Figure 5 shows the comparison of the average monthly fee for Fio bank current account and GE Money Bank current account (when discrete uniform distribution for the number of activities has been used) for the active client.

It is also possible to view statistics for each result or for more – at Figure 6 you can see the comparison of the best three accounts. The results are similar to the Excel ones.

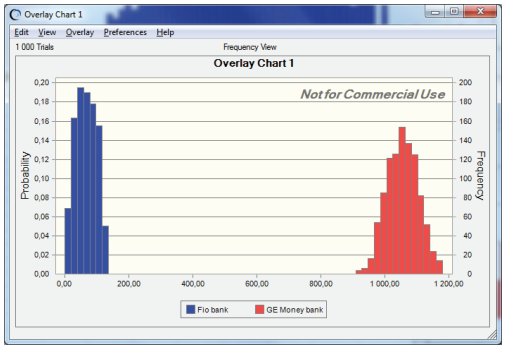


Figure 5: Results for the average fees for Fio bank and GE Money Bank accounts – active client

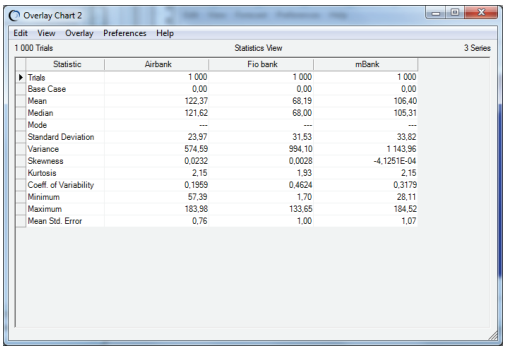


Figure 6: Statistics for 3 best accounts – active client

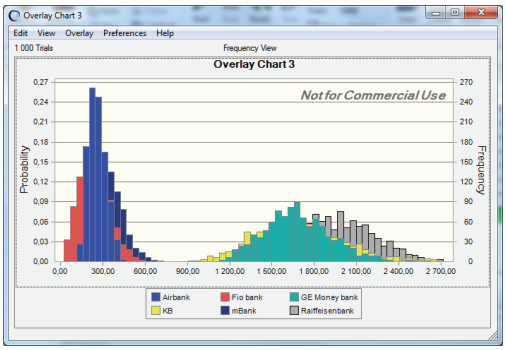


Figure 7: Results for three best and three worst bank accounts (triangular distribution) – active client

As there are a lot of possibilities while using Crystal Ball, we have tried an extreme case when an active client may need each transaction 10 times a month at maximum. For this situation we have chosen triangular distribution where minimum was the same as in previous case, maximum is 10 and the likeliest value is the mean taken from the previous (uniform) distribution for the number of transactions. On the figure 7 you can see the difference between the best and the worst accounts – the situation is still the same, for

the active client it is good to have a current account in Fio bank, mBank or Airbank.

Back to the first asumptions we have tried the simulation with the uniform distributions for the three types of clients – and for all of them Fio bank account seems to be the best one – see Figure 8.

For the branch office client and average internet client the situation is the same – the best three accounts are from Fio bank, mBank and Airbank – Figures 9 and 10. So if we think only about the fees (not about the number of offices or number of years the bank is on the Czech market) we may say that the small new banks compete succesfully with the 4 biggest. The best from the big four banks are Unicredit Bank (for active and average internet client) and CS (for brach office client).

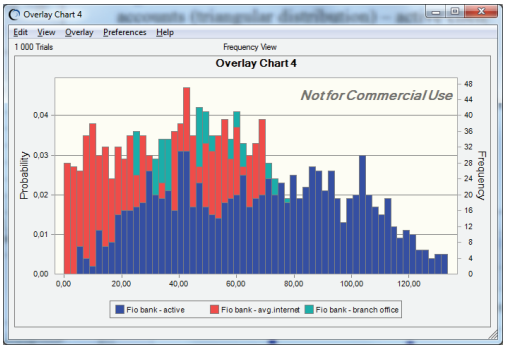


Figure 8: Results of the average monthly fee for all types of clients and Fio bank account

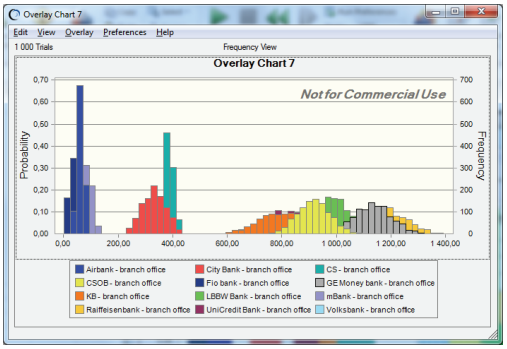


Figure 9: Average monthly fees for the branch office client (Volksbank is hidden behind CSOB and GE)

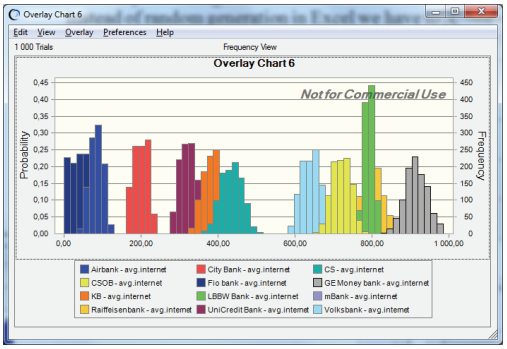


Figure 10: Average monthly fees for the average internet client (mBank is hidden behind Fio and Airbank)

5.3. Results from WSA method

To be able to calculate the order of the bank accounts via WSA method we need weights of the criteria. As the sum of weights must be equal to 1, it is better to use points for each criterion and then recalculate it into weights. For each type of the client the most important criterion is the first one (monthly account charges), so we put there 10 points. Second important might be the debit card (5 points). For the rest of the criteria we have added as much points as the mean of the expected number of transactions is. The weight vector can look like in the Figure 11.

The order of the bank accounts according of the total utility (calculated by WSA method in Sanna add-in application) is similar to the simulation results. The order for the active client is on the Figure 12, other results are in Table 5. Also WSA method chooses Fio bank account as the best one for all the types of clients.

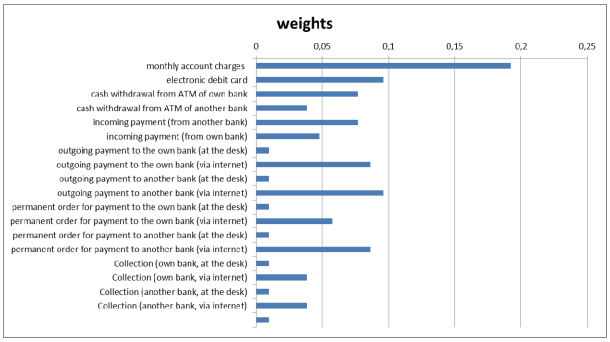


Figure 11: The weights of the criteria (active client)

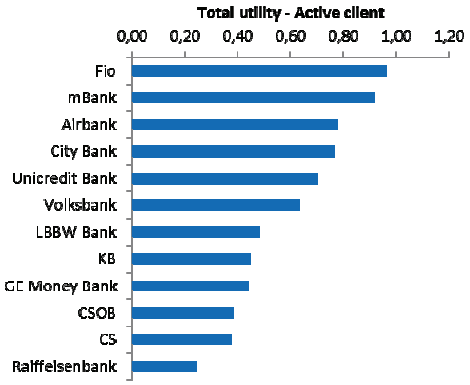


Figure 12: Total utility calculated by WSA method (active client)

	active client		avg. internet client		branch office client	
	utility	order	utility	order	utility	order
KB	0,4486653	8	0,5121974	7	0,4533868	10
CSOB	0,3837324	10	0,4183387	9	0,5322561	7
CS	0,3751688	11	0,3827847	11	0,6398752	5
mBank	0,9177598	2	0,9168804	2	0,937661	2
Fio bank	0,9643944	1	0,9760353	1	0,9801268	1
Raiffeisenbank	0,2423374	12	0,2779611	12	0,3212198	12
City Bank	0,7665772	4	0,6985495	5	0,6492946	4
Airbank	0,7766996	3	0,7687051	3	0,9083081	3
GE Money bank	0,4380828	9	0,4107128	10	0,4358596	11
Volksbank	0,6325093	6	0,5911011	6	0,5754712	6
UniCredit Bank	0,6989662	5	0,7397417	4	0,4814653	9
LBBW Bank	0,4815537	7	0,4286697	8	0,4969971	8

Table 5: Results from the WSA method

6. CONCLUSIONS

Monte Carlo simulation is a good tool that might help with the complicated decision especially when some factors that influence the decision are random. In this article we showed a simple model in MS Excel and Crystal Ball that uses Monte Carlo simulation to find the best current account offered on the Czech market from the bank charges (or fees) point of view. The results of all the models (included the multi-criteria evaluation of alternatives WSA method) show that the small new banks like Fio, mBank and Airbank are the best and all types of clients can save about 90% of money spent on fees. It is true that the situation can be different if we think about the number of branch offices of the bank, about the history the credibility or the time the bank operates on the Czech market. But as it has been mentioned, nowadays fees are very important for customers and our results show that the big banks should do something to be more competitive. The simulation model can be changed according to the different types of clients or by the selection of another statistical distribution that corresponds more with the client behavior – these facts can influence the results but the principle of the model stays the same.

REFERENCES

Airbank, 2012. *Rate tariff*, Available from: http://www.airbank.cz/cs/o-air-bank/dokumenty/cenik-jako-soucast-pribalu-k-rs/Contents/0/1141A1C4F3CE47123C7C4FDE5C08D9E4/resource.pdf [accessed 26 March 2012]

Bankovni poplatky – kalkulator, 2012. *Calculator of the bank accounts.* Available from: http://www.bankovnipoplatky.com/kalkulator.html [accessed 20 March 2012]

Banks, J., 1998. *Handbook of Simulation*. USA, John Willey & Sons

Capgemini, 2011: *Retail Banking World Report*. Available from :http://www.capgemini.com /services-and-solutions/by-industry/financial-services/solutions/banking/wrbr11/ [accessed 4 April 2012]

City Bank, 2012. *Rate tariff*, Available from: http://www.citibank.cz/czech/gcb/personal_banking/czech/static/pdf/citikontoplus_cz.pdf [accessed 26 March 2012]

CNB – Czech National Bank, 2012. *Zpráva o dozoru nad finančními trhy (Announcement about the Finantial Markets Control),* Available from: http://www.cnb.cz/miranda2/export/sites/www.cnb.cz/cs/dohled_financni_trh/souhrnne_informace_fin_trhy/zpravy_o_vykonu_dohledu/download/dnft_2010_cz.pdf [accessed 4 April 2012]

CSAS, 2012. *Rate tariff*, Available from: http://www.csas.cz/banka/nav/osobni-finance/sazby-a-poplatky-d00013670 [accessed 26 March 2012]

CSOB, 2012. *Rate tariff*, Available from: http://www.csob.cz/cz/Csob/Sazebniky/Stranky/Sazebnik-pro-fyzicke-osoby-obcany.aspx#dbkarty [accessed 26 March 2012]

Evans, G.W.. 1984. An Overwiev of Techniques for Solving Multiobjective Mathematical Programs. *Management Science*. 30 (11), 1268-1282.

Fio, 2012. *Rate tariff*, Available from: http://www.fio.cz/bankovni-sluzby/bankovni-ucty/bezny-bankovni-ucet [accessed 26 March 2012]

GE Money Bank, 2012. *Rate tariff*, Available from: http://www.gemoney.cz/ge/cz/1/ucty/bezne-ucty [accessed 26 March 2012]

Hedvicakova, M., Soukal, I., 2011. *Možnosti sledování trendu ve zpoplatnění základních bankovních služeb (Possibilities of Monitoring Trends of Charges in the Core Retail Banking Services),* Available from:http://www.cjournal.cz/files/52.pdf [accessed 3 April 2012]

Jablonsky, J., 2009. Software Support for Multiple Criteria Decision Making Problems. *Management Information Systems.* 4(2), 29–34

KB, 2012. *Rate tariff*, Available from: http://www.sazebnik-kb.cz/cs/obcane/depozitni-produkty/mujucet-a-g2-2-balicky-v-konceptu-mojeodmeny.shtml [accessed 26 March 2012]

Kuncova, M., 2006. Practical Application of Monte Carlo Simulation in MS Excel and its Add-ons – The Optimal Mobile Phone Tariffs for Various Types of Consumers in the Czech Republic. *Proceedings of Mathematical Methods in Economics 2006 Conference*, pp. 323-332. September 13-15, Pilsen (Czech Republic)

Law, A., 2000. *Simulation Modelling and Analysis. Boston (USA),* MC-Graw Hill

LBBW Bank, 2012. *Rate tariff*, Available from: http://www.lbbw.cz/cs/nasi-klienti/osobni-bankovnictvi/sazebnik/index.shtml [accessed 26 March 2012]

mBank, 2012. *Rate tariff*, Available from: http://www.mbank.cz/informace-k-produktum/sazebnik-osobni-finance/#tabs=6 [accessed 26 March 2012]

Oracle, 2012. *Crystal Ball*. Available from: http://www.oracle.com/us/products/applications/crystalball/index.html [accessed 1 April 2012]

Penize.cz, 2012. Bank accounts rates. Available from: http://www.penize.cz/bankovni-poplatky [accessed 1 April 2012]

Raiffeisenbank, 2012. *Rate tariff*, Available from: http://www.rb.cz/attachements/pdf/obecne-dokumenty/cenik-pi/cenik-produktu-sluzeb-soukrome-os_2011.pdf [accessed 26 March 2012]

Turban , E., Meredith, J., R., 1994. *Fundamentals of Management science.* 6th ed. USA, Richard D.Irwin Inc.

UniCredit Bank, 2012. *Rate tariff*, Available from: http://www.unicreditbank.cz/cz/sazebnik/obcane/osobni-konta.html [accessed 26 March 2012]

Volksbank, 2012. *Rate tariff*, Available from: http://www.volksbank.cz/vb/jnp/cz/sazebniky/obc

ane/index.html?url=bezne-ucty-a-depozita#bezne-ucty [accessed 26 March 2012]

AUTHORS BIOGRAPHY

Martina Kuncova: She has got her degree at the University of Economics Prague, at the branch of study Econometrics and Operational Research (1999). In 2009 she has finished her doctoral study at the University of West Bohemia in Pilsen (Economics and Management). Since the year 2000 she has been working at the Department of Econometrics, University of Economics Prague, since 2007 also at the Department of Economic Studies of the College of Polytechnics Jihlava. She is a member of the Czech Society of Operational Research, she participates in the solving of the grants of the Grant Agency of the Czech Republic, she is the coauthor of three books and the author of many scientific papers and contributions at conferences. She is interested in the usage of the operational research, simulation methods and methods of multi-criteria decision in reality

Lizalova Lenka: She has got her degree at the Technical University in Brno, at the branch of study Economy and management (1989). After studies she has start to work for bank „Ceska sporitelna a.s.“ as a programmer and she worked there for 8 years. Since the year 2004 she has been working at the Department of Economic Studies of the College of Polytechnics Jihlava. In 2007 she has finished her doctoral study at the Mendel University in Brno Faculty of Business and Economics (Finance). She is a councillor of the Higher Education Development Fund, she participates in the solving of the grant of the Grant Agency of the Czech Republic, she is the author of many scientific papers and contributions at conferences. She is interested in the company performance, finance, banking and insurance.

INERTIAL ACCELERATION APPLICATION FOR WHEEL SLIP MEASUREMENT OF MOBILE ROBOTS

Daniel Szocs[(a)], Teodor Pană[(b)], Andrei Feneșan[(c)], Ioana Vese[(d)]

[(a), (b), (c), (d)]Technical University of Cluj-Napoca, Department of Electrical Drives and Robots

[(a)]Daniel.Szocs@edr.utcluj.ro, [(b)]Teodor Pana@edr.utcluj.ro, [(c)]Andrei.Fenesan@edr.utcluj.ro, [(d)]Ioana.Vese@edr.utcluj.ro

ABSTRACT

Smartphone three-axis acceleration sensors mounted on a mobile robot in conjunction with motor wheel encoders are used to determine wheel slip. Inertial acceleration is obtained by means of gravity compensation. Signal drift is observed when determining the vehicle speed by integrating the inertial acceleration data. A Kalman filter is used to correct noise from sensor data and attenuate the drift effect. Vehicle speed determined from acceleration sensor data is compared to wheel encoder data to determine wheel slip.

Keywords: sensor, filter, traction, robot

1. INTRODUCTION

Currently there is a wide spread of sensor equipped phones mainly used in game applications. In this paper, a three-axis acceleration sensor is mounted on top of a National Instruments Robotics Starter Kit™ mobile robot and uses the inertial component of the acceleration in conjunction with motor wheel encoders to determine wheel slip.

The principle of the wheel slip controller (Hori, Toyoda and Tsuruoka1997) for the Robotics Starter Kit is the following: when slip coefficient values exceed a certain value, the behavior of the active wheel is changed to increase grip. A PID controller imposes a rotational speed correction for the wheels and traction is regained. As opposed to other methods which require torque data (Cai 2010) or estimated tire force (Hong 2006), this method uses vehicle speed in conjuction with wheel speed to determine wheel slip.

The encoder speed is the speed measured by the optical encoder on the motor (Bräunl 2003). Sensor speed is the speed from the integrated acceleration sensor values. The sensitivity of the motor controller that commands the motor by translating PWM signal width to rotor speed is neglected in this model. This means the desired wheel speed and the encoder speed are considered to have the negligible differences. This is not true for small values of desired wheel speed, where small PWM command values are unable to determine motor spin.

A model of the vehicle is generated in MATLAB/Simulink™ and slip is simulated. Raw acceleration data is calibrated for amplitude and offsets that are dependent on the orientation of the accelerometer. Euler angles are used to determine gravitational acceleration. Inertial acceleration is extracted from the raw data by removing the gravitational component. A Kalman filter is used to smooth out the signal. Speed is obtained by integrating inertial acceleration. Wheel slip is measured using the speed data from inertial acceleration and wheel speed from the encoders.

2. TRACTION CONTROL METHODS

Traction control can be classified into the following: longitudinal control improves adhesion by controlling the traction force, and lateral control of yaw by varying the steering angle. Recently a lot of electric vehicles have been developed, mainly hybrid models using both electric and fossil fuel energy. The main catalyst for the development of electric vehicles is to solve current energy and environmental issues caused by combustion engine vehicles (ICV). Electric motors have instantaneous high torque, quickly and precisely controlled as opposed to internal combustion engines.

One paper (Hori, Toyoda and Tsuruoka1997) proposes two methods for traction control for electric vehicles: "Model Following Control (MFC)" and "Optimal Slip Ratio Control". In the MFC, the actual speed is compared to the simulated speed output of the vehicle model; motor torque is reduced and traction is increased.

The side force has a maximum value when λ=0 and decreases rapidly for bigger λ. Sudden decrease of road friction causes λ to increase, the side force decreases as well. This has serious consequences: front wheel drive cars drift, rear wheel drive cars spin, and four wheel drive cars drift and rotate.

In the Optimal Slip Ratio Control, the road condition estimator decides the optimal slip ratio, the slip ratio controller uses this information to control the torque and obtain the desired slip ratio.

Slip ratio is defined as:

$$\lambda = \frac{(V_W - V)}{V_W} \quad (1)$$

where V_W is wheel speed and V is vehicle speed.

The kinematic equations of the wheel and of the vehicle are as follow:

$$(F_m - F_d)/(M_W \cdot s) = V_W \quad (2)$$

$$F_d/(M \cdot s) = V \quad (3)$$

where F_m is motor torque (force equivalent); F_d is friction force; M_W is wheel inertia (mass equivalent); M is vehicle weight.

The vehicle body can be seen as one inertia system with equivalent inertia moment J, by introducing the slip ratio λ,

$$J = J_W + M \cdot r^2(1 - \lambda) \quad (4)$$

where J_W, M and r are the shaft inertia moment, vehicle weight and tire radius.

2.1. Wheel Slip Ratio Control For Electric Vehicle

The principle of the wheel slip controller for the electric vehicle prototype Robotics Starter Kit is the following: when slip coefficient values exceed a certain value, the behavior of the active wheel is changed to increase grip. A PID controller imposes a rotational speed correction for the wheels and traction is regained (Fig. 1).

The desired wheel speed is the command for the rotational speed of the wheel. The encoder speed is the speed measured by the optical encoder on the motor (Bräunl 2003). Sensor speed is the speed from the integrated acceleration sensor values. The sensitivity of the motor controller that commands the motor by translating PWM signal width to rotor speed is neglected in this model. This means the desired wheel speed and the encoder speed are considered to have the negligible differences. This is not true for small values of desired wheel speed, where small PWM command values are unable to determine motor spin.

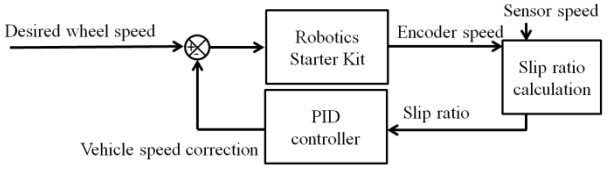


Figure 1: Wheel Slip Ratio Controller for the Robotics Starter Kit Electric Vehicle Prototype

3. ROBOT COMPONENTS AND FUNCTIONALITY

For the purpose of modeling the behavior of an electric automobile, a four-wheeled robot was used. Two motors drive the wheels by means of reduction gears so that the rotational speed at the wheels is half the speed of the spinning DC motor shaft. All the electrical components are placed on a solid plane attached to the robot chassis. An ultrasonic sensor mounted on a servomotor serves for orientation purposes. The robotic platform NI Robotics Starter Kit is comprised of a sbRIO9631™ embedded device with AI, AO, DIO, 1M Gate FPGA. The programming is realized using the LabVIEW™ graphical development environment, programs are compiled using Xilinx™ tools and written to field programmable-gate arrays on NI reconfigurable I/O hardware through the Ethernet port.

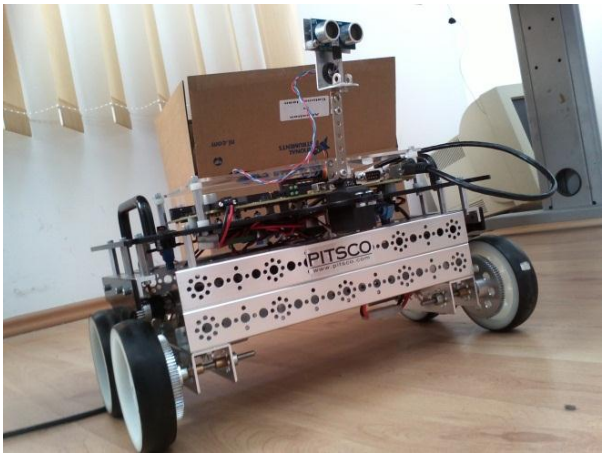


Figure 2: Robotic Starter Kit Electric Vehicle Prototype Platform – Photo Taken Inside the PhD Office at National Instruments Cluj-Napoca Headquarters

Digital I/O Port P2 is connected to a mezzanine board, which further connects to the motors controller, encoder sensors, ultrasonic sensors and its servomotor controller. A 12 V, 3000 mAh Ni-MH battery serves as a power supply and a 12 V - 24V DC converter adjusts the power requirements for the sbRIO9631™. Two DC motors are powered and controlled by a dual DC motor controller. Rotation is sent to the wheel using a 1/2 gear ratio.

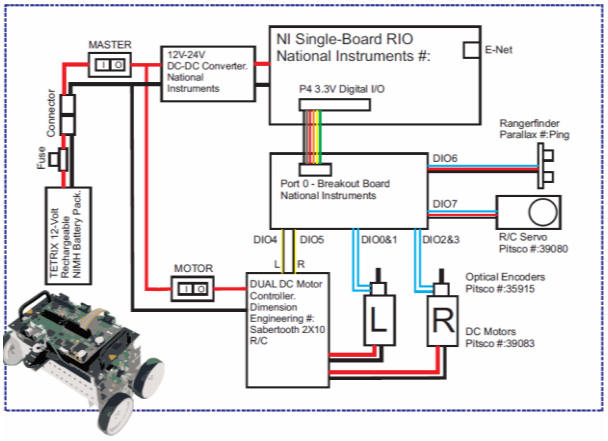


Figure 3: Labview Robotics Starter Kit Block Diagram

Motors are two RC servo controlled continuously rotating motors. A PWM signal is generated to control the motors: 1000(μs) pulse width corresponds to full backward speed, 1500(μs) to motor stop and 2000(μs) to full forward speed. Only the pulse width affects the speed, not the amplitude or the frequency. The duty

cycle and current amperage affect torque. Amplitude of current dictates the amplitude of torque, and the duty cycle affects the jerkiness of the torque. For example, a 0.5 duty cycle means half of the time torque will be one hundred percent and half of the time it will be zero. Frequency of the signal affects the command rate, or how often one can change the command in a period of time.

The signal command is generated through ports DIO4 and DIO5 for the left and right motor respectively. Note that the motors have opposite motor movement, due to the opposite physical position of the motors shaft, this affects the PWM command.

Rotary optical encoders, one for each of the two motors, give absolute position feedback with 400 position increments. Data is sent through DIO0, DIO1 for the left motor and DIO2, DIO3 for the right.

In LabVIEW™, forward and angular velocity are transformed into left and right wheel velocities. A differential steer and fixed wheel frame is defined, wheel radius, wheel base and ratio are set.

4. SLIP SIMULATION IN MATLAB/SIMULINK™

The vehicle is modeled using equations (1), (2), (3) and the Matlab/Simulink graphical programming environment. From the vehicle and wheel model as well as a given motor torque, we determined wheel drift.

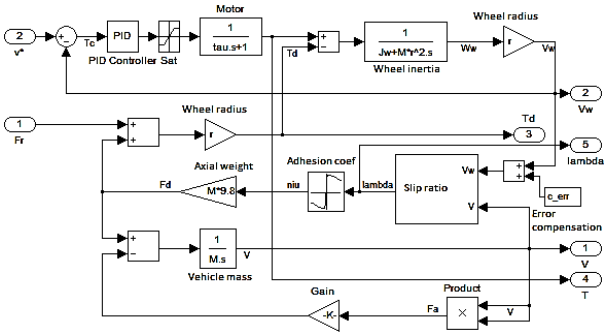


Figure 4: Vehicle Model Structure in Simulink. Generates Wheel and Vehicle Speed and Determines Slip Ratio

The following parameters have been defined:

Table 1. Wheel Slip Simulation Parameters

Simulation parameter name	Value(unit)
resistant wheel friction force, F_r	6.125(Nm)
wheel radius, r	0.0508(m)
wheel inertia moment, J_w	0.00077(kg·m^2),
vehicle mass, M	3.6(kg)
air density, ρ_a	1.205(kg⁄m^3)
air friction coefficient, c_{fa}	0.3
front vehicle area, A	0.035(m^2).

From 0 to 800(ms) there is an acceleration interval where the vehicle accelerates to the maximum speed: 0.3(m⁄s) (Fig. 5) and the torque is at maximum constant value (Fig. 6).

Speed is constant from 800 to 1000(ms); motor torque decreases and has same value as resistant torque. Until the 1550(ms) mark, the vehicle brakes and stops, motor torque is negative and less than resistant torque. Total simulation time is two seconds.

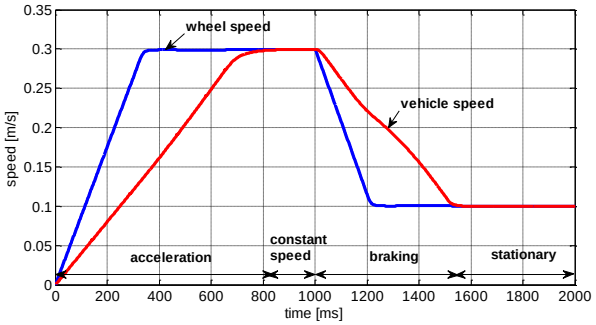


Figure 5: Wheel and Vehicle Speed Characteristic Simulation in Simulink; Four Phases of Movement: Acceleration, Constant Speed, Braking and Stationary

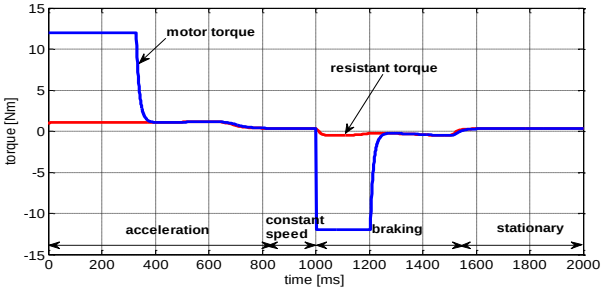


Figure 6: Motor and Resistant Torque Characteristic Simulation in Simulink

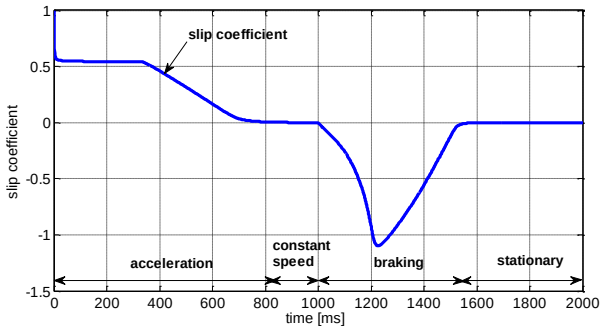


Figure 7: Slip Coefficient Characteristics Simulation in Simulink

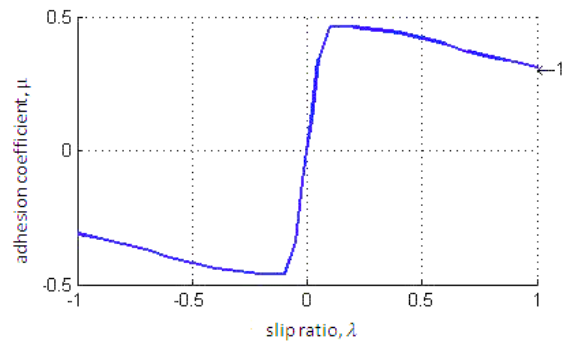


Figure 8: Adhesion Coefficient as a Function of Slip

ratio, by interpolating discrete values in a reference table.

In figure 8, the μ−λ characteristic used in the simulation is generated by interpolating discrete values in a reference table. Wheel speed, vehicle speed, motor torque and resistant torque are the main parameters defining the vehicle behavior. Using (1), slip coefficient is found (Fig. 7).

5. SENSOR SIGNAL ACQUISITION AND CONDITIONING

Sensors are provided from an LG P970™ smartphone and data is transmitted via Bluetooth. The phone contains a gyroscope, an accelerometer, a magnetic compass sensor.

Currently acceleration sensor, magnetic sensor and orientation information is available to the LabVIEW™ environment, using an Android™ application that transmits wireless Bluetooth sensor information.

Raw acceleration data contains the gravitational acceleration as well as inertial linear acceleration. From the acceleration data, the gravitational component is removed by means of orientation compensation and the result is linear acceleration. Gravitational acceleration has a constant value as long as the sensor moves parallel to the ground plane; inertial acceleration or linear acceleration along one axis is the acceleration generated by movement.

5.1. Amplitude Calibration for Gravitational Acceleration

Consider a stationary accelerometer, oriented with one of its axis in the vertical place, so the measured values for each axis must read +g for the positive direction of the down pointing oriented axis, or -g for the negative axis pointed in the same direction. After measurements are done using the accelerometer, there is a noticeable deviation from the gravitational constant. Local gravitational acceleration is considered, obtained from latitude and altitude values (Boynton, 2001).

Furthermore, the amount of deviation can vary from one accelerometer device to another. Calibration will result in three linear equations, one for each axis. The equations can then be applied to correct the acceleration values and are unique for each device.

Orientation of the device is done using the right hand rule, considering the z-axis perpendicular to the mobile device's surface and oriented downwards, the y-axis is aligned to the width and oriented to the right side, the x-axis is aligned to the length of the device and oriented upwards, where the start button resides. The axis system is considered fixed to the mobile device.

This method is based on measurements in static conditions, with the accelerometer at rest and with the read axis oriented vertically. One hundred discrete values are taken from the axis that is calibrated. After the values have been collected, the device is rotated 90 degrees and the measurements are repeated for the next axis. The process is repeated until the device has done a full rotation.

By using values in the acquired measurement database, an average and standard deviation of the samples is determined for positive and negative values of acceleration.

$$\overline{a_x} = \frac{\sum_{j=1}^{n} a_{xj}}{n} \tag{5}$$

where a_x represents the arithmetic mean of acceleration values for the x-axis, a_{xj} are values of the x-axis acceleration, and n represents the number of discrete values of the digital signal for acceleration.

Standard deviation of the sample is defined:

$$\sigma_x = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(a_{xj} - \overline{a_x})^2} \tag{6}$$

It is a measure of signal quality; it describes the absolute value of deviation from the mean value.

For each axis, the "two point formula" is used to determine the slope and the intercept, considering two points of coordinates (mean value for negative x-axis, negative value of gravitational acceleration). By applying the two point formula, one equation for each axis results in the form of:

$$y = mx + n \tag{7}$$

where m is the slope and n is the intercept.

Slope is determined with the formula:

$$m = \frac{Y_2 - Y_1}{\overline{a_{x+}} - \overline{a_{x-}}} \tag{8}$$

where m is the slope, Y_2=9.80678 is the value for local gravitational acceleration, $Y1 = -Y_2$, a_{x+} and a_{x-} are mean values of sample acceleration for positive x-axis and negative x-axis, respectively. The slope gives a value of the scale between the compensated and not compensated signal.

The intercept n aligns the mean value of acceleration to the nominal value of acceleration, resulting in a signal that is in the ±9.80678(m/s^2) interval.

$$n = m\overline{a_{x+}} - Y_2 \tag{9}$$

The corrected value of the x-axis acceleration, after calibration, is: $a_{xcor} - a_{xe}\ (m/s^2)$

$$a_{xcal1} = ma_x - n \tag{10}$$

Value for acceleration at horizontal orientation is given correctly from the accelerometer, but after calibration, an error is introduced because of the n parameter. Also, positive and negative semi-periods have different amplitudes. For this reasons, the method of calibration will be applied separately on each semi-period, one of the two points will be considered the origin of the axis.

After calibration, the values for ±90 degrees are identical to the Euler determined values. This is a necessary condition for calibration, but it is not sufficient. After further examination, it seems the calibrated signal still presents a difference relative to the ideal signal. Other measurement values for this method of calibration in semi-periods will be presented in the next subchapter.

5.2. Gravitational Calibration in Quarter-Periods

The δ_x difference from the ideal value has a double sinusoidal shape when it is represented as a function of orientation angle θ, as presented in figure 9.

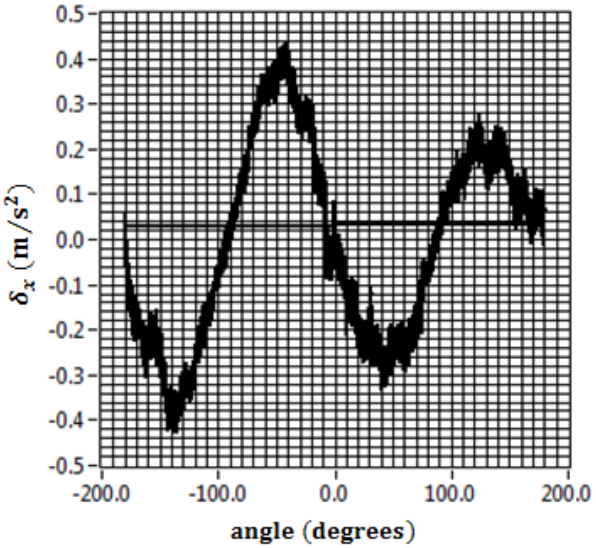


Figure 9: Shape of the Semi-Period Calibrated Gravitational Acceleration Deviation from the Euler Determined Gravity Component

$$\delta_x = a_{xcal1} - a_{xe} \tag{11}$$

Readings are done for δ_x at ±45 and ±135 degree values. An additional calibration signal is defined:

$$a_{xcal2} = \delta_{xqp}\sin(2\theta), \tag{12}$$

Where

$$\delta_{xqp} = \begin{cases} \delta_{x-135}, -180 < \theta \leq -90; \\ -\delta_{x-45}, -90 < \theta \leq 0\,; \\ \delta_{x+45}, 0 < \theta \leq 90\,; \\ -\delta_{x+135}, 90 < \theta \leq 180. \end{cases} \tag{13}$$

$$a_{xcal} = a_{xcal1} - a_{xcal2} \tag{14}$$

Values are compared for the uncompensated signal and the compensated signals in the following table:

Table 2. Values For X-Axis Gravity Components at Different Angles and Calibration Methods

θ [degrees]	-135	-45	45	135
$\overline{a_x}$ (m/s^2)	6.56	7.24	-7.32	-6.73
$\overline{a_{xcal1}}$ (m/s^2)	6.70	7.35	-7.26	-6.66
$\overline{a_{xcal}}$ (m/s^2)	6.92	6.93	-6.91	-6.94
$\overline{a_{xe}}$ (m/s^2)	6.93	6.94	-6.93	-6.93

where $\overline{a_x}$ is the mean value of the uncompensated gravitational acceleration, $\overline{a_{xsp}}$ is the mean compensated value for semi-periods, $\overline{a_{xds}}$ is the mean double sinus value, $\overline{a_{xe}}$ is the mean Euler derived acceleration value. Experimental measurements have been done for ±45 and ±135 degrees θ angle orientation.

After the signal is calibrated at ±45 and ±135 grade, it has a ±0.01(m/s^2) error of the calibration points and overall the error is ±0.15(m/s^2) as seen in figure 10. Large variations of error for a particular angle value are due to the fact that measurements were taken by manually rotating the accelerometer device and inertial acceleration has been added to the signal.

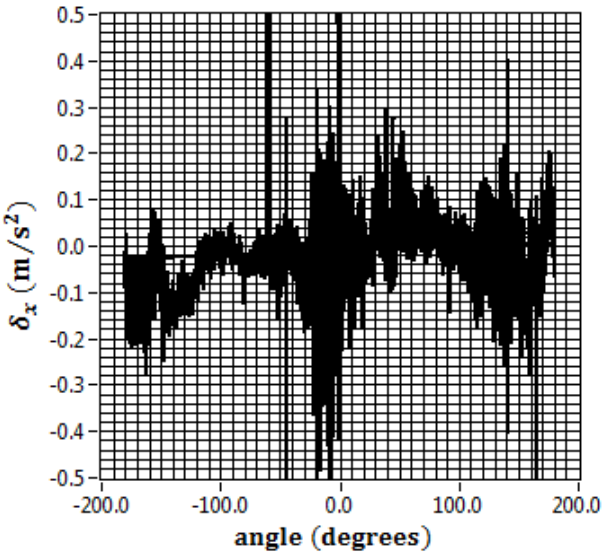


Figure 10: Shape of the Quarter-Period Calibrated Gravitational Acceleration Deviation.

5.3. Determining Inertial Acceleration

The Euler angles are a set of three angles that describe a number of three consecutive rotations around predefined axes. The mobile device that contains an accelerometer, also gives Euler angle values for an *XYZ* type rotation. It has a particular condition, which is that the final orientation is generated using only the first two rotations, meaning the rotation around *z* is considered null.

Rotation matrices describe rotation around a particular axis. By multiplying them, we get a generalized rotation matrix that is comprised of a rotation around *x* by φ degrees, then around *y* by θ degrees and lastly around *z* by ψ degrees.

$$R_{XYZ}=R_x(\psi)R_y(\theta)R_z(\varphi) \tag{15}$$

The generalized form is:

$$R_{XYZ} = \begin{bmatrix} R_{00} & R_{01} & R_{02} \\ R_{10} & R_{11} & R_{12} \\ R_{20} & R_{21} & R_{22} \end{bmatrix} \quad (16)$$

where

$$\begin{cases} R_{00} = \cos\theta\cos\varphi \\ R_{01} = -\cos\theta\sin\varphi \\ R_{02} = \sin\theta \\ R_{10} = \cos\psi\sin\varphi + \cos\varphi\sin\psi\sin\theta \\ R_{11} = \cos\psi\cos\varphi - \sin\psi\sin\theta\sin\varphi \\ R_{12} = -\cos\theta\sin\psi \\ R_{20} = \sin\psi\sin\varphi - \cos\psi\cos\varphi\sin\theta \\ R_{21} = \cos\varphi\sin\psi + \cos\psi\sin\theta\sin\varphi \\ R_{22} = \cos\psi\cos\theta \end{cases} \quad (17)$$

By multiplying the rotation matrix with the gravitational vector, we get the projections of the gravitational acceleration on the three axes.

$$R_{XYZ} \times \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} = \begin{bmatrix} g_x \\ g_y \\ g_z \end{bmatrix} \quad (18)$$

Now that the gravitational acceleration has been determined, we apply the difference between the accelerometer readings and gravitational acceleration to obtain pure inertial acceleration. This is also called gravity compensation.

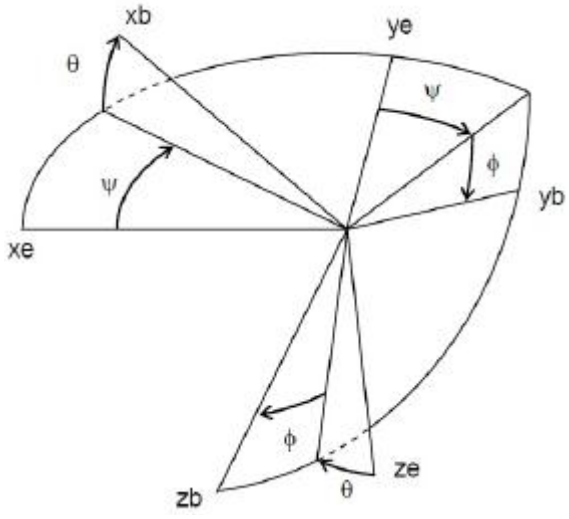


Figure 11: Euler Angles for a Fixed Earth and a Fixed Body Frame of Reference

Prior to compensation, the accelerometer is calibrated for latitude and altitude influences (Boynton, 2001), also the amplitude is calibrated using the two point formula:

$$a_{xcal} = ma_x - n \quad (19)$$

where *m* is the slope and *n* is the intercept.

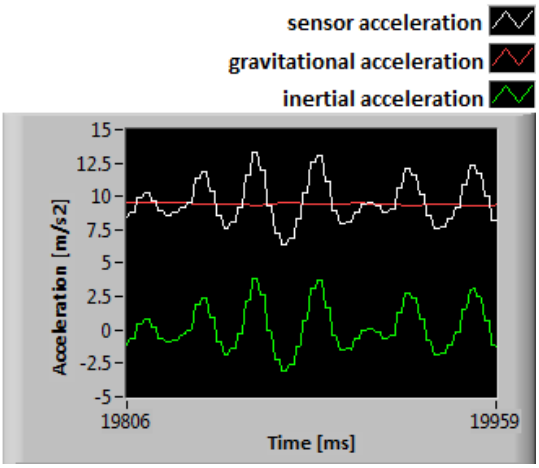


Figure 12: Sensor Acceleration in the *X* Direction, Contains Raw Sensor Data (Combined Acceleration); Gravitational Component and Inertial Component

In figure 12, the mobile device was oriented with the *x* axis oriented perpendicular to the ground and pointing downward. The variations in acceleration are caused by moving the phone up and down along its x axis. *Sensor acceleration* is the raw data from the sensor; *gravitational acceleration* has a constant value as long as the sensor moves parallel to the ground plane; *inertial acceleration* or linear acceleration along one axis is the acceleration generated by movement.

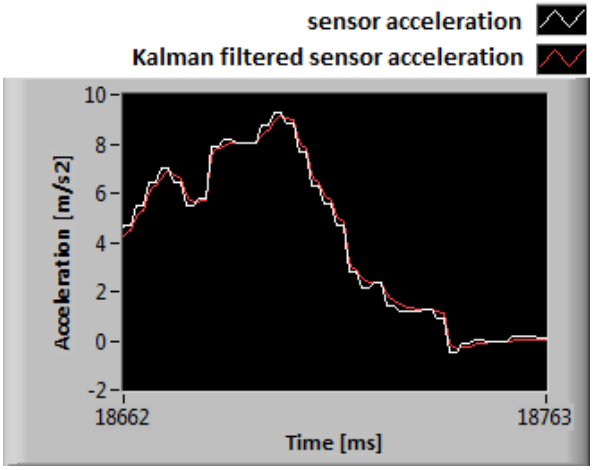
5.4. Signal Noise Cancelation

To measure the robot state, i.e. the position, velocity or acceleration, sensors are used. Measurement is done by sampling data every period of time; measurements include errors determined by alignment, sensitivity, noise.

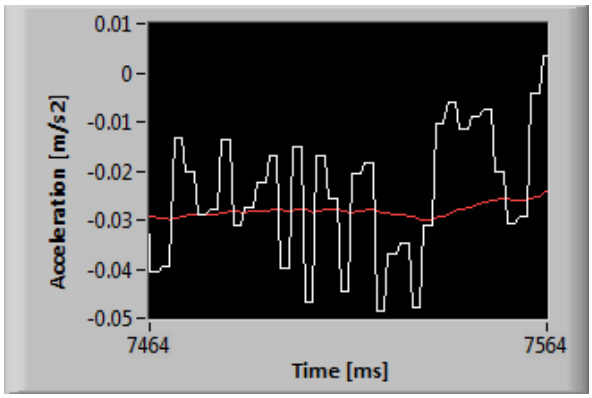
Low-pass filters give signals a smooth form, by removing short-term fluctuations and providing the desired long-term component.

For the sake of comparison, five filters have been tested: Kalman (Bizup 2004), Median, Butterworth, Chebyshev and Bessel. With the exception of the Kalman filter, the experiment results show they have little effect on small variation, the amplitude of the error remains the same and the filters introduce lag, probably from the extra time it takes to calculate the new values from raw data. The Kalman Filter proved to be the best at removing noise error, reducing it by an order of magnitude.

The Kalman filter is used to smooth the signal by removing noise. Small changes of linear acceleration in the cm/s^2 range are indistinguishable from the sensor error, so these are filtered out also. It is observable that the noise error is reduced by a factor of ten. Large variations, in the order of units of m/s^2 are not filtered out.



(A)



(B)

Figure 13: Measured Sensor Acceleration (m/s^2) and Kalman Filtered Data for Large (A) and Small (B) Variations

5.5. SPEED BY INTEGRATION

By integrating the linear acceleration along the *x* axis, linear speed results:

$$\int [a_x(t) + err_{n.x}(t)]\, dt = v_x(t) + coeff_{drift.x}(t) \quad (20)$$

An integration constant, C, is also added to the second term of the equation.

Integration of the acceleration samples is done point by point, dividing two small variations in one period of time. Last acceleration value is given at the beginning of the new period and measured again in the current. Subtracting the last value from the previous gives us a measure of acceleration in one period of time. Times values are also subtracted, present time from last period time. Shift registers are used for this purpose.

Initial conditions are v_{x0}=0 and t_0=0.

$$a_{x1} = v_{x1}/t_1 \quad (21)$$

$$a_{x2} = v_{x2} - v_{x1}/t_2 - t_1 \quad (22)$$

$$a_{xj} = (v_{xj} - v_{xj-1})/(t_j - t_{j-1}), for\ j \epsilon N \geq 1 \quad (23)$$

From (15), the value for v_{xj} is extracted:

$$v_{xj} = a_{xj}(t_j - t_{j-1}) + v_{xj-1} \quad (24)$$

$$v_{xj} = \Delta v_{xj} + v_{xj-1} \quad (25)$$

The raw inertial acceleration signal contains measurement errors in the range of ±0.02(m/s^2).

Drift is observed in figure 14 and is caused by integration of read errors $err_{n.x}$ (10). Although only speed measurements are required for our purpose, by further integrating the speed signal, the position can be obtained. The drift would further be amplified and the simulated signal value would drift quickly from the real value.

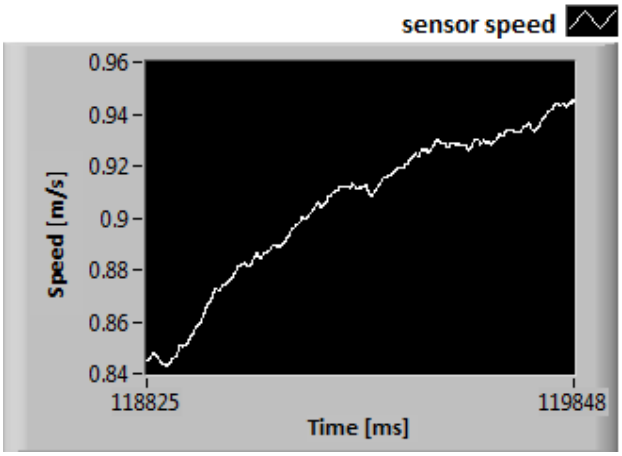


Figure 14: Sensor Speed Data Obtained by Integrating the Sensor Acceleration Data

6. WHEEL SLIP MEASUREMENT APPLICATION

Linear speed is obtained both from the sensor and encoders: the first one, by integrating the sensor linear acceleration and the second one from derivation of wheel motor position. In Figure 15 the vehicle accelerates until a constant speed of negative 0.3(rad/s) is achieved. Note that this is not reverse movement; the negative value is due to inverse position of the right motor with respect to the left motor; same shaft axis, opposite rotational direction.

Wheel slip is determined from equation (1), vehicle speed is determined from inertial acceleration integration and wheel speed is obtained from the digital encoders mounted at each motor axis.

The first period of positive wheel slip for the vehicle lasts for 15(ms), until the 65545(ms) mark: λ=0.8 The second one is shorter, only lasts 5(ms) until the 65550(ms) mark with larger slip: λ=1.5. Negative slip (wheel block) is observed from 65550 to 65560(ms), λ=–0.24. After this point in time, slip stabilizes to a value close to zero.

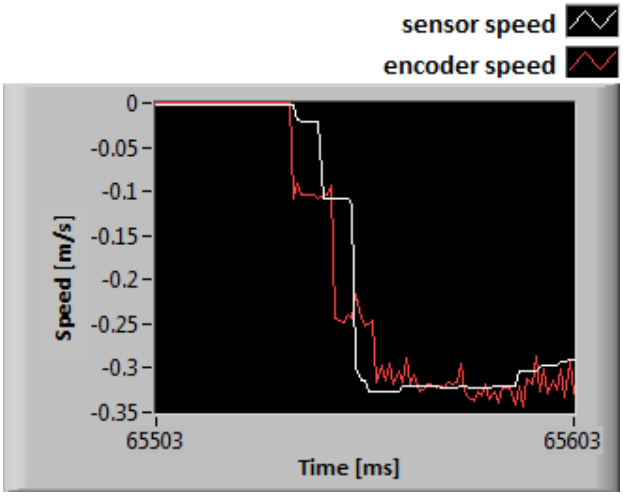


Figure 15: Measured Sensor and Encoder Speed. Acceleration Phase is Shown.

Using the wheel slip data in a closed loop, wheel control can be obtained. Depending on the position of the slipping wheel, corrections can be made to the slipping wheel or any of the other wheels to correct the behavior and regain grip.

7. CONCLUSIONS

Theoretical absolute value for gravitational acceleration is calibrated for latitude and altitude, signal amplitude is calibrated in semi-periods and further calibration is done in quarter-periods.

Gravitational component is determined using Euler angles. It then must be removed from raw acceleration sensor data to obtain inertial acceleration.

Noise errors accumulate during integration, the Kalman filter reduces noise by a factor of ten.

Inertial acceleration data is transformed into vehicle speed data and compared to wheel speed to measure wheel slip.

Slip is very small in value and time for the Robotics Starter Kit. To better visualize the wheel slip and the effect of the slip control, changes are proposed. In the future, the slip coefficient value will be forced to rise by changing the road surface to one with low adhesion coefficient and by changing the wheel tire surface with a more slippery one, like plastic instead of rubber.

REFERENCES

Bizup, D.F., 2004. A generalized acceleration model for Kalman filter trackers. *Proceedings of the Thirty-Sixth Southeastern Symposium on System Theory*, pp. 31-35, September 27, Atlanta, Georgia, USA.

Boynton, R., 2001. Precise Measurement of Mass, Proceedings of the 60th Annual Conference of the Society of Allied Weight Engineers, pp.4-8, May 21-23, Arlington, Texas, USA.

Bräunl, T., 2003. Embedded Robotics, Berlin Heidelberg: Springer

Cai, Z., Ma, C., Zhao, Q., 2010. Acceleration-to-torque Ratio based Anti-skid Control for Electric Vehicles, *Proceedings of IEEE/ASME International Conference on Mechatronics and Embedded Systems and Applications (MESA)*, pp 577-581, July 15-17, Qingdao, ShanDong, China.

Hong, D., Yoon, P., Kang, H.J., Hwang, I., Huh, K., 2006. Wheel Slip Control Systems Utilizing the Estimated Tire Force. *Proceedings of the American Control Conference*, pp. 1-6. June 14-16, Minneapolis, Minnesota, USA.

Hori, Y., Toyoda, Y., Tsuruoka Y., 1997. Traction Control of Electrical Vehicle based on the Estimation of Road Surface Condition – Basic Experimental Results using the Test EV 'UOT Electric March'. *Proceedings of the Power Conversion Conference*, pp. 1-7, August 3-6 Nagaoka, Japan.

AUTHORS BIOGRAPHY

Szöcs Daniel is an electrical engineer and third year PhD candidate at the Technical University of Cluj-Napoca, Department of Electrical Machines and Drives. Research for the doctoral thesis was made at National Instruments Romania, the hardware basis for the research is the National Instrument Robotics Starter Kit. Research topics: direct-drive electrical vehicles, anti-slip control.

Fenești Andrei is an electrical engineer and second year PhD candidate. Research topics: direct-drive electrical vehicles, obstacle avoidance, trajectory control with LADAR and GPS technology.

Pană Teodor is head of the Department of Electrical Machines and Drives and professor of Microprocessor and Computer Systems, as well as two other digital control based courses at the Technical University of Cluj-Napoca, Faculty of Electrical Engineering and Research topics: vector control systems, electric vehicle drive control, industrial robots.

Vese Ioana is an electrical engineer and PhD at the Technical University of Cluj-Napoca, Department of Electrical Machines and Drives**.** Research topics: modeling and simulation of electrical machines, design, modeling and control of linear motors and actuators, design and control of small electric motors.

VISUALIZATION IN BUSINESS PROCESS SIMULATION

Xiaoming Du[(a)], Terrence Finandor [(b)], Kuzhen Wu [(c)], Jialiang Yao [(d)]

School of the Built Environment (Thinklab)
University of Salford, Manchester, UK

[(a)]x.du@salford.ac.uk, [(b)]fernando @salford.ac.uk, [(c)]k.c.wu @salford.ac.uk, [(d)]j.yao @salford.ac.uk

ABSTRACT

Visualization in Business Process Simulation (BPS) studies techniques and methods for graphically representing abstract business concepts and data set which are produced in simulation design, execution and analysis. Its main goal is to enhance, simplify and clarify the understanding of BPS. Although data/ information visualization has been actively studied, current research is lack of complete and systemic description on how visualization is used in BPS. This paper focuses on the visualization techniques in BPS. Firstly, the procedure of BPS is defined and the role of visualization in BPS is analyzed. Then, three typical visualization techniques are summarized. According to the BPS steps, different visualization usages in different BPS phase are illustrated. Finally, challenges of enhancing the impact of visualization in BPS are concluded.

Keywords: business process improvement, business process simulation; information visualization; process visualization

1. INTRODUCTION

Simulation is thought of as a key technique for business processes improvement. The technique of using simulation in this context is referred to as Business Process Simulation (BPS). It can be used to test decisions prior to their implementation in real business environment. Furthermore, it allows for the integration of variability and uncertainty into the anticipation of business process performance (April, Better, Glover, Kelly, and Laguna 2006). Such capabilities are often aided by BPS software tools which not only provide users with a variety of analysis possibilities based on the key simulation performance metrics, but also with a visual interface to aid decision making (Bradley, Browne, Jackson, and Jagdev 1995; Vullers and Jetjes 2006). Such user-friendly visual interface for process modeling, simulation execution and result analysis is seen as key criteria for enhancing the usability of BPS tools.

There are many literatures concerning visualization techniques which can be categorized into two major groups: "scientific visualization" and "information visualization" (Tory and Moller 2004). However, most these papers are related to general ideas or concrete domain-specific applications (Streit, Pham, and Brown 2005; Aouad, Ormerod, and Sun 2000; Steel and Iliinsky 2000). Current research is lack of complete and systemic description on how visualization is used in BPS. Visualization of BPS can be regarded as a kind of information visualization. It interprets abstract or behavioral data into visual image that represents as an analogy or metaphor in the problem space.

In this paper, we focus on the visualization techniques in BPS. Our goal is to review the visualization methods and point out its current and potential usage in BPS. To achieve this goal, we firstly define BPS procedures, then summarize the visualization techniques and analyze its application patterns in BPS as a whole. Finally, the challenges of enhancing the impact of visualization in BPS were presented.

2. PROCEDURE OF BPS

Tumay (1996) provides an overview of BPS research work. Regarding the simulation implementation, some basic steps can be distinguished. Figure 1 essentially illustrates four steps in conducting BPS: 1) choosing and defining business, 2) building business model, 3) running business model and 4) analyzing performance and making decision.

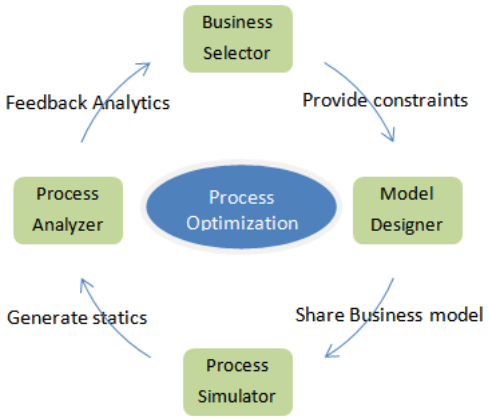


Figure 1: Implementation Steps of BPS

Before starting simulation, the real business process should be chosen and mapped into a process model,

supplemented with business process scenarios and necessary documents. In this context, the sub processes and activities are identified; the control flow is created by entities and connectors which abstractly show the system structure and execution order; the resources are assigned to the activities where they are necessary. In addition, the performance indicators, such as throughput time and resource utilization, also need to be defined.

In order to statistically obtain valid simulation results, typically, business models are executed for many times during the simulation. A simulation run should consist of multiple sub-runs which require substantial time in order to get a valid result. The simulation tool may show an animated picture of the process flow or real-time fluctuations of the key performance measures. After simulation is finished, the simulation results can be collected to aid drawing useful and correct conclusions. Meanwhile, statistical data analysis and visualization should be performed to assist user in decision making. The feedback of analysis result can bring optimization to the real business process.

3. VISUALIZATION TECHNIQUES USED IN BPS

The main advantage of BPS is its ability to incorporate the variability and inter-dependence factors in order to obtain and compare the different process performance. However, such capability simultaneously require an efficient visual interface to support stakeholders who need a communication channel to manipulate complex data, validate the model and evaluate the operation of (re)designed business process in a fast and effective way.

Visualization is "the process of representing data as a visual image (Latham 1995)". In the context of BPS, the data is normally composed of business resources, activities, constraints, running information, results output etc. The purpose of the visualization in BPS can be categorized as validation, analysis and marketing (Bijl 2009; Balci 1997). Visualization can be regarded as a dynamic validation technique because displaying animation of a model and comparing it with the real life operation can help user to identify discrepancies between the model and the real system. Visualized analysis refers to the process of gathering information, displaying 2D&3D graphs and drawing conclusions from the simulation. Consequently, it can be used to locate business process bottlenecks. The third term, marketing, refers to increasing the confidence of people with the help of attractive and understandable graphs. This means convincing users of the validity of the simulation result.

Various visualization techniques play an important role in business knowledge transformation from text-based data to visual characteristics, patterns and tendencies. They appear to be simple, intuitive and natural, break the barrier between the business stakeholders and the knowledge of a specific subject domain. According to appearance style in BPS, we groups visualization technique into three categories:

3.1. Static Graph

Static graph is used to transform an initial representation of a data structure into a non-action graphical one. Though the graph can be visually examined and interacted with, the corresponding graph has no variation itself. The BPS analysis result can be shown as a bar or pie diagram. In term of vision perception, they can be 2D, 3D or multi-dimensional (Adams 2012).

During business process design phase, model could be described as a static flow chart, which consists of nodes (with different node types) connected by edges (of different edge types) (Owen and Raj 2012). In combination with node and edge labels, automatically arranging process graphs and reducing edge crossings to a minimum are complex tasks (Bobrik and Reichert 2005). Some commercial workflow management systems, like MQ Workflow, Lotus Workflow, Staffware, or Oracle Process Manager can support process model visualization (Leymann and Roller 2000). One flaw of above approaches is that model conceptual design and model execution monitoring should depend on the same platform. Process model cannot be visualized by different simulation engines. Other shortcoming is the poor options offered for customizing the way while process is being visualized.

As for the static visualization of analysis data, no matter original input data or result output data, it is mainly provided by automated graphing tools in the form of lines, boxes, arrows, various symbols and pictograms. Such tools could be independent such as MS Graph, Excel or integrated in a BPS tool suite. A scale and labels are also common. The elements of a graphic do not have to be an exact or realistic representation of the data, but can be a simplified version. Spatial layout and graph-drawing algorithms play a fundamental role in visualization.

3.2. Dynamic Animation

While static visualizations of business data are still valuable for the objectives of many studies, there is an increasing demand for dynamic, interactive visualization capabilities to facilitate the validation and understanding of complex business processes.

Dynamic (i.e. time-dependent) animation might help a viewer work through the logic behind an idea by showing the intermediate steps and transitions, or show how data changed over period of time. A moving image might offer a fresh perspective, or invite users to look deeper into the data presented. BPS should have its ability to capture and visualize the dynamic behavior of a process. There are two dynamic aspects, when implementing visualization in BPS, these needs to be addressed (Greasley 2003):

- Variability. Most business processes contain variability both in the demand on the system (e.g. customer arrivals) and in durations of processes (e.g. customer service times). The simulation permits the incorporation of

statistical distributions and thus BPS visualization should provide an indication of such dynamic behavior of the process.

- Inter-dependence. Most processes contain a number of decision points that affect the overall performance of the system. The simulation technique can incorporate various input data to model the likely decision options taken. Also the "knock-on" effect of such interdependent decisions can be dynamically assessed and shown over a time period.

Two application purposes of animation are presented: exploration and presentation (Steele and Iliinsky 2000). Some forms of animation are most suited to presentation, while others work well for exploration. It also discusses a hierarchy of different types of animation which shown as Table 1, ranging from changing the view to changing the axes. Since such classification is not especially designed for BPS, some subtypes may be not a real sense of dynamic visualization (i.e. subtype 1). In this paper, dynamic visualization of BPS is mainly for exploration which covers the second, third and sixth subtype.

Table 1: Type of Animation

No	Type	Function
1	Change the view	Pan over or zoom in on a fixed image, such as a map or a large data space
2	Change the charting surface	On a plot, change the axes (e.g., change from linear to log scale). On a map, change from, for example, a Mercator projection to a globe.
3	Filter the data	Remove data points from the current view following a particular selection criterion.
4	Reorder the data	Change the order of points (e.g., alphabetize a series of columns).
5	Change the representation	Change from a bar chart to a pie chart; change the layout of a graph; change the colors of nodes.
6	Change the data	Move data forward through a time step, modify the data, or change the values portrayed (e.g., a bar chart might change from Profits to Losses).

3.3. Virtual Environment

Although the quality of information visualization has been improved substantially due to the advanced 3D representation, there are several information visualization systems that were implemented in a virtual environment which gives user a feeling of being immersed in real-time 3D world (Kirner and Martins 2000). Virtual Environments (VE) provides unique advantages for information visualization such as "near-real-time" interaction and response (Bryson 2002). As a result, VE tends to be a more attractive and intuitive way in BPS visualization.

Within BPS context, VE based visualization particularly represents the transformation of abstract business data, normally shown as 2D graphs or spreadsheets, into 3D geometry which can be displayed as stereoscopic image and interactively manipulated by the user immersed within the simulation in term of visual experiences. Combined with VE technology, BPS visualization could integrate abstract visualization objects with real-world 3D spatial business sceneries, and support "near-real-time" interaction and response. This visualization metaphor can be used in most of data analysis domains such as tourist attractions, electricity consumption, water supply, crime distribution, spending power etc in which data to be analyzed can be organized by geographical location (Burdea and Coiffet 2003).

These three visualization techniques in BPS above can be categorized in term of user perception and interaction level. From static graph to virtual reality, it gets more complex, striking and intuitive for the user. In addition, advanced visualization techniques such as 3D and dynamic interaction can be used in any of above three visualization forms with different extent of application. However, advance technique does not mean better for problem solving. More suitable techniques need to be chosen according to the business domain characteristics.

4. VISUALIZATION APPLICATION IN BPS

Based on the procedure of BPS denoted in section two, different visualization techniques can be integrated into each steps of BPS.

4.1. Visualization in Process Modelling

The modeling of business processes has a significant role in BPS. Notably, business process modeling demonstrates two important functions (Lin, Yang, and Pai 2002): (1) to capture existing processes by structurally representing their activities and related elements; and (2) to represent new processes in order to evaluate their performance.

Business Process Modeling Notation (BPMN) is the new standard for graphically modeling business processes, which was defined by the Business Process Management Initiative (Owen and Raj 2003). The competing standard to BPML is the Business Process Execution Language for Web Services (BPEL4WS) created in a joint venture by BEA, IBM, Microsoft and others. Such notation and language are sets of graphical constructs and rules about how to combine these constructs. In addition, process modeling tools such as Microsoft Visio, IBM/Rational Rose, Arena etc can provide user and analyst with the ability to draw business processes. However, such tools only provide a 2D graphical static model editor making use of the appointed shapes such as rectangles, circles, and arcs etc which conform to these modeling grammars.

Modeling languages which use only 2D diagram for visualization limit the amount of information to be integrated into a process model in an understandable way. Betz, Eichhorn, Hickl, Klink, Koschmider, Li, Oberweis, and Trunko (2008) introduce a 3D technique into business process model which represents

information more compactly. In addition, 3D process model visualization allows user to change view-point interactively. As a result, this approach not only improves the layout of process models (e.g. by minimizing the number of crossings of arcs), but also increases the information content of a process model.

Recently, collaborative process modeling has appeared under some web-based experimental environments. In some circumstances such as large scale manufacturing system, modeling needs to be performed in a cross-organizational, distributed environment. Brown, Recker, and West (2011) describe a novel process modeling approach using recent 3D virtual world technology. Even without the support of 3D technology, some visualization tools can be extended to provide a complete and usable distributed environment for collaborative (re-) design of business processes. This approach increases user empowerment and adds significance to the collaboration and consensual development of process models even when the relevant users are geographically dispersed.

4.2. Visualization in Simulator Running

While modeling allows user to visualize the design of a business process, process running could offers a fresh perspective to look deeper into the data presented. There have been different approaches for visualizing the behavior of a simulator running in an understandable way for the relevant stakeholders.

Dynamic and interactive visualization techniques are mostly adopted during the execution of the business simulation. Similar to algorithm animation (Korhonen and Malmi 2002), the state change of data structure can be visualized at certain forms. For example, process simulation can be animated by continuously displaying the process state with a sequence of visual snapshots of data structure. In addition, user can control the process by interacting with system such as stopping or continuing the animation. Moreover, these animation steps could be recorded in order to give user the control of traversing the animation sequence back and forth. The replay or animation of the simulation shows the states of process model which might be used to reveal the bottleneck of process design (Vullers and Netjes 2006).

There are many commercial-off-the-Shelf (COST) tools, such as ARIS, Casewise, ExSpect, having sophisticated simulation capacity and animation function to support the visualization of time-based generation of process event. For example, Verbeek (2000) uses the concept of dashboards to track the key performance metrics' changes and reflect the dynamic behavior of a business process (e.g. by using flow meters, flashing lights, etc. as used in typical control panels). Moreover, interaction with the simulation via this dashboard is available. Comparing to the dashboard, Petri nets tend to support process visualization by showing the movement of graphical objects. Mimic library of Design/CPN supports the simulation system to manipulate graphical objects, allows user to interact with simulation via such graphical objects. As a result, user can get a good impression of the 'look and feel' of the final product. Other examples include the model of mobile phone communication and a graph transformation based animation such as GenGED (Kindler and Pa 1es 2004).

One recent example of process visualization, GapMinder, is an animated bubble chart designed to show trends over time in three dimensions. Both size and locations of bubbles smoothly animate as time passes. This technique appears to be very effective in presentations, where a presenter tells the observer where to focus by making the data come to life, and emphasizes the critical results of an analysis.

Since 3D visualizations can offer intuitive understanding of business processes simulations that every stakeholder can engage with easily, Brown and Cliquet (2008) develop a tool which can be used to provide 3D avatar-based visualizations, showing an avatar executing a workflow process in a 3D virtual environment. Thus, a workflow in principle can be designed within a standard 2D tool and then be visualized in a 3D environment for communication and validation processes. With 3D scene and human models, a typical service-based process can dynamically illustrated. Eichhorn, Koschmider, Li, Oberweis, St ürzel, and Trunko (2009) present an interactive 3D process animation based on Petri nets which support users to quickly identify the weak spots of the real business processes.

4.3. Visualization in Process Analyzing

In BPS procedures, process analyzing is considered as a post-simulation step after simulation has been completed. Apart from the static presentation and dynamic animation of business model, the simulation result of model execution should be appropriately harnessed and presented in order to aid decision making.

Static visualization technique is widely used in this phase. Of course, besides the traditional 2D form of tables, outlines, pie charts, line graphs, and bar charts, advanced information visualization such as multi-dimensional graphics has been used increasingly. The main purpose of the visualization in process analyzing is to support the last decision making after simulation result had been obtained. The result of BPS is different from other types of data which is normally abstract, discrete, multi-dimensional, hierarchical and networked. These characteristics make visualization more difficult (Tegarden 1999).

Although rich static graph is widely used for process analyzing, dynamic animation and visual reality technologies are getting more acceptable in some particular situation such as geography related business simulation in which the process analyzing is illustrated using 3D and animation dashboard along with 2D line graph and different color indication. Tegarden (1999) also presents some typical visual representations created by information visualization designers which include Kiviat diagrams, parallel coordinates, 3D scattergram,

3D line graph, volume rendering, floors & walls representation, maps and so on.

5. CHANLLENGES IN ENHANCING THE IMPACT OF VISUALIZATION IN BPS

Visualization apparently offers many benefits to BPS stakeholders. Academics, researchers and designers are continually striving to formulate new and creative methods for representing. However, they could face a number of challenges and choices summarized below.

5.1. Usability and Accuracy

Human factors play an important part in BPS visualization design. Creating an effective tool for BPS visualization should require a deep understanding of domain-specific problems and tasks (i.e. key performance metrics, metaphor designs etc). As most software developers, BPS developers also tend to do their development as quickly as. possible. Nevertheless, many efforts just focus on creating new visualization techniques with little attention on user needs and capabilities.

One common criticism of BPS visualization research is that it presents interesting techniques, rather than solutions to real problems. With new technology, the power and novelty of techniques can lead to using technology for technology's sake instead of providing real benefits to users and providing more information through visualization (Stone 2009). Few's Web site shows some examples of inadequate business visualization which require developers to further understand the data, the audience and the problem being solved (Few 2012).

5.2. User Interaction

Interaction in BPS visualization is essential for the users to conduct analytical reasoning, gain insight from complex data-set. BPS visualization tool has to provide not only effective visual representations but also effective user interaction to ease the exploration and help users to achieve understanding.

Due to the incorporation of advanced technology such as 3D, VE and Web into BPS, the user interaction becomes more complex. For example, in VE based BPS visualization, interaction with abstract data element and virtual environment requires a certain training and familiarity with nonconventional devices. On the other hand, another challenge for the interaction in BPS is to be able to dynamically build up visualization views along with certain criteria which may be specified via a direct manipulation interface. This technique involves graph reduction and graph aggregation which comprise choosing process objects and composing them in an appropriate way (Rinderle, Bobrik, and Reichert 2006). Furthermore, distributed user interaction in BPS visualization via Internet provides another challenge (Huang, Xiong, and Li 2004).

5.3. Collaborative Visualisation

Visualization is the key support during collaborative business process improvement in which visualization is evolving into mediators of human-to-human and human-to-data interaction. Current BPS visualization is user-centered and task (process) driven, which supports multi-disciplinary teams of users to collaborate and share ideas. However, how to design such digital information visualization systems that can adequately enhance collaborative business process improvement is still remain as a challenge.

Some exploring research work has being done in this area. For example, process modeling is a complex organizational task that requires many iterations and communication between business analysts and domain specialists, some state-of-the-art 3D virtual environments were developed with that in mind for collaborative (re-) design of business processes (West, Brown, and Recker 2010; Brown, Recker and West 2011). During the development of collaborative visualization of BPS, social interaction and information security could be the main challenges for researchers.

6. CONCLUSION

BPS visualization is a specialized area of information visualization, which focuses on improving business process understanding by providing visual representation of abstract business concepts and data set. No such a very common and proper visualization technology can be used over each phase of BPS and each application area. Visualization techniques need to be carefully chosen which conform to the problems and tasks of a particular domain. With the development of modern computer technology, new visualization approaches for solving business problems can go further. Though some challenges will be confronted with, visualization of BPS will continuously play an indispensable role in business process improvement.

REFERENCES

Adams, D., 2012. *Data Visualization*, Available from: www.CFOProject.com.

Aouad, G., Ormerod, M., and Sun, M., 2000. "Visualisation of construction information: a process view", *International Journal of Computer-integrated Design and Construction*, 2(4):pp206-214.

April, J., Better, M., Glover, F., Kelly, J. and Laguna, M., 2006. Enhancing business process management with simulation optimization, *Proceedings of the 38th Conference on Winter Simulation,* pp.642-649. December 03-06, Monterey, CA, USA.

Balci, O., 1997. Chapter Verification, Validation and Testing, in: *Handbook of Simulation. Principles, Methodology, Advances, Application and Practice*, John Wiley and Sons, Inc. 335–393.

Betz, S., Eichhorn, D., Hickl, S., Klink, S., Koschmider, A., Li, Y., Oberweis, A., and Trunko, R., 2008. 3D Representation of Business Process Models, *Modellierung betrieblicher Informationssysteme– Online Proceedings* (*MobIS 2008)*, pp.73-87, November 27-28, Saarbrücken, Germany.

Bijl, J. L., 2009. *How game technology can be used to improve Simulations. literature survey*, Delft University of Technology, 2009.

Bobrik R., and Reichert, M., 2005. Requirements for the visualization of system-spanning business processes, *Proceedings of the 16th International Workshop on Database and Expert Systems Applications.* pp.948-954, August 22-26, Copenhagen, Denmark.

Bradley, P., Browne, J., Jackson S., and Jagdev H., 1995. Business process reengineering (BPR) – A study of the software tools currently available, *Computers in Industry*, 25(3):309-330.

Brown, R., and Cliquet, F., 2008. Communication of business process models via virtual environment simulations, *BPTrends*, 12(9):1-7.

Brown, R., Recker, J., and West, S., 2011. *Using Virtual Worlds for Collaborative Business Process Modeling*, Available from: http://eprints. qut.edu.au/41816 [accessed 20 March 2012]

Bryson, S. , 2002. Information Visualization in Virtual Environments, in: Stanney, K., ed. *Handbook of Virtual Environments*, LEA Associates.

Burdea, G. C., and Coiffet, P., 2003. *Virtual Reality Technology*, 2nd Edition, Wiley-IEEE Press.

Eichhorn, D., Koschmider, A., Li, Y., Oberweis, A., Stürzel, P., and Trunko, R., 2009. 3D support for business process simulation, *33rd Annual IEEE International Computer Software and Applications Conference*, pp.73-80. July 20-24, Seattle, Washington.

Few, S., 2012. Available from: http://www. perceptualedge.com/examples.php [accessed 20 March 2012]

Greasley, A., 2003. Using business-process simulation within a business-process reengineering approach, *Business Process Management*, 9(4): 408-420.

Huang, B., Xiong D., and Li, H., 2004. An Integrated Approach to Real-time Environmental Simulation and Visualization, *Journal of Environmental Informatics*, 3(1):42-50.

Kindler, E., and Païles, C., 2004. 3D-Visualization of Petri Net Models: Concept and Realization, *25th International Conference on Application and Theory of Petri Nets*, pp.464-473, June 21-25, Bologna, Italy.

Kirner, T.G. and Martins, V. F., 2000. Development of an Information Visualization Tool Using virtual Reality, *Proceedings of the 15th ACM Symposium on Applyed Computing*, pp.604-607, March 19-21, Como, Italy.

Korhonen, A., and Malmi L., 2002. MATRIX – concept animation and algorithm simulation system, *Proceedings of the Working Conference on Advanced Visual Interfaces. ACM*, pp.109-144. May, Trento, Italy

Latham, R., 1995. *The Dictionary of Computer Graphics and Virtual Reality*, 2nd Ed. New York, NY, Springer-Verlag.

Leymann, F., and Roller, D., 2000. *Production Workflow – Concepts and Techniques*, Prentice Hall.

Lin, F., Yang M., and Pai, Y., 2002. A generic structure for business process modeling, *Business Process Management*, 8(1):19-41.

Owen, M., and Raj, J., 2003. *Popkin Software, BPMN and Business Process Management*, Available from: http://www.BPMI.org [accessed 20 March 2012]

Rinderle, S., Bobrik, R., and Reichert, M., 2006. Business Process visualization – use case, challenges, solution, *8th International Conference on Enterprise Information Systems*, pp.205-211, May 23-27, Paphos, Cyprus.

Steele, J., and Iliinsky, N., 2000. Animation for Visualization-Opportunity and Drawbacks, in: *Beautiful Visualization, O'Reilly Media*, Inc. 329-352.

Stone, M., 2009. *Information Visualization: Challenge for the Humanities*, Available from: http://www.clir.org/pubs/resources [accessed 20 March 2012]

Streit, A., Pham, B., and Brown, R., 2005. Visualization Support for Managing Large Business Process Specifications, *Proceedings of third International Conference on Business Process Management*, pp.206-219. September 6-8, Nancy, France.

Tegarden, D., 1999. Business Information Visualization, *Communications of AIS* , 1(4).

Tory, M., and Moller, T., 2004. Rethinking visualization: A high level taxonomy, *Proceedings of Information Visualization*, pp. 151–158. October 10-12, Austin, Texas.

Tumay, K., 1996. Business process simulation, *Proceedings of the 1996 Winter Simulation Conference*, pp.93–98, ACM Press, Coronado.

Verbeek, E., 2000. ExSpect 6.4x product infromation, *21st International Conference on Application and Theory of Petri Nets*, pp.39–41, June 26-30, Aarhus, Denmark.

Vullers, M. J., and Netjes, M., 2006. Business process simulation – a tool survey, *Proceedings of the Seventh Workshop and Tutorial on the Practical Use of Coloured Petri Nets and the CPN Tools*, pp.77-96, October 24-26, Department of Computer Science, University of Aarhus.

West, S., Brown, R., and Recker, J., 2010. Collaborative business process modeling using 3D virtual environments, *16th Americas Conference on Information Systems Sustainable IT Collaboration around the Globe*, August 12-15, Swissôtel, Lima.

LUMPED PARAMETERS MODELLING OF THE FURNACE AND STEAM SYSTEMS OF A 350 MW BOILER

Edgardo J. Roldan-Villasana[(a)], Ma. Cardoso-G. [(b)], Jose A. Tavira-Mondragon[(c)] Miguel Rossano[(d)]

Advanced Systems on Training and Simulation Department
Instituto de Investigaciones Electricas
62440 Cuernavaca, Mor., Mexico
web page: http://www.iie.org.mx

[(a)]eroldan@iie.org.mx, [(b)]mcardoso@iie.org.mx, [(c)]jatavira@iie.org.mx, [(d)]rossano@iie.org.mx

ABSTRACT
The models of the combustion and steam of a boiler real time simulator are presented in this paper. The simulator reproduces the actual behaviour of a steam generator of 350 MW power plant (a dual thermal power station currently operating in Mexico). The simulator is part of an analysis tool that reproduces the response of variables such as temperature, pressure, mass flow and composition of the combustion gases in different parts the boiler under diverse operating conditions. The models were developed using a concentrated parameters approach. The simulator response was validated with extensive tests. Here, the models responses under a couple of transients are discussed.

Keywords: boiler modelling; boiler simulation; combustion model; lumped parameters

1. INTRODUCTION
The present work was developed in the Advanced Systems on Training and Simulation Department (GSACyS, after its name in Spanish) of the Institute of Electrical Research (IIE) in Mexico. This is the second part of a previous paper (Roldan-Villasana et al. 2010a) where the simulator architecture, the software platform, the simplified models, and development of the waterwalls modelling were fully explained.

The GSACyS has developed several real time simulators for operators training of power plants for the Federal Electricity Commission (CFE, the Mexican Utility Company), a complete summary of these developments have been reported elsewhere (Roldan-Villasana et al. 2010b).

Due to the high incidence of problems of broken pipes in the steam generators, caused mainly by high temperatures in the different components of the boilers, the CFE requested the development of an analysis boiler simulator in order to be able to identify and predict potential failures.

This analysis tool is called boiler simulator with distributed parameters, and allows analysis of internal phenomena such as heat transfer, flow paths and combustion products, by using the commercial package FLUENT (Galindo-García, et al. 2010, Vázquez-Barragan et al. 2010).

The aim of this simulator is to provide a set of initial conditions from where the boiler simulator with distributed parameters, may initiate an analysis simulation.

2. LUMPED PARAMETERS SIMULATOR
The boiler simulator with lumped parameters can simulate steady state conditions, change variables values, and simulate malfunctions. The objective is to operate the boiler in any condition to reach a point where the analysis simulator may be initiated. A lumped parameter approach with both, storage and resistive modules was used (Colonna and van Putten 2007).

The lumped parameters approach simplifies the description of the behaviour of spatially distributed physical systems into a topology consisting of discrete entities that approximate the behaviour of the distributed system under certain assumptions (basically, a point has the properties of a finite volume). From a Mathematical point of view, the simplification reduces the state space of the system to a finite number, and the partial differential equations of the continuous (infinite-dimensional) time and space model of the physical system into ordinary differential equations with a finite number of parameters

The simulator session is guided through a simulation console having the main functions explained in next section.

3. SIMULATOR INTERFACES

3.1. Simulation Scenarios Interface
The user may interact with the simulator through a graphic interface (Fig. 1) where some specific functions are available. The interface is designed to guide a simulation session (the hardware/software configuration has been described in the previous part of this paper). The main available functions are:

Figure 1: Simulation Scenarios Interface

Run/Freeze. The user may start or freeze a dynamic simulation test.

Simulation Speed. Normally, the simulator is executed in real time, but user may execute the simulator up to ten times faster or ten times slower than real time.

Initial Condition. The instructor may select an initial condition to begin the simulation session. It is possible to record new initial conditions (snapshots) or to erase old initial conditions, and to specify the time interval to get the automatic snap-shooting function (up to 500 of them).

Malfunctions. It is used to introduce/remove a simulated failure of equipments.

External Parameters. They allow the instructor to modify the external conditions like atmospheric pressure and temperature (dry and wet bulb).

Distributed Parameters Simulation. Switch the program to simulate the boiler furnace with the commercial tool FLUENT.

Development Tools. The simulator has a series of implements helpful especially during the development stages: simulate just one step (1/10s), monitor and change on line the value of selected global variables, and tabulate any list of variables.

In particular, the malfunctions are a very important part of the simulator. They are defined as abnormal events associated with equipment failures. The simulator includes failures, such as trip of forced circulation pumps; trip of the fans of induced draft; malfunctions of the coal mills; rupture of the waterwalls, superheater, reheater and economizer pipelines; fouling of regenerative air heaters; trip of the soot blowers, trip of the primary air fan; and trip of the induced draft fans.

3.2. Operation Interfaces

The user may interact with the simulated plant using the operator interface. They are graphic displays that reproduce real operation consoles in power plant control rooms. Figs. 2 and 3 present two examples of operation screens where the user interacts with the simulator by acting, for example, on pumps, valves and controls, and may verify the operative parameters.

4. GENERAL SIMULATOR MODELS

The lumped parameters approach was used for the models development. It simplifies the description of the behaviour of spatially distributed physical systems into a topology consisting of discrete entities that approximate the behaviour of the distributed system under certain assumptions (basically, a point has the properties of a finite volume). Mathematically, the simplification reduces the state space of the system to a finite number, and the partial differential equations of the continuous (infinite-dimensional) time and space model of the physical system into ordinary differential equations with a finite number of parameters.

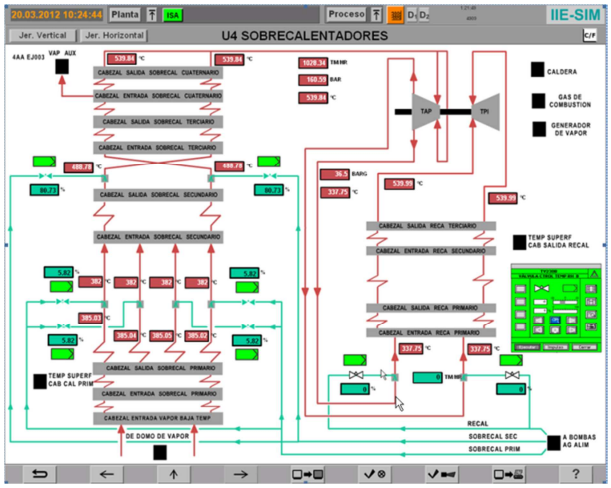


Figure 2: Superheaters

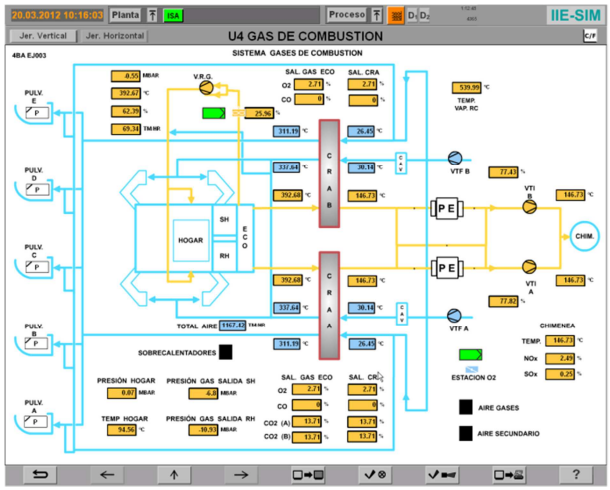


Figure 3: Gas Combustion System

The models were divided into three main parts: Simplified models; boiler part air and gases; and water and steam in the boiler.

The simplified models simulate the data required by the boiler using basic equations, but always accomplishing the first principles. These systems are: Fuel (Oil and Coal); Air and Combustion Gases; and Turbine and Auxiliary Steam.

Other systems such as lubrication oil, control oil

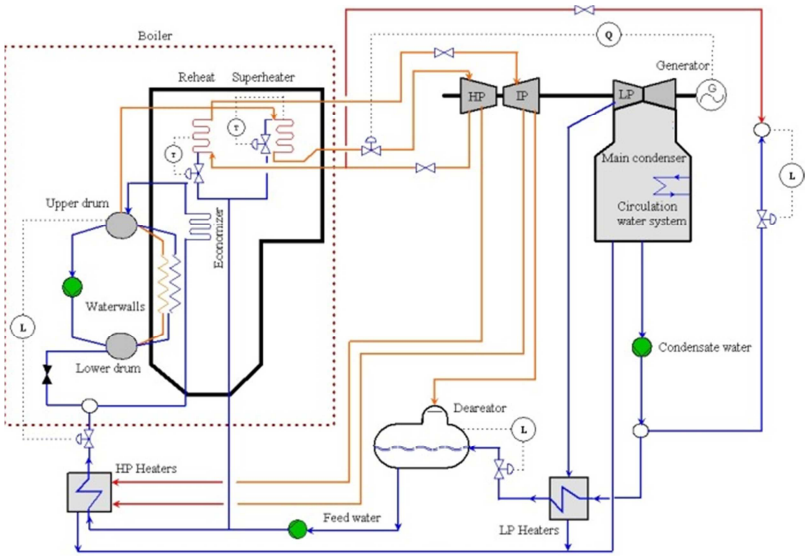


Figure 4: Simplified Diagram of the Simulated Systems

cooling water, air for instruments and electrical system are considered to always operate properly.

A basic diagram of the simulated plant is shown in Fig. 4. In this paper, the models of the boiler air and gases part and steam are presented. All others were explained by Roldan-Villasana et al. (2010a).

4.1. Causal Interaction of Models

For modelling purposes, the boiler has the boundary elements presented in Table 1. Fig. 5 is a schematic of the main interaction variables, where *h* is enthalpy, *S* is control signal, *T* is temperature, *L* is level, *P* is pressure, *w* is mass flowrate, c_i is fuel components, and *q* is heat.

Table 1. Interaction with other (simplified) models

Element	Frontier System
Economizer	Feed water
Deheaters (Super and Reheaters)	Feed water
Burners	Fuel System
Inlet Air Dampers	Combustion Air System
Fourth Superheater	High Pressure Turbine
Third Reheater	Intermediate Pressure Turbine
Boiler Waterwalls and Drum	Atmosphere
Boiler	Controls

5. GENERAL MODELLING

The momentum, mass and energy equations apply for all the simulated equipment and systems.

The mass balance is:

$$\frac{dm}{dt} = \sum w_i - \sum w_o \qquad (1)$$

Sub-indices *i* and *o* are the properties at the inlet and output, respectively. No mass accumulation exists for incompressible flowrates (water and liquid fuel). All flowrates through a resistive element are calculated as:

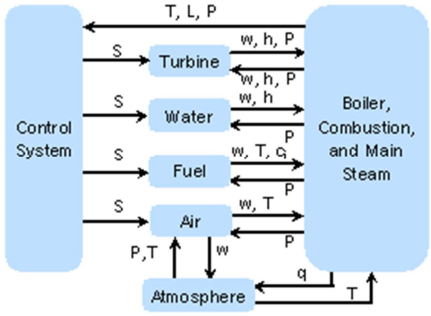


Figure 5: Boiler Main Variables Causal Diagram

$$w^2 = k\, Ap^{\gamma}\, \rho\, (\Delta P + \rho g \Delta z) \qquad (2)$$

Where *k* is the friction coefficient, *Ap* the valve aperture (if this exist in the line), ρ is density, *g* the gravity acceleration and Δz the height difference between the inlet and the output, and γ is the valve characteristic.

For pumps and fans and any other rotate equipment, the flowrate is:

$$\Delta P = \frac{K_1\, w^2}{\rho} + K_2\, w\, \omega + K_3\, \rho\, \omega^2 - \rho g \Delta z \qquad (3)$$

There, ω is the speed and K_1, K_2, and K_3 are constants.

A proper energy equation is established depending of the particular equipment as summarised later.

All the model equations form a set of differential equations and a non-linear system of algebraic equations, both solved numerically. The differential equations are integrated with the Euler method with an integration step of 0.1 s. The algebraic equations are solved with a Newton-Rapson method with relaxation factor.

The thermodynamic and transport properties are calculated as a function of pressure and enthalpy. The data source was the steam tables by Arnold (1967). The functions were adjusted by least square method. The

application range of the functions is between 0.1 *psia* and 4520 *psia* for pressure, and -10 ^0C and 720 ^0C (equivalent to 0.18 *BTU/lb* and 1635 *BTU/lb* of enthalpy). The adjustment was performed to assure a maximum error of 1% respecting the reference data; to achieve this it was necessary to divide the region into 14 pressure zones. The functions are applied to three different cases: subcooled liquid saturated and superheated steam.

6. SIMPLIFIED MODELS AND CONTROLS

6.1. Simplified Models (Boundaries)

These are models that simulate the input data required by the boiler. The approach is to state simple equations, but following first principles.

The fuel (oil and coal) system supplies any type of fuel. It is ready to go into operation when required. The fuel flow is calculated depending on the demand of the control. The boiler has four burner elevations for both, fuel and carbon, and the flowrate depends on the number of elevations in operation. The fuel properties are considered constant.

The air system serves to provide air for the combustion in the furnace. No heaters are modelled, thus, the air temperature is a function of the main steam flowrate with an empirical equation.

Turbine is driven by the steam generated in the boiler and it is directly connected to the electrical generator for power generation. The high pressure and low pressure turbines were simulated in a simplified way. The output conditions of the turbines are obtained by calculating the enthalpy at isentropic conditions and correcting it considering a constant efficiency.

The auxiliary steam is a system that provides steam to the turbine seals to prevent entry of air or steam leak in the cylinder of the turbine.

6.2. Controls

A full distributed control system is not necessary for this simulator. All control models are simplified in both parts analogical and logical.

PI controls where considered. They include the derivations of superheater and reheater steam, superheater and reheater temperatures (deheat water), fuel demand, excess air for combustion, and gas recirculation. The signal control is, where K_P is proportional constant, ε error, and τ integral time constant:

$$S = K_I\,\varepsilon + \frac{1}{\tau}\int d\varepsilon \tag{4}$$

As a simplification, the fuel and excess air set points were considered to be a function of the generated electric power.

The logical control was simplified, the start/stop sequence of burners, for example, does not check for the pilot ignition or if the fuel valve is open. A perfect behaviour was assumed for the simplified systems.

7. BOILER STEAM MODEL

The steam produced by the waterwalls flows through the superheater, the high pressure turbine, and then to the reheater and the intermediate and low pressure turbines to be discharged into the condenser.

Both, the superheater and the reheater are finned tubes taking heat from the combustion gases to be transformed in mechanical energy in the turbines to produce electricity.

Small nodes (for example, nodes where the feed-water deheat flow discharges) are considered resistive and their pressure is calculated with the equations set produced by (1) to (3). Their energy balance is:

$$\frac{dh}{dt} = \frac{\sum w_i\,(h_i - h) + q_m}{m} \tag{5}$$

Where the mass m is considered constant in these cases and just represents the inertia of the node. Heat q_m is transferred from the pipeline metal (subindex m) that depends on physical and operative conditions and metal temperature:

$$\frac{dT_m}{dt} = \frac{\sum q}{(m_m\ \ Cp_m)} \tag{6}$$

Being Cp the heat capacity and the heat sum the difference between the heat flow from the furnace and to the fluid.

For big nodes (like headers) two state equations are established for the enthalpy and pressure. Different equations apply according the state of the fluid (Roldan-Villasana, E and Verduzco Bravo, A, 2009).

For saturated conditions, the equations applied on a volume V are:

$$\frac{dh}{dt} = \frac{1}{\left(\frac{\partial\rho}{\partial P}\right)_h \rho + \left(\frac{\partial\rho}{\partial h}\right)_P} \left\{ \begin{array}{l} \frac{\left(\frac{\partial\rho}{\partial P}\right)_h}{V}\left[\sum (wh)_i - \sum (wh)_o - \sum q\right] \\ + \left[1 - \left(\frac{\partial\rho}{\partial P}\right)_h h\right]\frac{d\rho}{dt} \end{array} \right\} \tag{7}$$

$$\frac{dP}{dt} = \frac{1}{\left[\frac{\partial\rho}{\partial P}\right]_h}\left(\frac{d\rho}{dt} - \left[\frac{\partial\rho}{\partial h}\right]_P \frac{dh}{dt}\right) \tag{8}$$

For saturated steam:

$$\frac{dh}{dt} = \frac{dP}{dt}\left\{ \begin{array}{l} \left[x\left(\frac{dh_g}{dP}\right) + (1-x)\left(\frac{dh_f}{dP}\right)\right] + \\ \frac{(h_g - h_f)}{(v_g - v_f)}\left[\frac{x}{\rho_g^2}\left(\frac{d\rho_g}{dP}\right) + \frac{(1-x)}{\rho_f^2}\frac{d\rho_f}{dP}\right] \end{array} \right\} - \frac{(h_g - h_f)}{(v_g - v_f)}\left[\frac{1}{\rho}\left(\frac{\sum w_i - \sum w_o}{V}\right)\right] \tag{9}$$

$$\frac{dP}{dt} = \rho \frac{dh}{dt} + \left[\frac{\sum w_i \left(\sum h_o - \sum h_i \right) - \sum q}{V} \right] \quad (10)$$

Here, the subindexes *f* and *g* represent condition of saturated gas and saturated liquid, respectively and *x* is the vapour quality.

8. FURNACE

The model simulates the combustion in the furnace and calculates the heat transfer from the flue gas to the economizer, superheater, reheater, waterwalls, regenerative air heaters and electrostatic precipitators. The preheated air from the air/steam heater and the Ljunström regenerative heater is mixed with the fuel and recycled gas in the furnace. After the combustion, a fan located at the exit of the boiler, induce the flow of gases to the chimney.

In summary, the model is based on the solution of the combustion stoichiometric equation considering as reactive agents C, H_2, O_2, H_2O, S, N_2, CO_2, CO, and SO_2. The products depend on the molar flow of air and fuel and the composition of the fuel. Molar flow for each combustion agent considers the contribution flow of fuel or carbon, air to burners, air to pilots, atomization air, steam on breakage of tubes from waterfalls, steam soot blowers, and recirculated gas.

There is not taken into account the kinetics of combustion; it is considered a perfect mixing in the furnace, and it is stated that the combustion gases follow the ideal gases law. A chemical balance is obtained to calculate the excess air in the furnace.

The emission of carbon monoxide in the combustion depends on excess air: more excess air, lower amount of CO and vice versa. CO emissions also depend on the flow regime, i.e. less turbulence decreases efficiency of mixing and CO emissions increase due inefficient combustion.

The flame temperature, as a function of the formation enthalpies, is used as the basis for calculating the heat transfer in each area of the gas path, from the furnace up to the exhaust.

The furnace (*fu*) conditions are obtained with two equations for the derivatives of the state variables (pressure *P* and temperature *T*):

$$\frac{dP_{fu}}{dt} = R T_{fu} \frac{d\rho_{fu}}{dt} + R \rho_{fu} \frac{dT_{fu}}{dt} \quad (11)$$

$$\Sigma(w\,h)_i - \Sigma(w\,h)_o + \Sigma q = V_{fu} \left(\rho_{fu} Cp_{fu} \frac{dT_{fu}}{dt} + h_{fu} \frac{d\rho_{fu}}{dt} - \frac{dP_{fu}}{dt} \right) \quad (12)$$

Where *R* is the gas ideal constant, *V* is the volume, *Cp* is the heat capacity, and *q* is the heat flow.

Note that density is not a state variable and its derivative is used only as a value for the solution of (11) and (12). The sum of the heat flow refers to heat transferred by radiation and convection to the waterwalls.

After the furnace, the combustion gases heat other fluids and metals in the steam generator: upper sections of the waterwalls, superheater, reheater, economizer, regenerative air heaters and electrostatic precipitators.

If along the path of the gases a series of control volumes are established (each occupied by a specific heat transfer equipment), assuming that there in not accumulation of gases at any point in his path, with a balance of energy may be calculated the outlet temperature of each control volume. In all cases, heat transfer coefficients were calculated with the help of correlations found in the literature.

The heat transfer along the circuit water/steam considers the thermal resistance of the metals.

9. TESTS AND RESULTS

9.1. Proposed Test

One test was proposed to be presented in this paper in order to the general scope of the models. However, during the acceptance tests, the simulator was probed in all the normal operation range, from 25% to full charge, including the response under malfunctions and abnormal operation procedures. The simulator variables were compared with the plant data at every steady state (25%, 50%, 75%, and 100% of load). In all cases the response satisfied the existing ISA S77.20 (1993).

The starting point for this test was the full load condition (350 MW). The oil fuel consumption is 21.3 kg seg^{-1} and the produced steam is 285 kg seg^{-1}. The user made no actions on the simulator during the test. In Figs. 6 and 7 the response of some variables is presented.

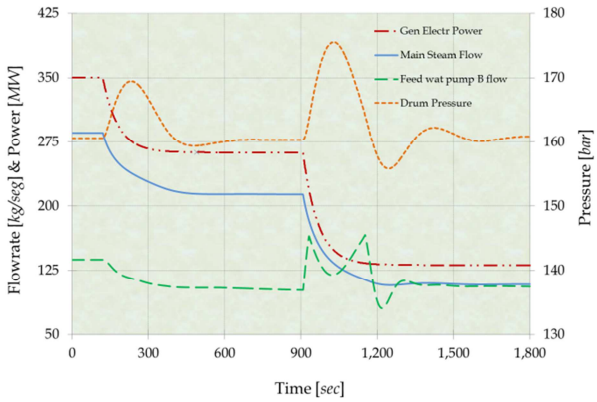


Figure 6: Flowrates, Electrical Power and Pressure

After 120 s of steady state, the user, by means of the proper control, ask to download the load to 75%. The system is stabilised by the control approximately at 500 s and then at 900 s, a feedwater pump is trip.

9.2. Graph Results

During the first part of the test, the electrical power (Fig. 6) is gradually changed from 100 % to 75% during 260 s approximately. As a response of the turbine admission valves, the main steam flow descends continuously.

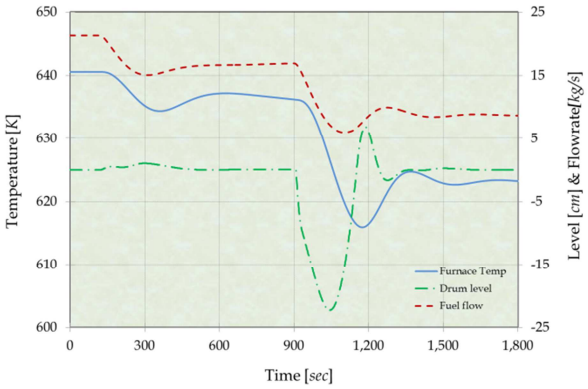


Figure 7: Temperature, Level and Flowrate

The feedwater flow also descends following the drum level control (the level remains practically stable).

The change in the electrical load origins a notorious response in the drum pressure caused by reduction in the consumption of steam.

The fuel flowrate goes down trying to control the changing electrical power. Due the control parameters, when the charge finally stabilises, the fuel flowrate presents a small oscillation. The furnace temperature, as expected, moves practically with the same tendency that the fuel flowrate.

Once the transient is controlled, the pressure returns to its initial value. At 600 s the plant is stable at 75% of electrical load.

For the second part, at 900 s, one of the feedwater pumps is trip. This action is detected by the control and a runback is activated (the charge demand is automatically set at 50%). The electrical power descents smoothly.

The main steam flow declines its value gradually following the aperture of the turbine valves that partially close due the runback.

Initially, the feedwater pump B flow increases its flow when the pump A is trip (they are connected in parallel). Then, the flowrate changes according the control valve aperture trying to stabilize the drum level that presents, as expected, the stronger response (note that the total feedwater flow decreases during the transient).

The drum pressure initially increases its value due the diminution of the main steam flowrate and then it is affected by the changes in the feedwater flowrate and the temperature in the furnace.

Again, the fuel flow is adjusted according the value of the electrical load, and a mild oscillation is presented. The furnace temperature follows the fuel flowrate but a delay may be appreciated mainly due to the thermal inertia of the boiler.

At 1500 s approximately, the plant is stable again in its new state (50%).

10. CONCLUSION

The presented boiler simulator is a replica, high-fidelity, plant specific simulator. Realism is provided by the use of detailed modelling. The lumped parameter approach is adequate for the purposes of the simulation that was tested and validated by the client according plant data and international norms. The tests that were followed to validate the modes were performed in the defined operation range of the plant and were successful. This makes the simulator a reliable tool for analysis purposes.

REFERENCES

Arnold, E. 1967. *Steam tables: thermodynamic properties of water and steam, viscosity of water and steam, thermal conductivity of water and steam*, Electrical Research Association, London.

Colonna P., van Putten H. 2007. "Dynamic Modeling of Steam Power Cycles. Part I - Modeling Paradigm and Validation", *Applied Thermal Engineering*, 27, 467-480.

Galindo-García, I.F, Vázquez-Barragan, A.K, Rossano-Román, M, 2010. Numerical Study of the Flow and Temperature Distribution in a 350 MW Utility Boiler. *Proceedings of ASME, Power Conference, POWER2010*, pp. 767-775. July 13-15, Chicago (Illinois, USA).

ISA S77.20, 1993. *Fossil-Fuel Power Plant Simulators Functional Requirements*.

Roldán-Villasana, E.J, Cardoso-Goroztieta, Ma. J, Tavira Mondragón J.A, Rossano Román M.B, 2010a. Lumped Parameters Modelling of the Waterwalls of a Power Plant Steam Generator, *European Modelling Symposium, Fourth UKSim European Symposium on Computer Modelling and Simulation*, UK Simulation Society, pp 283-288. November 283-288. Pisa (Tuscany, Italy).

Roldán-Villasana, E.J, Mendoza, Cardoso M, Jimenez Sanchez, V and Cruz-Cruz, R, 2010b. *Gas Turbine Power Plant Modelling for Operating Power Plant*, eds. *Gas Turbine* . Ijeti Gurrapa. India, 169-214.

Roldan-Villasana, E.J and Verduzco Bravo, A, 2009. *Diseño e Implementación de los Modelos de la Caldera. Modelo de Vapor Recalentado con Parámetros Concentrados.* Reporte interno, Gerencia de Simulación, IIE. Cuernavaca, (Morelos, México)

Vázquez-Barragan, A.K, Galindo-García, I.F, Mani-González A.G, Rossano–Román, M, 2010. Simulación CFD, una alternativa para el análisis de emisiones contaminantes en calderas de plantas termoeléctricas. *VIII Congreso Internacional sobre Innovación y Desarrollo Tecnológico CIINDET 2010*, pp. 693-698. Noviembre 23-26, Cuernavaca (Morelos, México).

CFD SIMULATIONS AS A TOOL FOR FLOW AND THERMAL ANALYSIS IN BOILERS OF POWER PLANTS

Iván F. Galindo-García[(a)], Ana K. Vázquez-Barragán[(b)], Miguel Rossano-Román[(c)]

[(a), (b), (c)]Advanced Training Systems and Simulation Department
Institute of Electrical Research
62490, Cuernavaca, México

[(a)]igalindo@iie.org.mx, [(b)]akvb@iie.org.mx, [(c)]rossano@iie.org.mx

ABSTRACT
Computational Fluid Dynamics (CFD) simulations of the gas flow inside the boiler of a power plant are presented. The CFD simulations can be employed for a very detailed analysis where spatial and local effects can be important. CFD calculations were performed for a 350 MW utility boiler at 100% of total load using either pulverized coal or heavy oil as fuels. Then, two case studies with variation in working conditions are presented: for the first case the effect of the amount of combustion air is investigated and for the second case the failure of one burner is simulated. Simulation of these test cases demonstrate the general capability of the simulator and that CFD methods are recommended as a viable computational tool to evaluate the flow and thermal performance in the gas side of the boiler of a power plant.

Keywords: CFD, boiler modelling, power plant

1. INTRODUCTION

A computational model has been developed in order to analyse flow, temperature and species distributions for the gas side of a 350 MW utility boiler. It is assumed that identification of high temperature or high velocity zones will help in the prevention of failures in the boiler walls tubes, superheaters, reheaters and economizers. Failures such as stress rupture, erosion, and thermal fatigue are associated with short- and long-term overheating, fly ash, coal particle impingement, and overfiring or uneven firing of boiler fuel burners. Identifying and correcting the cause of these failures is essential to ensure low availability loss and to eliminate repeat failures. Careful attention has been given to assemble tools that utilities need to eliminate these failures (Dooley and Chang 2000).

One of these tools is numerical modelling, and CFD methods provide a potentially accurate and cost effective tool that can help in the analysis of the gas side of a boiler. Modelling of the chemical and physical phenomena inside the boiler is important for the analysis of failures, because one of the known causes of tube failure is the non-uniform heating of the tubes, which strongly depends on flow and temperature distributions of the combustion gases. CFD is becoming a critical part of the design process of different power plant equipment. During the last 20 years CFD has been applied to study pulverized coal combustion in furnaces to predict the combustion phenomena and to troubleshoot flow, mixing, combustion, and heat transfer problems (Boyd and Kent 1986, Fiveland and Wessel 1988). However, even if in recent years great progress has been achieved, the predictions of CFD models for combustion should be considered as qualitative trends and for parametric analysis (OIT 2002). Nevertheless, careful use of these codes as an engineering tool can help to obtain a reliable prediction of the combustion behaviour of utility boilers.

In the present work computer simulations have been performed to model steady state, 3-D combustion for a 350MW utility boiler burning either pulverized coal or heavy oil. The main aim is to develop a computational tool to investigate flow, temperature and species distributions within the gas side of a boiler that can help to predict zones with abnormal operation for a particular set of operation conditions. The CFD model is part of a simulation tool that integrates the current CFD model with a real time lumped-parameter module of the boiler including all associated systems and controls (feedwater system, steam turbine, controllers, etc.), which permits the user to dynamically simulate different operational conditions and to establish a particular condition to be simulated with the CFD code. In Roldan-Villasana et al. (2010) the real time simulator architecture, the software platform, and mathematical models are fully explained.

This work has been developed at the Advanced Training Systems and Simulation Department (GSACyS, after its name in Spanish) of the Institute of Electrical Research (IIE) in Mexico. The GSACyS has developed several real time simulators for training of power plant operators for the state-owned utility in Mexico (Federal Electricity Commission CFE). For the present work CFE requested the development of a simulation tool that can provide a deeper insight into the physical and chemical phenomena in a power plant boiler in order to be able to identify and predict potential boiler tube failures.

2. GENERAL DESCRIPTION OF THE BOILER

A boiler, or steam generator, is a key component in power plants. The boiler extracts energy from the gases product of the combustion, and transfers thermal energy to the water/steam that flows inside multiple sections of heat exchanger tubes (superheater, reheater and

economizer). The steam produced in the boiler is then supplied to a steam turbine to generate power.

The boiler under consideration is part of a 350 MW commercial power plant operating in a subcritical steam cycle. The combustion chamber is rectangular in shape (dimensions 12.7 x 14.15 x 45.6 m), and is fired tangentially using five levels of pulverized coal burners or four levels heavy oil fuel burners in each corner. The furnace geometry and burner arrangement are shown in Fig. 1.

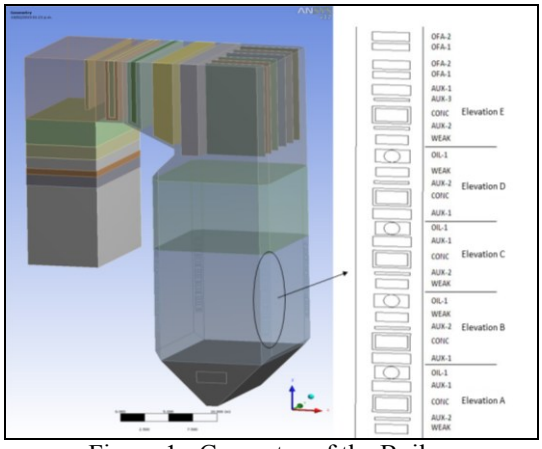


Figure 1: Geometry of the Boiler.

3. CFD MODEL

Four stages have been followed for the development of the CFD model: 1. Building the geometry, 2. Mesh generation, 3. CFD simulation and, 4. Post processing and analysis of results. The simulations in the present work were done using the general purpose CFD software "ANSYS FLUENT" (ANSYS 2009a).

3.1. Geometry

The solution of any CFD process begins with the generation of the geometry. Technical drawings of the reference boiler have been consulted to generate the computational 3D geometry that represents the actual equipment as closely as possible. The model domain consists of the combustion chamber from the burner nozzles at the furnace corners, up to the exit of the economizer.

3.2. Meshing

The accuracy of a CFD simulation depends on the quality of the mesh. Numerical error is a combination of many aspects, for example the grid density, discretization method, and convergence errors (Ferziger and Peric 2002). Numerical error can be minimized using denser grids, higher order discretization methods and suitable time step size. The limitation for these factors is computation time, as time required to get a converged solution for a CFD problem depends directly on the size of the mesh. However in all CFD computations results should be ensured to be grid independent. Generally, it is important to find an optimum between acceptable results and computational time.

In this work a mesh sensitivity analysis was performed in order to evaluate the effect of the mesh on calculations. Beginning with a coarse mesh, simulations have been carried out for different mesh sizes. Target quantities (temperature and velocity) have been obtained as a function of the grid density. The final result of the calculations should be independent of the grid that is used. This is usually done by comparing results of calculations on grids with different grid sizes. Fig. 2 shows the velocity in a horizontal line inside the boiler using different grids sizes.

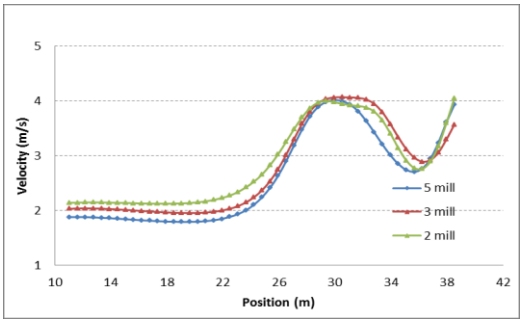


Figure 2: Velocity for Three Different Mesh Sizes.

The computational mesh adopted for these calculations consist of tetrahedral and hexahedral elements and has approximately 3 million elements of unequal size. The regions close to the burners were assigned a denser mesh. Figure 3 shows two views of the mesh.

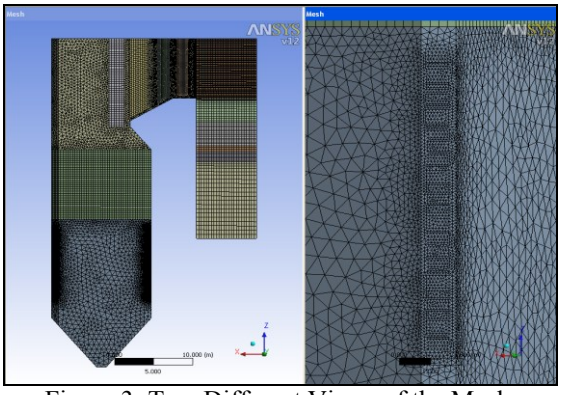


Figure 3: Two Different Views of the Mesh.

3.3. CFD Simulation

The CFD simulation involves defining the mathematical models and establishing the boundary conditions for the problem to solve.

The simulation of combustion systems includes modelling a number of complex, simultaneous, interdependent processes such as fluid flow, turbulence, particle transport, combustion and radiation. The time averaged conservation equations (mass, momentum and energy) are solved for predicting the flow, temperature and concentration of gas species. Turbulent quantities are calculated using the standard high-Reynolds-number k-ε turbulence model. Standard wall functions are used to bridge the regions adjacent to solid boundaries, the forms adopted taking $k^{1/2}$ as the velocity scale (ANSYS,

2009b). Lagrangian particle trajectories of the pulverized coal particles or heavy oil droplets are calculated throughout the computational domain. The dispersion of particles due to gas turbulence is predicted using the stochastic tracking model which includes the effect of instantaneous turbulent velocity fluctuations of the gas on the particle trajectories. The P1 radiation model is used to simulate radiation heat transfer. Absorption coefficients of the gas phase are calculated using the weighted-sum-of-grey-gases model (WSGGM). The impact of reacting particles or droplets on the continuous phase can be examined using heat and mass transfer relationships, available in ANSYS FLUENT. For coal particles the model includes particle heating, evolution of volatiles and swelling, char reaction and cooling of the particle. For droplet combustion the droplet evolution includes heating to vaporization temperature, evaporation, boiling and cooling. All models mentioned above have been extensively used for an efficient modelling of the complex phenomena in large-scale boilers. The governing equations for the mean flow in tensor notation are (detailed formulations can be found in the ANSYS FLUENT Theory Guide (ANSYS 2009b)):

Continuity:

$$\frac{\partial(\rho Y_i)}{\partial t}+\frac{\partial(\rho U_j Y_i)}{\partial x_j}=-\frac{\partial(J_i)}{\partial x_j}+R_i+S_i \quad (1)$$

Momentum:

$$\rho\frac{DU_i}{Dt}=-\frac{\partial p}{\partial x_i}+\frac{\partial}{\partial x_j}\left((\mu+\mu_t)\left[\frac{\partial U_i}{\partial x_j}+\frac{\partial U_j}{\partial x_i}\right]\right)+S_f \quad (2)$$

Energy:

$$\frac{\partial\rho H}{\partial t}-\frac{\partial P}{\partial t}+\frac{\partial\rho U_j H}{\partial x_j}=\frac{\partial}{\partial x_j}\left(\lambda\frac{\partial T}{\partial x_j}+\frac{\mu_t}{\sigma_t}\frac{\partial H}{\partial x_j}\right)+S_h \quad (3)$$

The continuity equation predicts the local mass fraction of each species, Y_i, in a mixture. Here R_i is the net rate of production of species by chemical reaction and S_i is the rate of creation by addition from the dispersed phase. J_i is the diffusion flux of species i. The eddy-dissipation model is used to calculate R_i.

A two-step mechanism involving oxidation of volatiles (*vol*) to *CO* in the first reaction and oxidation of *CO* to CO_2 in the second reaction is employed:

$$vol + v_{O2}O_2 \rightarrow v_{CO}CO + v_{CO2}CO_2 + v_{H2O}H_2O + v_{SO2}SO_2 + v_{N2}N_2$$

$$CO + 0.5O_2 \rightarrow CO_2 \quad (4)$$

where the stoichiometric coefficients, *v,* are estimated from the ultimate and proximate analyses.

3.4. Heat Exchanger Modelling

For the simulation of the tube bundles (superheaters, reheaters, economizers and hanger tubes) located downstream of the furnace, it is not feasible to model each tube individually as this would result in a very large and complicated computational mesh. Instead, a porous media approach is adopted to model pressure drop and heat transfer. The porous media model adds two source terms to the momentum equations, a viscous term and an inertial loss term, which depend on the molecular viscosity and the square of velocity, respectively.

$$S_i=-\left(\frac{\mu}{\alpha}U_i+C_2\frac{1}{2}\rho|U|U_i\right) \quad (5)$$

where coefficients α and C_2 represent the permeability and the inertial resistance factor, respectively.

For the treatment of the heat transfer inside the porous zone, the energy equation (Eq. 3) is modified in the heat conduction term, using an effective thermal conductivity, λ_{eff} that takes into account the fluid, λ_f and solid, λ_s conductivities and the porosity, β of the medium.

$$\lambda_{eff}=\beta\lambda_f+(1-\beta)\lambda_s \quad (6)$$

The porosity is the volume fraction of fluid within the porous region (i.e., the open volume fraction of the medium). The formula for porosity factor is,

$$\beta=1-\frac{\pi D_o^2}{4S_T S_L} \quad (7)$$

where D_o is heat exchanger tube diameter, S_T is the transversal length (pitch), and S_L is the axial length. The porous model is employed to model three superheaters (SH1, SH3 and SH4), two reheaters (RH2 and RH3), two economizers (ECO1 and ECO2) and hanger tubes for SH1. Superheater SH2 is modelled as plates and reheater RH1 is not modelled. The geometric data used for the calculation of the porosity of the heat exchangers are given in Table 1.

Table 1: Heat Exchanger Parameters

	Rows	Tubes	Diam mm	S_T, mm	S_L, mm	Porosity
SH3	26	33	48.6	522	58.1	0.938
RH2	40	20	63.5	348	115	0.920
RH3	80	8	60.3	174	115	0.857
SH4	80	8	48.6	174	100	0.893
SH1	54	34	54	130.5	100	0.824
ECO 1	27	12	45	130.5	115	0.894
ECO2	27	12	45	130.5	115	0.894
Hanger SH1	54	5	48.6	130.5	100	0.857

Total heat absorbed in each exchanger is modelled by adding an energy source term to the energy equation. The value of the source term is calculated based on the percentage of heat absorbed in each heat exchanger. In Table 2, for example, the percentage of heat absorbed in each heat exchanger for 100% load is shown.

Table 2: Heat Absorbed in Each Heat Exchanger

Heat exchanger	Coal	Heavy Oil
Boiler Walls (%)	34	36
Superheaters (%)	33	31
Reheaters (%)	14	14
Economizer (%)	9	10
Air Pre-heater (%) (not modelled)	10	9
Total absorbed heat (%)	100	100

3.5. Boundary Conditions

Once the mesh has been generated, appropriate boundary conditions need to be applied for the surfaces. This step includes defining the inlet, outlet and walls and specifying the zones for the heat exchangers.

Boundary conditions were obtained from the plant's design data sheets. The air and fuel nozzles were the inlets and the boiler final duct after the economizers was the outlet. The boundary conditions required by the model include primary and secondary air flow rates and temperatures, fuel mass flow rates and temperatures, and fuel properties. The outlet boundary was set as a pressure outlet. The boiler walls were assigned wall boundary conditions for flow and thermal properties. Table 3 shows the main boundary conditions for the simulation cases. Properties of fuels, proximate and ultimate analyses, as well as heating value were taken into account to specify the fuels (coal proximate analysis: moisture 9.5%, ash 12.2%, volatiles 31%, fixed carbon 47.3%; ultimate analysis: C 82.5%, H 5.6%, O 8.96%, N 1.8%, S 1.1%, Cl 0.04%; heating value 26,497.27 kJ/kg. Heavy oil: C 83.64%, H 11.3%, S 4.2%, O+N 0.86%; heating value 41,868 kJ/kg).

Table 3: Boundary Conditions

Parameter	Boundary condition
Coal firing	
Load	100% (350 MW)
Coal flow rate	33.786 kg/s
Primary air flow rate	86.666 kg/s
Primary air temperature	70 °C (343 K)
Secondary air flow rate	256.388 kg/s
Secondary air temperature	321 °C (594 K)
OFA flow rate	21.04 % of secondary air
Heavy oil firing	
Load	100% (350 MW)
Heavy oil flow rate	21.98 kg/s
Heavy oil temperature	117.4°C (390 K)
Air flow rate	305.55 kg/s
Air temperature	325 °C (598 K)
Gas recirculation flow rate	30.8 kg/s
OFA flow rate	13.78 % of total air

4. MODEL VALIDATION

The data needed for model validation, in particular data for CFD-type calculations, are usually not available in commercial utilities. As stated in Fiveland and Wessel (1988), it is impractical and unlikely that enough experimental data could be collected to provide the detailed information needed for CFD modelling. Therefore the global parameters available from the equipment manufacturer and from routine measurements by plant operators may serve as a guide for model validation. In this context validation refers more to agreement in trends than comparison of absolute values.

For the validation calculations the boiler is assumed to be operating at 100% load. Simulations are compared to some global design parameters available from boiler data, mainly values at the furnace exit such as the average temperature and the average O_2 mass fraction. It should be noted that the data is assumed to be an average in a plane at that region. Two different fuels were employed: pulverised coal and heavy oil. The boiler is operated with the lower A–B–C–D levels in operation and the upper E level out of service. For the heavy oil case all four levels are in operation. The tilt angles of the A–D burners were assumed to be 0°. Data from calculations are compared to plant data in Table 4.

Table 4: Comparison Calculations and Reference Plant

Variable	Calculation	Ref.
Fuel: Coal		
Flue gas at outlet, kg/s	372.2	376.66
Gas Temperature		
Flue gas at furnace exit, °C	1171 – 1376	1007
Reheater outlet, °C	646 – 826	779
Superheater outlet, °C	466 – 680	527
Economizer inlet, °C	466 – 676	524
Economizer outlet, °C	336 – 576	343
O_2 at outlet (dry vol %)	2.6 – 6.5	3.6
Fuel: Heavy Oil		
Flue gas at outlet, kg/s	358.20	357.5
Gas Temperature		
Flue gas at furnace exit, °C	1146 – 1606	1017
Reheater outlet, °C	727 – 846	773
Superheater outlet, °C	376 – 562	517
Economizer inlet, °C	376 – 566	546
Economizer outlet, °C	376 – 426	352
O2 at outlet (dry vol %)	6.8 – 8.8	1.1

It was found that calculated values show a big variation in temperature at each plane, which indicates that temperature is not uniform and that high temperature regions exist within the boiler. This behaviour can be expected due to the very complex flow that develops inside the furnace, as can be seen in Fig. 4, where flow streamlines through the boiler are shown. In general the reference data is within the range found in the calculations, with the exception of the gas temperature at the furnace exit, where the calculated values were significantly higher than plant's data.

However, the general trend in gas temperature as the gas flows through the boiler is similar for calculation and reference data. It is acknowledged, however, that this comparison is only a rough approximation towards model validation and that more plant data is needed for better analysis.

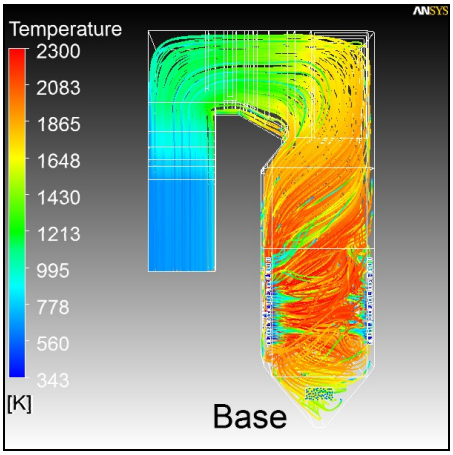


Figure 4: Streamlines Coloured by Temperature

5. TEST CASES

CFD simulations can be employed for numerical simulation of existing and possible operation situations and for the purpose of solving problems in power plants operating in working conditions subjected to change (change of the fuel characteristics, load, etc.). In this work two case studies are presented: Case "a", where the effect of varying the amount of combustion air is investigated, and case "b", in which the failure of one burner is simulated.

Computations have been conducted in a relatively basic PC (2.66 GHz, 16 GB RAM, 4 cores running in parallel). Computation time has varied approximately from 15, 25 to 40 hours of clock time for meshes of 2, 3.5 and 5 million cells, respectively.

A real time lumped-parameter simulator of the boiler, developed in a parallel work, may be employed to establish operational conditions for the CFD model. The real time module allows the user to perform operational manoeuvres such as increasing and reducing load, or operating air or fuel control valves, with the aim of obtaining the dynamic response of some of the main variables such as pressure and temperature of the steam flow to calculate heat absorbed and temperature of combustion gases. For the cases presented here, however, the conditions were simply assumed and the real time module not used.

5.1. Case "a" – Variation of the Air Flow

For the first test case CFD simulations have been performed in order to analyse the effect the amount of air has on the combustion process. It is assumed that variation of the combustion air can have a significant influence on the formation process of some pollutant species such as NO_x. Here a small variation of 10% more than and 10% less than the normal amount of air is specified.

Figure 5 shows a plot of the average gas temperature in horizontal cross-sections at different distances in the path of the gas where 0 m is the furnace bottom and 25 m is the furnace exit. The plot shows lower temperature for the case with less air at the bottom and lower temperature for the more air case at the top of the furnace. Figures 6 and 7 show temperature and velocity contours in a horizontal cross-section at the height of burner level D (10 m) for the three cases. Small differences can be observed, for instance in Fig. 6, the high temperature region is more defined as the amount of air increases, which corresponds to the higher velocity, shown in Fig. 7, as more air is being injected.

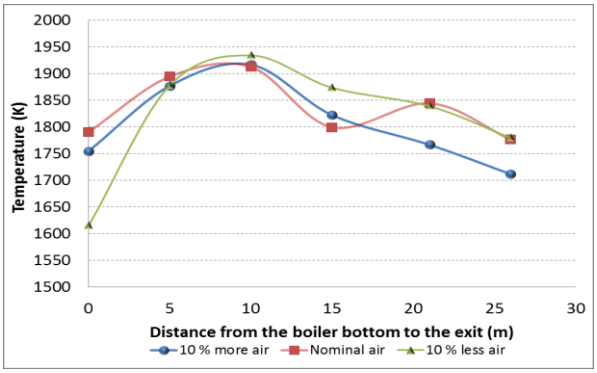


Figure 5: Gas Temperature Varying the Amount of Air

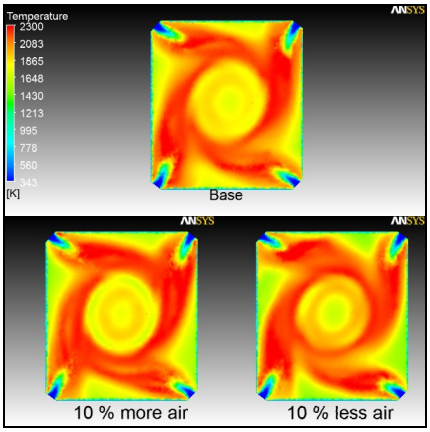


Figure 6: Temperature at 10 m for Test Case "a"

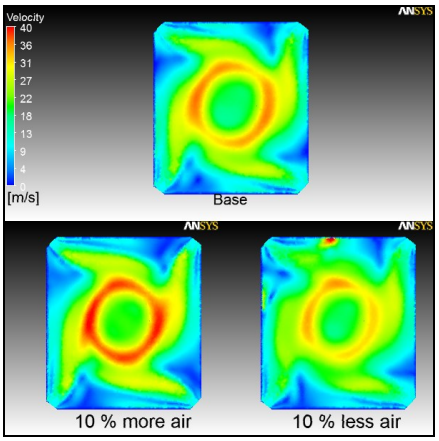


Figure 7: Velocity at 10 m for Test Case "a"

5.2 Case "b" – Failure of One Burner

For the second case the boiler is assumed to be at 100% load and a failure on one of the burners is postulated. The failure is specified as injection of air without coal content. Figure 8 shows temperature, velocity and CO_2 mass fraction contours at horizontal planes at the failed burner height. It clearly shows the effect on temperature and production of CO_2 in the corner with the failed burner. The velocity contour does not show a significant variation because air is still being injected.

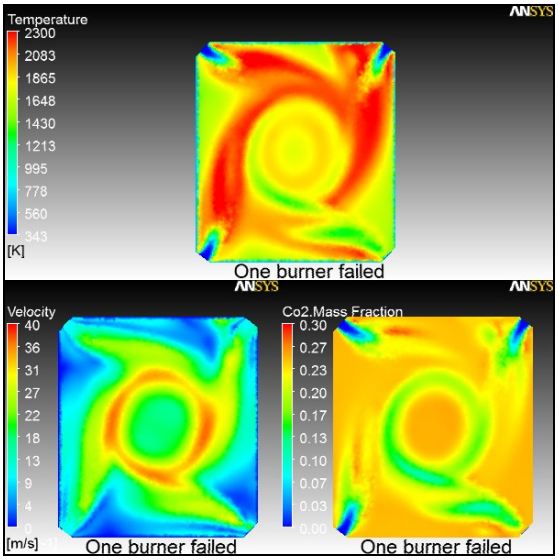


Figure 8: Temperature, Velocity and CO_2 at failed burner height (5 m) for Test Case "b"

6. CONCLUDING REMARKS

A CFD model has been developed to simulate the gas side of a utility boiler. The CFD model forms part of a simulation tool that includes a real time simulator of the boiler and associated systems that may be employed to establish operational conditions for the CFD model. The CFD model has been validated comparing simulation results to design parameters from the reference plant, where validation refers more to agreement in trends than comparison of absolute values. Two case studies have been presented in which numerical simulations were conducted varying the operational conditions: the amount of air available for combustion, and the failure of one burner. These test cases demonstrate the general capability of the simulator and that CFD methods are recommended as a viable computational tool to evaluate the flow and thermal performance in the gas side of the boiler of a power plant.

ACKNOWLEDGMENTS

Financial support for this work was provided by CFE (the Mexican utility, Laboratorio de Pruebas a Equipos y Materiales, LAPEM), and IIE (Institute of Electrical Research).

NOMENCLATURE

C_2	inertial resistance factor (1/m)
D_o	tube diameter (m)
H	specific enthalpy (j/kg)
J_i	diffusion flux of species *i* (kg/m^{2}s)
k	turbulent kinetic energy (m^2/ s^2)
P	pressure (Pa)
R_i	rate of production by chemical reaction (kg/s)
S_i	rate of creation from the dispersed phase (kg/s)
S_f, S_h	source terms (N/m^3, W/m^3)
S_L, S_T	axial and transversal length heat exchanger(m)
T	temperature (K)
U_i	velocity components (m/s)
x_i	coordinates direction (m)
Y_i	mass fraction of species *i* (-)

Greek letters

α	permeability (m^2)
β	porosity factor (-)
ε	rate of viscous dissipation (m^2/ s^3)
λ	thermal conductivity (j/s m K)
λ_{eff}	effective thermal conductivity (j/s m K)
λ_f, λ_s	thermal conductivity fluid/solid phase (j/s mK)
μ	dynamic viscosity (kg/ms)
μ_t	eddy viscosity (kg/ms)
v	stoichiometric coefficients (-)
ρ	density (kg/m^3)
σ_t	turbulent Prandtl number (-)

REFERENCES

ANSYS, 2009a. *ANSYS FLUENT 12.0 User's Guide*, ANSYS, April 2009.

ANSYS, 2009b. *ANSYS FLUENT 12.0 Theory Guide*, ANSYS, April 2009.

Boyd, R.K., Kent J.H., 1986. Three-dimensional furnace computer modelling. *Proceedings of 21st Symposium (international) on combustion*, Munich, West Germany, 1986. pp. 265–274

Dooley, B. and Chang, P.S., 2000. The current state of boiler tube failures in fossil plants. *Power Plant Chemistry*, 2(4), 197-203.

Ferziger, J.H., and Peric, M., 2002. *Computational methods for fluid dynamics*, 3rd rev ed., Springer, Berlin.

Fiveland, W.A., Wessel RA, 1988. Numerical model for predicting performance of three dimensional pulverize-fuel fired furnaces. *J Eng Gas Turb Power*, 110, 117–126.

OIT, 2002. *Improving industrial burner design with computational fluid dynamics tools: progress, needs and R&D priorities*, Workshop report, U.S. Department of Energy's Office of Industrial Technologies (OIT) and the Sandia National Laboratories (SNL).

Roldán-Villasana, E.J, Cardoso-Goroztieta, Ma.J, Tavira Mondragón J.A, Rossano Román M.B, 2010. Lumped Parameters Modelling of the Waterwalls of a Power Plant Steam Generator, *Fourth UKSim European Symposium on Computer Modelling and Simulation*, pp. 283-288. November 17-19, Pisa (Tuscany, Italy).

MATHEMATICAL MODELING OF A HEAT RECOVERY STEAM GENERATOR AND ITS INTEGRATION TO A COMBINED CYCLE POWER PLANT SIMULATOR

[(a)]José Tavira-Mondragón, [(b)]Luis Jiménez-Fraustro, [(c)]Fernando Jiménez-Fraustro

Instituto de Investigaciones Eléctricas (IIE)
Reforma 113, C.P. 62490, Cuernavaca, Morelos, México

[(a)]jatavira@iie.org.mx, [(b)]ajimenez@iie.org.mx, [(c)]fjimenezf@iie.org.mx

ABSTRACT

To expand the scope of a gas turbine power plant simulator to a combined cycle power plant simulator, the mathematical models of process and control and the operation interfaces were developed and integrated in the available hardware-software architecture. To accomplish these goals, software tools were developed and applied to build and simulate a typical distributed control system of this type of power plants with a suitable operation interface for the trainees. Relating to the elaboration of process models, this paper highlights the development of the heat recovery steam generator and the steam turbine. The static and dynamic response of the simulator has been tested and validated regarding plant data. This full-scope simulator provides a suitable human machine interface to train operators of modern power plants. Currently the simulator has been installed in an operators training centre and it is being utilized as a part of the corresponding training courses.

Keywords: mathematical modelling combined cycle power plant simulator, operator training.

1. INTRODUCTION

Currently, full-scope simulators are recognized worldwide as the only realistic method to provide real time and hands-on training of operators (International Atomic Energy Agency, 2004). Full scope simulators include a detailed modeling of all reference plant systems, with which the operator interacts from the control room, and usually include replicas of consoles operation (Instrument Society of America, 1993). In these simulators, and under similar conditions of operation, the response of the simulated unit is very similar in time and indications to the response obtained from the real plant control room. A significant part of expenditure made in these simulators is due to the high fidelity simulation software development. The training quality using a full-scope simulator is clearly better to the one obtained with other simulators, since the operator is acting in an environment that is "identical" to the actual control room, therefore, the operators can be effectively trained due to the variety of initial conditions, malfunctions and operation scenarios.

Since more than 25 years, the simulation department of the Mexican Electric Research Institute (IIE) has designed, developed and started-up all the power plant simulators required for the Mexican Federal Commission of Electricity (CFE) to train its operators of thermal, geothermal, gas turbine and combined cycle power plants. These simulators have different scopes, therefore in their training centres, the CFE have classroom simulators, full-scope replica and no-replica simulators. At this time, the training centre of the CFE with the largest number of simulators is the Ixtapantongo National Training Centre (CENAC-I), which is dedicated to train the operators of thermal power plants. In 2007 the IIE finished the development of a simulator to train operators of gas turbine power plants. This simulator is installed at the CENAC-I, and it replicates the behaviour of a generating unit of 150MW based on a gas turbine, with their control loops and additional systems as: electrical generator, fuel gas, compressed air, lubricating oil and control oil, among others (Roldán, Mendoza, Zorrilla, Cardoso and Cruz 2008). The simulator has a personal computer-based operation interface. As is described by Melgar, Tavira, Rodríguez and Cruz [2008], this type of interface allows operator (trainee) to view and perform the operation actions required from Interactive Process Diagrams (IPD). With this architecture the operator navigates through any IPD, and operates at any moment each one of the simulated systems. Due to the additional functions of the interface, the operator can monitor the behaviour of the simulated unit by means of trend graphs and be aware of any problem thanks to the alarms monitor. Furthermore, the operator has interactive windows to modify the status of any equipment, for instance: turn pumps on, open valves, or modify set points of controls.

To expand the scope of its gas turbine power plant simulator, CFE requested to the IIE to develop the process and control models, as well as the operation interfaces necessary to integrate a simulator of a 450 MW combined cycle power plant. This article describes the major features of the software and hardware architecture and presents the main results achieved in this development.

2. HARDWARE AND SOFTWARE ARCHITECTURE

The current gas turbine power plant simulator is installed at the CENAC-I on a computers network, therefore the 450 MW combined cycle power plant simulator was implemented in the same computers, which are personal computers with Pentium 4 (3.6 GHz) processor with 1 GB of memory and Windows XP operating system. Figure 1 shows the arrangement of these computers. The OC-1 and OC-2 are the operator consoles, the OC-1 has two 20" monitors and the OC-2 has two 20" and two 50" monitors. The OC-1 and OC-2 are utilized by the trainee to perform the required operation actions and/or monitoring the simulated process. On the other hand, the IC is the instructor console, and it is where the instructor of the training session guides the simulation practice.

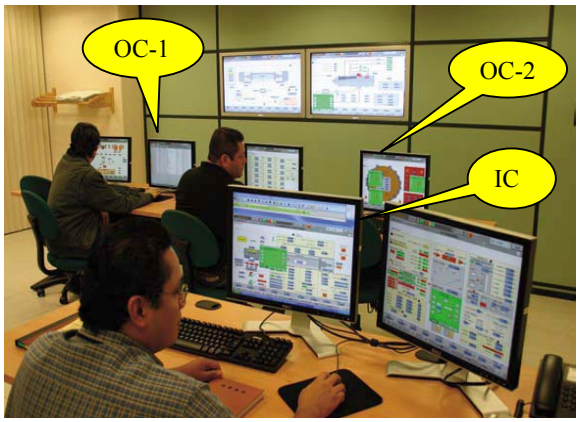


Figure 1: Computer Platform

2.1. Software Architecture

The general architecture of the simulator software is shown in Fig. 2. The Executive, Operator and Instructor Console modules are described in detail by Tavira, Jiménez, and Romero [2010]; therefore just a brief description of these modules is given here. The Executive module is responsible of sequencing the real-time tasks requiring a cyclical execution as: mathematical models of process and control, interactive process diagrams, Transducer module (Models-Control Communication Manager), and the managers of the Instructor Console and Operator modules. For its part, the Operator module represents the Human-Machine Interface (HMI) of the simulator user, and it consists of a group of IPDs with which the user operates the simulated process. Finally the Instructor Console module is the interface of the instructor who is leading the simulation session, from this interface the instructor establishes the initial condition, supervises the user actions, introduces malfunctions or performs some other type of actions required during simulation. The real time operation of the simulator is guaranteed because of the implementation of a "hard" real-time module. This module allows the creation of timers (from 0.1 to 1 millisecond) and system threads. This module also carries out the managing of resources such as memory, ports and interrupts.

2.1.1. Control Simulation Modules

While previous modules are the basis for the implementation of the simulator, the complexity and size of modern control systems, requires the use of generic tools that allow optimizing development and also obtaining an appropriate representation of the simulated system regarding to the control utilized as reference. Control Editor and Control modules represent these tools, which for purposes of this project were adapted to the requirements of the control system of a combined cycle unit.

The Control Editor module is a program whose main goal is to make a graphical modelling of modern control systems; these control algorithms are organized into components with a very specific function (PID, high/low detector, set/reset, etc.). This organization represents a network of components which transfer information through connections. These networks are organized hierarchically. At the lowest level of the hierarchy are basic elements or gates, which are organized in diagrams, a diagram or a group of diagrams constitutes a control loop, and a group of loops a module. Finally, a set of modules simulates the control system of the power unit.

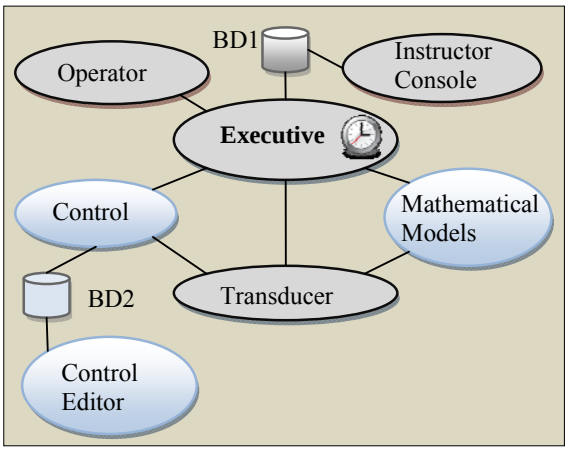


Figure 2 : Software Architecture

For its part the Control module is a tool that allows the dynamic assembly of predesigned software components. This methodology is based on the design of components which have a generic description of their dynamic behaviour; this allows generating an executable model in real time. The use of this development facilitates the construction of the model to simulate the control system, and has the ability to view during simulation the flow of logic and analogue signals of control algorithms. This allows understanding the functional conduct of the control, tuning parameters and helping with detection of problems associated with the

model. Once the control models are assembled, they can be executed in parallel on multiple processors because they can be distributed in several computers and/or multi-core computers.

The control system developed for the simulator has the following features: logical protection equipment, alarms system, turbine interlock, automatic sequencing for turning equipment (e.g. turbine start-up) on, and automatic controls of level, pressure and temperature for each Heat Recovery Steam Generator (HRSG). On the other hand, the operation interface of the control system is based on the IPD of the simulated process areas (e.g. HRSG1, HRSG2, steam turbine, etc.). These IPDs contain information such as: indications of equipment status, major variables displays, trends graphs, control stations for turning equipment on, and auto/manual transfers, among others.

3. HUMAN MACHINE INTERFACES

The HMI for the instructor console is a windows-friendly application with pull-down menus and icons, as it is shown in Fig. 3.

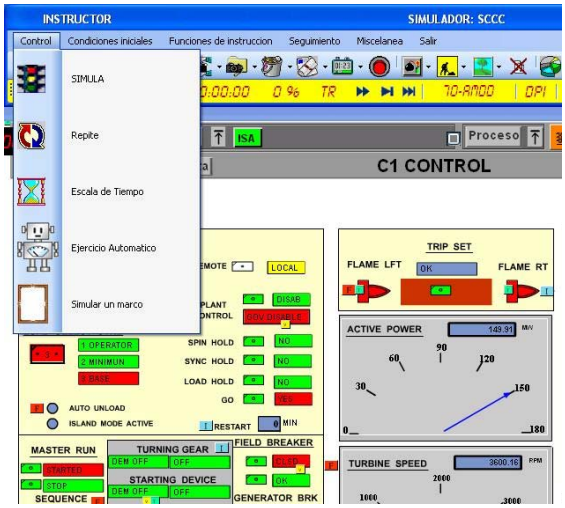


Figure 3: Partial View of Instructor HMI

The main functions of this interface are:

1) Run/Freeze. The instructor starts or freezes a dynamic simulation, during this simulation, the mathematical models of control and process respond to the actions of the trainee in a similar way as it happens in the actual plant.
2) Initial Condition. This function has the following options:
    a. Selecting an initial condition to begin the simulation session. Each instructor has its own set of one hundred initial conditions;
    b. Recording a new initial condition (snapshot) or to erase a previous initial condition;
    c. Specifying the time interval to get automatic snapshots.
3) The option of malfunctions is used to introduce/remove an equipment malfunction at any time during the simulation session. There are 205 malfunctions. Some examples are: pumps trip, tubes break in heat exchangers, and valves obstruction. All the malfunctions are grouped in systems and subsystems for an easy location. For binary malfunctions, the instructor has the option of defining its time delay and its duration. For analog malfunctions, besides the former time parameters, the instructor can also define their intensity and their time evolution.
4) The instructor has the option of internal parameters for simulating the operative actions not related to automated equipment. These operative actions are associated with the local actions performed in the actual plant by an auxiliary operator, e.g., open (close) valves. Similarly to the malfunctions, they are grouped in systems and subsystems and they have time and intensity parameters as instructor options.
5) The option of external parameters allows the instructor to modify the external conditions to the process. Some of these conditions are: atmospheric pressure, ambient temperature, voltage and frequency of the external electric system.
6) The instructor also can create automatic training exercises. Each one of these exercises can include: initial conditions, malfunctions, local actions, and a time sequence. The exercises are stored in a database for their subsequent use.
7) In its default mode, the simulator is executed in real time, but the instructor can execute the simulator up to ten times faster or up to ten times slower than real time. This option is especially important when the trainee wants to analyse a fast transient, allowing him to simulate slower, or in the case of a slow thermal process, like state turbine's iron warming, it can be simulated faster.

Some additional important features of the instructor console are its portability and its auxiliary functions, like trends graphs and variables monitoring, so during the development stages of the project, the personnel in charge of the mathematical models utilized these functions to simulate, tune and debug the models. On the other hand, the HMI for operator is also completely graphical and based on a multi-window environment with interactive process diagrams, these diagrams are organized in hierarchical levels following the organization of the power plant systems, i.e. HRSG-1, HRSG-2, steam turbine, etc.

There are two main types of diagrams: information diagrams and operation diagrams. The first one shows to the trainee the values of his selected variables, or a predefined set of variables. The values are presented as trends graphs. The operation diagrams are utilized by the trainee to control and supervise the whole process; with them he turns equipment (pumps, fans and compressors) on(off), opens (closes) valves, modifies

set points of automatic controls, and carries out any feasible operation in a similar way as he would do in the actual power plant. When the trainee needs to perform an action, he selects the suitable icon, and then a pop-up window appears with the corresponding operation buttons. Figure 4 shows a partial view of the operation diagram of steam lines for HRSG-1 and HRSG-2.

4. MATHEMATICAL MODELLING

Figure 5 shows a schematic of the main parts of a combined cycle power plant. Each one of the gas turbines delivers 150 MW, while the two HRSGs and the steam turbine generate 150 MW of electric power.

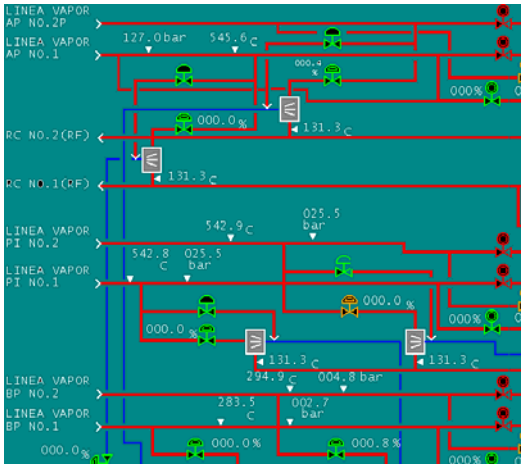


Figure 4: Partial View of HMI for operator

The Mathematical Models simulate the behaviour of the actual power plant in any feasible operation condition. Depending on its complexity, the mathematical models may be constituted by linear algebraic equations, non-linear algebraic equations, ordinary differential equations, or in most cases, by a combination of them. Usually linear equations are solved by LU decomposition methods, non-linear algebraic equations by Newton-Raphson methods and differential equations are solved by fixed step integration methods. The mathematical models of the simulator are divided into electric models and process models. The complete simulator includes all models required to represent the full conduct of the reference unit, in this paper, as an example of the type of models developed, the main equations related with the simulation of the HRSG and the steam turbine are presented.

Figure 6 shows a simplified diagram of the two HRSG array. Each HRSG is in charge of receiving the output hot gases (620 deg C) of the respective gas turbine, and because of a group of heat transfer equipment, liquid water from feedwater system is transformed into superheated steam of three pressure levels: high pressure (126.1 bar and 538 deg C), intermediate pressure (25.5 bar and 536 deg C) and low pressure (4.5 bar and 273 deg C). For each one of the HRSGs, their steam lines of a similar pressure are joined to feed the steam turbine as appropriate. Each one of the HRSGs has sixteen sections divided in three pressure levels, and each level has their corresponding economizers, evaporators and superheaters.

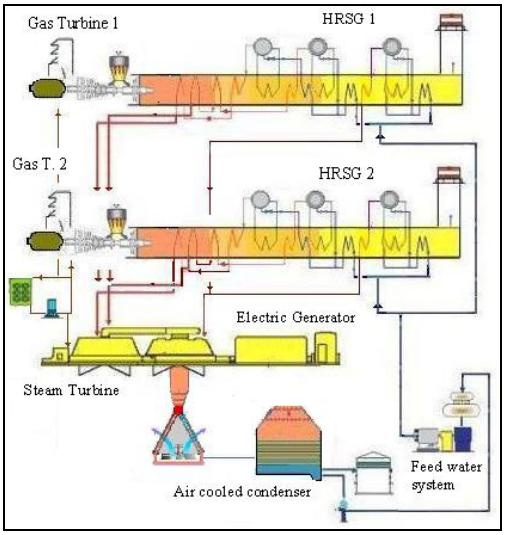


Figure 5: Combined Cycle Power Plant

For each section of the HRSG the main calculations carried out are:

1) Flow rates of gas and water (liquid/steam).
2) Heat flow rates from gases and to the water.
3) Levels of water in drums and evaporators.
4) Temperature of gases
5) Internal energy of water
6) Temperature of tubes (metal)

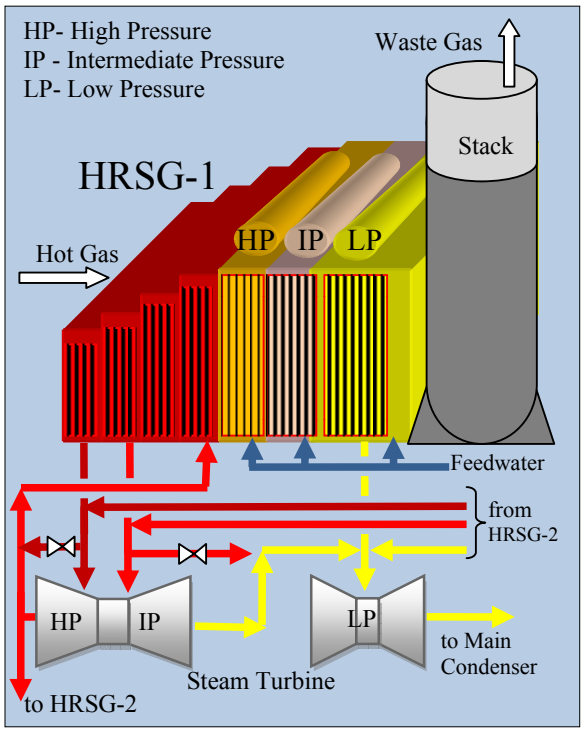


Figure 6 : HRSGs and Steam Turbine

The model was developed under the approaches of lumped parameter and perfect mixing, in this way; the gas temperature variation for each element of the HRSG is represented by:

$$\frac{dT_g}{dt} = \frac{W_g Cp_g (T_{g(e)} - T_g) - Q_g}{\rho_g V_g Cp_g} \quad (1)$$

In a similar way, the water energy variation inside the tubes for each element of the HRSG is calculated by:

$$\frac{dU_f}{dt} = \frac{W_f (h_{f(e)} - h_f) + Q_f}{\rho_f V_f} \quad (2)$$

For the tubes, the metal temperature variation is:

$$\frac{dT_m}{dt} = \frac{Q_g - Q_f}{m_m Cp_m} \quad (3)$$

Heat flow rates from gas and to the water are calculated as:

$$Q_g = j_g A_g \left(T_g - T_m\right) \quad (4)$$

$$Q_f = j_f A_f \left(T_m - T_f\right) \quad (5)$$

The above equations are evaluated for each one of the elements of the HRSG. The model is complemented with the calculation of: thermodynamic properties of water, properties of combustion gases, heat transfer coefficients, and flow rates. In this case the equation utilized is:

$$W = K \quad Y \sqrt{X \ \rho \ P_e} \quad (6)$$

where, $X = 1 - Pe / Ps$, and Y is the expansion factor. This factor compensates changes in the fluid properties by pressure variations. For liquids $Y = 1$, for gases the calculation is based on the verification of the conditions of critical flow (Crane Co. 1988).

The pressure and liquid mass variations for each one of the drums are calculated from the mass and energy balances over the corresponding equipment.

$$\frac{dP}{dt} = \frac{\sum h_e W_e - \sum h_s W_s + x1\left[\sum W_e - \sum W_s\right]}{x2 \ \ m_L - V} \quad (7)$$

$$\frac{dm_L}{dt} = \frac{x3 \ \frac{dP}{dt} - x4\left[\sum W_e - \sum W_s\right]}{x5} \quad (8)$$

where: *x1, x2, x3, x4* and *x5* are partial derivatives of density and enthalpy of saturated states of the water evaluated to the pressure *P* of each one of the drums, and the summations are applied over all the inlets and outlets streams of the drums.

On the other hand, the model for the steam piping and steam turbine is based on the development presented by Tavira, Melgar, García, and Cruz [2008]. This model solves the continuity equation for all points where steam flows are mixed or separated, these points are called nodes.

$$\sum W_e - \sum W_s = 0 \quad (9)$$

The flows from each one of the elements (valves and turbine stages) of the flow network are always associated with a pressure drop, in the case of the valves (governing, stop and bypass), Eq.(6) is utilized too. The calculation of the flow rate through the turbine stages is based on the following equation:

$$W = J \sqrt{2 \frac{m}{m-1} P_e \rho_e} \sqrt{p^{2/m} - p^{(m+1)/m}} \quad (10)$$

where m is the polytrophic coefficient for steam and p is defined as:

$$p = \frac{P_i - \frac{P_e - P_s}{\delta}}{P_e} \quad ; \delta \geq 1 \quad (11)$$

δ is the stage number where the flow rate is being evaluated

In this way, if each one of the flow rate equations is replaced into the continuity equation for every node, a system of non-linear algebraic equations is obtained, where the pressures of the nodes are the independent variables. This system is resolved by the Newton-Raphson method.

The turbine model is complemented with the calculation of the enthalpy output for every stage:

$$h_j = h_{j-1} - \eta_j \left(h_{j-1} - h_{id.j}\right) \quad (12)$$

and finally with the enthalpy differences between stages, the mechanical power produced is calculated.

All parameters of the models as: volumes inside and outside of tubes, mass of metals, valves conductance, turbine stages efficiencies, etc., are calculated with equipment design data and operation data of the combined cycle unit utilized as reference.

5. RESULTS

The developed simulator for a combined cycle unit has a distributed control system with an operation interface similar to the interface available in the actual power plant. The software tools used for these developments facilitate the elaboration, integration and debugging of the simulated control system.

In order to test and validate the dynamic response of the simulator, the instructors of the CENAC-I applied a group of acceptance tests, whose main objectives were: a) Verifying the simulator response in any feasible operation condition; b) Validating the simulator

response when a malfunction is inserted; c) Testing the right operation of the user interface.

Regarding the simulator response, it has been tested and from cold shutdown to nominal power and malfunctions. Table 1 shows the absolute error percentages of the temperatures and flow rates through all equipment elements for a 100% generation state; these errors are calculated between simulated values and design data. In general, the errors are lower than 10%, which is a reasonable measure of the model accuracy.

Table 1: Absolute Error Percentages of HRSG Model

HRSG		Water Side		Gas Side	
Pressure Level	Element	Outlet Temp.	Flow Rate	Outlet Temp.	Flow Rate
High	Superheater-2	2.08	0.37	0.91	0.18
Intermed.	Reheater-2	0.17	1.86	0.20	0.00
Intermed.	Reheater-1	0.04	1.88	0.59	0.00
High	Superheater-1	1.58	0.16	0.33	0.18
High	Evaporator	8.89	7.56	2.57	0.18
Intermed.	Superheater	0.76	0.96	0.01	0.02
Low	Superheater	3.89	5.73	0.60	0.02
High	Economizer-3	2.61	0.08	2.43	0.18
Intermed.	Evaporator	0.41	2.21	0.11	0.00
High	Economizer-2	1.66	0.08	0.74	0.18
Intermed.	Economizer-2	10.10	5.01	6.72	0.00
Low	Evaporator	0.95	12.7	2.44	0.00
High	Economizer-1	0.27	0.08	0.48	0.18
Intermed.	Economizer-1	5.79	5.01	4.90	0.00
Low	Economizer	6.60	0.96	2.30	0.00
Low	Preheater	8.40	7.89	4.07	0.00

About the dynamic response, Figs. 7, 8 and 9 show the results obtained during a start-up, from cold iron up to full load.

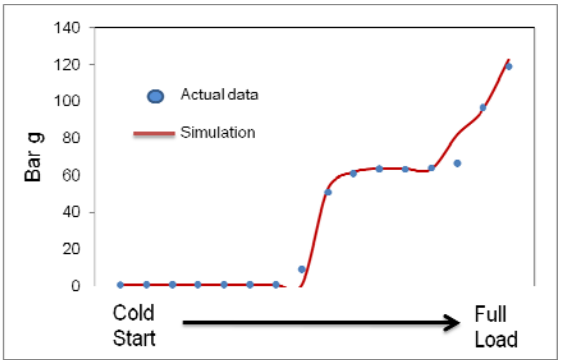


Figure 7: HRSG High pressure drum

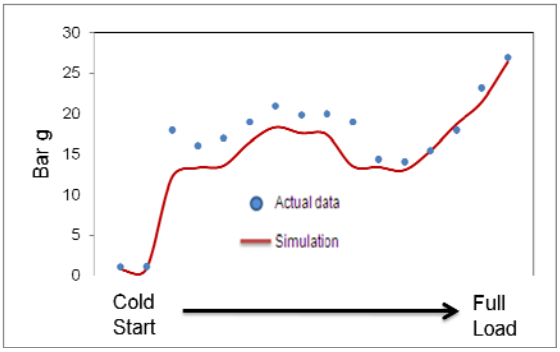


Figure 8: HRSG Intermediate pressure drum

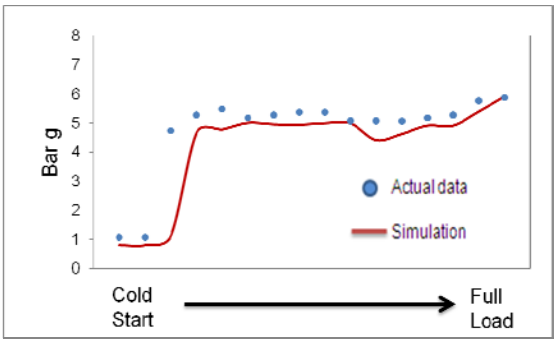


Figure 9: HRSG Low pressure drum

The simulated behaviours of the high, intermediate and low pressures of the HRSG follow the expected trends regarding actual data of the reference power plant.

6. CONCLUSIONS

In order to expand the capacity of a gas turbine power plant simulator, to a combined cycle power plant simulator to train operators, we integrated the control algorithms of a distributed control system, the operation interfaces, and the mathematical models of HRSG, steam turbine and auxiliary services into the prior simulator. The static and dynamic response of the simulator has been extensively tested and validated regarding plant data. This simulator has full-scope and real time features, and provides a suitable HMI for the operators of modern power plants.

The combined cycle power plant simulator is currently installed at the training centre for operators of the CFE, and it is being utilized as a part of the training courses for the operators.

ACKNOWLEDGMENTS

This project was supported by the Federal Mexican Commission for Electricity. The authors would like to thank all the personnel of the IIE and CENAC-I who participated in the project development. The simulation environment is proprietary software of the IIE, and it was customized to the particular requirements of this project.

NOMENCLATURE

A	Area
Cp	Heat capacity
h	Enthalpy
HP	High Pressure
IP	Intermediate Pressure
j	Heat transfer coefficient
J	Turbine stage conductance
K	Valve conductance
LP	Low Pressure
m	Mass
P	Pressure
Q	Heat flow rate
t	Time
T	Temperature
U	Internal energy
V	Volume
W	Flow rate

Greek letters

η	Stage turbine efficiency
ρ	Density

Subscripts

e	Inlet
f	Liquid, steam or mixture
g	Gas
id	Isentropic
j	Turbine stage
L	Liquid in the drum
m	Metal
s	Outlet

REFERENCES

International Atomic Energy Agency, 2004. *Use of control room simulators for training of nuclear power plant personnel*. Austria: IAEA.

Instrument Society of America, 1993. *Fossil-Fuel Power Plant Simulators–Functional Requirements*. U.S.A.: IAEA.

Roldán, E. J., Mendoza, Y., Zorrilla, J., Cardoso, M., Cruz, R., 2008. Development of a Gas Turbine Full Scope Simulator for Operator's Training. *Proceedings of the Second UKSIM European Symposium on Computer Modeling and Simulation*, pp. 376-381. September 8-10, Liverpool (England).

Melgar, J.L., Tavira, J.A., Rodríguez, S., Cruz, R., 2008 Desarrollo de Simuladores para Entrenamiento de Operadores de Centrales Eléctricas Modernas (Development of Operator Training Simulators for Modern Power Plants), *Proceedings of the* 6º *Congreso Internacional en Innovación y Desarrollo Tecnológico*. October 8-10, Cuernavaca, Mor. (Mexico).

Tavira, J., Jiménez, L., Romero, G. 2010. A Simulator for Training Operators of Fossil-Fuel Power Plants with an HMI Based on a Multi-Window System. *International Journal of Computer Aided Engineering and Technology*, 2 (1), 30-40.

Crane Co., 1988. *Flow of Fluid Through Valves, Fittings, and Pipe, Technical paper, No.410*, U.S.A.: Crane Co.

Tavira, J., Melgar, J., García, J., Cruz, R. 2008. Upgrade of a Full-Scope Simulator for Fossil-Fuel Power Plants. *Proceedings of the Winter Simulation Conference*, pp. 1410-1418. December 7-10, Austin, Tx. (U.S.A.).

A SIMULATION BASED DESIGN FRAMEWORK FOR LARGE SCALE INFRASTRUCTURE SYSTEMS DESIGN

Yilin Huang[(a)], Mamadou D. Seck[(b)], Michele Fumarola[(c)]

[(a,b)] Delft University of Technology
Faculty of Technology, Policy and Management
Systems Engineering Group
[(c)] Macomi Systems Simulation

[(a)] y.huang@tudelft.nl, [(b)] m.d.seck@tudelft.nl, [(c)] m.fumarola@macomi.nl

ABSTRACT
This paper discusses an enhanced use of Modeling and Simulation (M&S) in the design cycle of large scale infrastructure systems. After a short review of the use of M&S in the current systems design process, we identify some issues in multi-actor design environment. A Simulation Based Design (SBD) framework is proposed to tackle these issues using a multi-methodological approach. The framework is specified with formal description, and the tree-like structuring and refinement of design alternatives are explained. A case of applying the framework to container terminal design is presented and the evaluation is reported.

Keywords: simulation based design (SBD), systems engineering (SE)

1. INTRODUCTION

Infrastructure design is often complex and has critical implications at strategic, tactical and operational levels for organizations in public and private sectors. Design faults are difficult to foresee and can entail high costs when they emerge later at development and operational phases. A lot of modern infrastructure design processes have many actors involved. We can identify different actors, e.g., as problem owners, policy makers and clients. They may have diverging or even conflicting interests and opinions. For them, the decision making process is actually also a learning process wherein they are engaged in identifying shared problems and objectives, and understanding the others' positions and views.

In engineering, designing systems is a part of the field of Systems Engineering (SE). During the past decades, theories and methods in SE have brought many beneficial changes in how engineered systems are designed and developed. When we take a broad look at the complete engineering process, we can often identify a set of phases that follow an iterative and incremental path. The problem is first structured; then the solutions are formulated based on the selected criteria; and finally the best alternative is selected (Simon 1996). Along the iterative process, the problem definitions and system requirements supposedly become more lucid and complete, contributing to solution finding.

In order to render SE processes more effective and efficient, we need to apply design and engineering methods with supporting tools. Following (Shannon 1975) and (Sol 1982), simulation can be seen as a method of inquiry that provides a possibility to study systems. It produces quantitative measures that allow users to study the dynamic behavior of the systems and make informed decisions of the design. Simulation Based Design (SBD) is promising in terms of providing designers and decision makers with insights into the engineered system at an early stage in the product lifecycle. This in turn increases the chance to correct design faults before development and operation.

Although the SBD concept holds a lot of potential, the realization in practice is hindered by the fact that Modeling and Simulation (M&S) is highly specialized and time-intensive. Given the current available M&S methodologies and tools, it is difficult to fully integrate M&S into the design process of complex systems. The reason is straight forward: modeling design alternatives of complex systems often takes long time and model building is expensive. Therefore, SBD methodologies and accompanying tools are required and should be developed if we are going to give SBD (currently merely a concept in theory) practical implications that could benefit designers and decision makers.

SBD methodologies and tools should allow designers, who often are not M&S experts, to simulate and evaluate their design without the concern of simulation model building. A SBD framework should offer easy-to-use tools with functionalities such as automatic model generation and calibration, version management, advanced data visualization, analysis and comparison.

In this paper, we present a SBD framework for large scale infrastructure design. We first give a short review of the use of M&S in current systems design processes, and identify some issues that often occur in multi-actor design environments. In Section 3, four constructs that are used in the design of the proposed framework are discussed. Section 4 specifies the

proposed SBD framework and explains the design process. A case of applying the framework to container terminal design is presented in Section 5 and the evaluation was conducted both with novice users (graduate students) and experts.

2. NEED FOR SIMULATION BASED DESIGN

2.1. Simulation in Traditional Design Processes

The traditional role of simulation models in the design process is in the analysis step of systems definition: evaluate design alternatives obtained in the conceptual design phase. Modeling often needs the following steps (Law 2007). First, the conceptual designs along with the system requirements and the definition of Key Performance Indicators (KPI) are provided to the M&S specialists. In the succeeding steps, the corresponding conceptual models are built and translated into M&S specifications. After model verification and validation with the domain experts, experiments are designed and performed to evaluate the design alternatives. This typical use of M&S has some weaknesses as stated below (Fumarola 2011).

Because of the time and high cost entailed by modeling, the number of design alternatives actually assessed through simulation is not high. M&S tasks are often conducted by professionals outside of the design team. They need some time to acquire proficiency in the problem domain in order to develop reasonable models. In many cases, when the simulation models are domain, organization or case specific, they need to be developed from scratch; hence these models have poor reusability, especially outside of an organization.

The acquisition of domain knowledge may be alleviated by choosing M&S consultants specialized in the relevant field. However, the transfer of concepts from designers to modelers is not without risk because they do not necessarily share the same mental model of the problem situation, the current state of the system and the envisioned solution.

Outsourced M&S consultants generally only deliver specific answers to the initial questions that are requested by the problem owners. The models, if they are a part of the deliverables at all, are often unable to answer new questions that possibly emerge at a later stage in the design process. Extending the model needs professional modeling expertise so that the designers have to commission the task again.

Although the compartmentalization of conceptual design and detailed design phases helps structure and reduce design complexity, it also forces designers to move back and forth the two phases. The designers often do so in an unstructured way. Blending these two phases may save design and modeling resources. But in doing so, supporting tools are required to reduce the complexity.

(Saanen 2003) advocates for the use of simulation throughout the design process. However, the current practice suggests that little has been changed due to applicatory difficulties. Some are mentioned above.

2.2. Multi-Methodological Approach

M&S has been considered as a methodology within the hard systems approach, a traditional school in systems thinking. Many studies show that the hard systems approach has some limitations, for its basic assumption of the existence of an optimal or quasi-optimal solution which can be solved by algorithmic methods without much consideration for the human actors involved in the system (Ackoff 1979, Simon 1996). Some attempts hence ensued to mitigate the perceived limitations of the hard systems design processes. This resulted in a tendency to pay more attention to the diverging views of different actors involved in the process. This school of thought is referred to as soft systems methodology (Checkland 1999). (Robinson 2001) advocates for a "hard" to "soft" continuum instead of bipolar extremes. To benefit from both, multi-methodological approaches have been introduced.

Many survey results reported a tendency towards combining hard and soft systems methods, e.g., in (Munro and Minger 2002). Multi-methodology has been used in problematic situations to support the model building phase of simulation studies (Mingers and Rosenhead 2004). Qualitative methods and techniques are suggested to be embedded into multi-actor design processes that leave more room for negotiation and mutual learning. They have the potential to support design processes (1) in a multi-actor environment with diverging stakes and (2) using simulation technology to foster discussion based on explicit knowledge.

2.3. Requirements for Simulation Based Design

The multi-methodological view (Robinson 2001, den Hengst et al. 2007) could benefit the SBD process in which designers and stakeholders can learn the others' perspectives by participating in the steps of the process. This feature is not yet present in the current design or modeling frameworks, e.g., (Shannon 1975, Zeigler et al. 2000). Some issues identified in the current systems design processes include (but may not be limited to) the following (Fumarola 2011):

1. **Unstructured Design Process**: without a well-defined structure, the design process may not be able to support designers' abstract and divergent thinking.
2. **Design Management**: without the ability to manage different designs in the design process, designers may not be able to compare them.
3. **Documentation**: without the ability to document assumptions, decisions and argumentations, actors may not be able to follow and audit the design process.
4. **Compare Designs**: without the ability to compare different designs in the design process, actors may not be able to identify the preference.
5. **Consistency**: Actors may have difficulties in keeping the conceptual designs and the detailed designs mutually consistent.

6. **Involvement**: not every actor is involved in both the conceptual design process and the detailed design process.
7. **Prediction**: without simulation models, actors may not able to predict the design decisions.
8. **Understandability**: the design is often presented in a way that not all actors can understand.
9. **Process Flow**: The system design process is often neither interactive nor iterative.

These issues are identified as relevant in many systems design projects. They reflect the requirements for the design of a SBD framework and tools. A design process that involves multiple actors requires a SBD framework that facilitates collaboration to enhance shared understanding. The framework should support automatic model generation so that a design can be transformed into executable simulation models without advanced modeling knowledge. The models should be available at different levels of completion and abstraction reflecting the nature of the design process. The model experiments should be manageable by the designers, and the simulation outputs generated should be presented in a way that they allow non-simulation specialists to compare different alternatives. And the different design alternatives (and models) need version management so that the designers can trace different versions and their logical relations systemically.

3. A SYSTEMS ENGINEERING APPROACH FOR SIMULATION BASED DESIGN

SE has evolved from being underpinned by a rather mathematical and hard systems approach of design towards an approach of transforming users' needs into an operational system via inter-disciplinary efforts. SBD influences systems development models, e.g., Royce's Waterfall model, Boehm's Spiral model and Forsberg and Moog's V-model. The former two models have many applications in software engineering, while the V-model remains dominant in SE. In current systems design process, some distinct phases can be recognized, e.g., requirements analysis, conceptual design and detailed design.

We posit that SBD could cover the three phases by blending them into a participative design phase. Using SBD, the design process on its own is both interactive and iterative such that all actors can be involved in discussions to exchange their opinions, and they should be able to explore the system requirements and the designs using the supporting tool. The conceptual design and detailed design phases are currently often separated because the details needed in the latter are considered as a burden in the former. However, in a SBD process, the model building blocks should be able to encapsulate the complexity in detailed designs and expose only the relevant information that is necessary to the conceptual design.

Based on these concepts and the issues addressed in Section 2, we propose four constructs that can be used for the design of a SBD framework. (We use the term construct to refer to the related methods or models that can be used to tackle a set of the issues.) Each construct tackles some issues addressed.

1. Component-based Modeling (issues 2, 4, 7, 9)
2. Different Levels of Specification (issues 1, 5, 6)
3. Structure Alternatives (issues 1, 2, 4, 9)
4. Participatory Design (issues 3, 6, 8, 9)

A SBD framework requires domain specific tool support. An extensible model component library is needed to allow the modeling of design alternatives. Pre-defined simulation components let users build new simulation models by focusing on the composite structure without the concern of the inner details of the components. Components support model modularity by using port-based communication (Paredis et al. 2001). Once this approach is adopted, model reusability is also enhanced. Therefore, we propose a component-based architecture as shown in Figure 1.

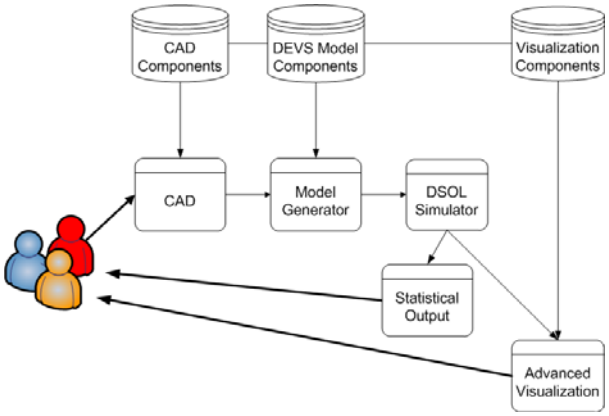


Figure 1: A Component-based Architecture.

Three libraries are in place: the design components (CAD), the simulation model components and the visualization components. The CAD components are used by designers in a design environment. The design is passed on to the Model Generator which transforms it automatically into a simulation model (Fumarola et al. 2010, Huang et al. 2011). This step should be transparent to the designers. The model will be executed and visualized. The results form an instant feedback loop to the designers and the other actors involved in the design process.

A typical design process starts with an incomplete picture of the design problem. Therefore, it is infeasible to start directly with a detailed solution; instead, the design is refined in a couple of steps to deal with uncertainty, complexity, and the cognitive limitations of designers (Hoover and Rinderle 1991, Goel 1995, Ullman et al. 1998). Correspondingly, along the design process, simulation models should be constructed to fit the pre-defined abstraction levels. This allows designers analyze simulation results throughout the process, even in the initial phases characterized by rough and incomplete designs. Further down the process, abstract models are replaced by more concrete

implementations until the end where detailed models are used to gather precise simulation results.

Creating alternatives is one of the fundamental steps in design methods. In a multi-actor design environment, the amount of alternatives in the solution space could become very high. They are often reduced to a possible minimum at an early stage of the design process due to the high cost of maintaining and comparing the alternatives. This unfortunately induced designers to systematically explore the solution space using trees, matrices, rankings and charts. Structured experimentation (Dwarakanatha and Wallacea 1995) can bring more insight into the way designers think. With it, designers can follow the paths in the tree-like structure to assess alternatives. Whenever a solution seems undesirable, the decision branch can be removed. Structured experimentation guides designers towards an iterative process to specify alternatives, evaluate them (formal or informal) and make selections.

A system design process is seldom performed by a single actor or even a single person. In a multi-actor design process, each actor tries to achieve his/her own goals. Hence, the design process should in general support a certain convergence of interests of all actors. The conceptual designs are the outcomes of negotiation and agreement between the actors. The major decisions (in the conceptual design) are made based on informal assessments because the simulation studies are often performed at the detailed design process. By bringing formal methods closer to the conceptual design process, the reasoning for choices can be more grounded and shared understanding can be enhanced. This leads to an approach that is "soft with a hard centre" (Robinson 2001). Moreover, explicit documentation ensures the consistent provision of rationale behind the decisions and provides insight into the design making process (Girod et al. 2003) in a multi-actor decision making environment.

4. THE PROPOSED FRAMEWORK

Based on the four constructs discussed in the previous section, this section proposes a SBD framework. The framework is specified with a formal description that assumes the use of the DEVS formalism (Zeigler et al. 2000) to model the system. The SBD framework

$$\mathcal{F} = \langle\ P, A, \mathcal{L}, \zeta, \eta, Z, \mathcal{Z}, \rho\ \rangle$$

where

P is the set of triples (N, L, t) where
N is a coupled model, as defined in (Zeigler et al. 2000)
L is the set of components of N, $L \in \mathcal{L}$
t is the creation time of N, $t \in \mathbb{R}^+$

A is a tree defined as $(P, <_{t,z,L})$ where
$<_{t,z,L}$ is the ordering relation of the creation time t, the applied decisions z, and the set of components L,
$z \in \mathcal{Z}$ and $L \in \mathcal{L}$

$\mathcal{L} = \{L_0, L_1, \ldots, L_k\}$ is the set of component sets where $\forall L \in \mathcal{L}$ and $\forall c \in L$, c is a coupled model

$\zeta: L_i \rightarrow L_j, L_i, L_j \in \mathcal{L}$ is the component transformation function where $\zeta(c_{Li}) = c_{Lj}$, $c_{Li} \in L_i$, $c_{Lj} \in L_j$

$\eta : S \times R \rightarrow S$ is the initialization selection function that selects the set of states such that $\eta(s, r) = s_{initial}$ with $s, s_{initial} \in S$, $r \in R$ and R is the set of sequences for which each sequence selects a specific state

$Z = \{A_{components,r}, R_{components,r}, A_{coupling}, R_{coupling}\}$ is the set of decisions where
$A_{components,r}$ adds the component $\eta(s, r)$
$R_{components,r}$ removes a component
$A_{coupling}$ adds a coupling
$R_{coupling}$ removes a coupling

$\mathcal{Z}$ is the set of all k-tuples $(a_1, a_2, \ldots, a_k)$ with $a_1, a_2, \ldots, a_k \in Z$

$\rho : P \times \mathcal{Z} \rightarrow P$ is the is the branching function
$\rho(p_i, b_j) = p_h$ where
$(p_i, p_h \in P) \wedge (p_i <_{t,z,L} p_h)$, $b \in \mathcal{Z}$

The structure $\mathcal{F}$ consists of eight elements: system models P, a tree A, a set of sets of components $\mathcal{L}$, a component transformation function ζ, an initialization function η, a set of decisions Z, a set of all sequences of decisions $\mathcal{Z}$, and a branching function ρ. To follow the design process, we need to have the ability to branch a model into a couple of new models that contain the possibilities one wants to assess using a predefined set of components.

A set of sets of components $\mathcal{L}$ is provided during the design process. Each node, from the set of possible nodes P, has a model constructed using components from a set of components. A component transformation function ζ transforms a model into another, each of which contains a set of components.

The branching function ρ constructs a new model from a (existing) model and a sequence of decisions in $\mathcal{Z}$. The set of decisions applies to the model structure. Structural changes need four operations: add or remove components, and add or remove couplings. Adding components needs the initialization function η that sets the initial state of the model from the set of states S.

The models generate by the branching function are organized in a tree structure A. The root represents the initial model. Each node contains an alternative design (or model) which is derived from its parent node by applying some change. The leaves of the tree contain models with no further change. The simulation of each model (in the tree) generates outputs, i.e., the KPIs, of a design alternative. The range of each KPI is related to the model's position in the tree; namely, a less detailed model has outputs with a broader range, and a more detailed model has outputs with a narrower range.

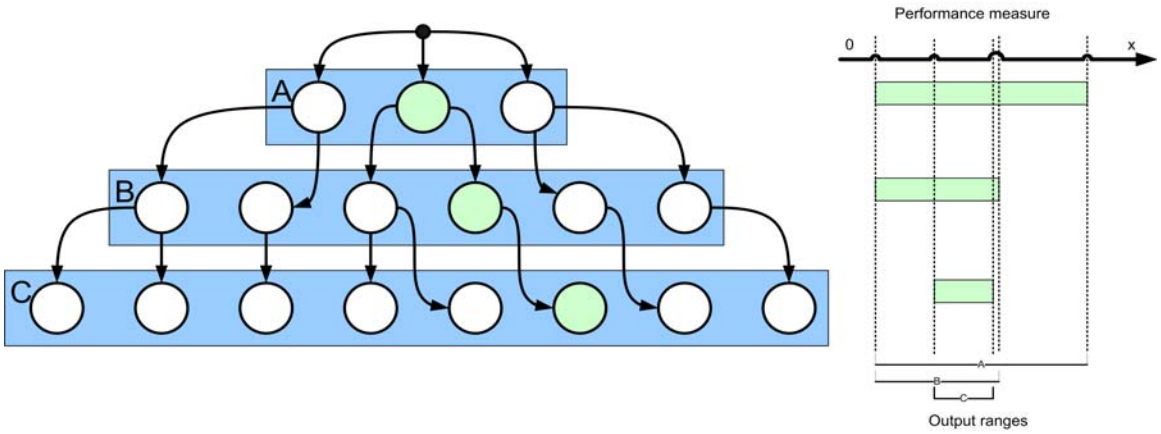


Figure 2: Output Ranges of the Model in the Tree Structure (Fumarola 2011)

The designs (or models) closer to the root of the tree contain more uncertainties designers have during the early design phases. Further down the tree, more details are added so that they reduce the solution space. This relation is exemplified in Figure 2 which has designs of three levels.

The tree guides actors or designers through the design process. At each level of the tree, an appropriate class of decisions needs to be made to refine the current design. Each decision leads the designers to a lower level node that contains a refined design that is derived from a higher level node. This helps the comparison of the designs at every level. The designers can steer the design process by choosing decisions that can reduce the KPI range of their interest. The design process will in turn gradually refine the design and narrow the outputs to a desired combination of KPI ranges.

5. APPLYING THE FRAMEWORK TO CONTAINER TERMAIL DESIGN

The SDB framework is applied to studies in automated container terminal design (Fumarola et al. 2010, Fumarola 2011). In the studies, simulation component libraries are integrated with a CAD design environment so that simulation models can be directly generated and experimented from the design alternatives. Using the tree structure to organize and guide the design process showed promising results.

5.1. Simulation Components

The model library contains a collection of DEVS components that represent a large variety of material handling equipments, such as quay cranes, automated guided vehicles and rail mounted gantries that are commonly used in automated container terminals. The components can be divided into two groups: the high level components and the low level components. They specify the equipments at high or low abstraction level. Relations are defined between the high level and the low level models. A dynamic build-up mechanism can add and remove sub-models to or from a model, and establish and remove couplings correspondingly. The simulation outputs are collected for visualizations and statistics as design evaluation criteria.

5.2. The Design Process

The design process, as shown in Figure 3, starts with the design requirements specification and the collection of documentations about the equipment, terminal plot, demand forecasts, etc. These documents define the amount of investment, the physical constrains and the required throughput. With the information, designers can identify a list of evaluation criteria. It is needed at a later design stage to compare alternative designs using multi-criteria decision analysis.

Once the criteria list is prepared, the first high level design can be constructed. At this level, the models are relative simple so the simulation rum time is in general short. Designers can create a high number of designs with relative low experimentation cost. Using the results of the high level models, the designs can be refined by using the low level models, allowing for more detailed analysis. The low level models can be used (in the container terminal design studies) to assess the terminal layout and the choice of equipments. The alternatives can differ in cost, performance and environmental footprint. The position, number and type of the equipments are decided at separate design levels. This allows the designers to concentrate on one decision factor at a time.

The tree structure used in the framework offers users the possibility to return to the higher design levels, alter their decisions, experiment with them, and discover the interactions between them. At each level, designers can evaluate the designs using multi-criteria decision analysis. When they had assessed different design alternatives based on their individual and collective preferences, they can focus on a set of final designs. That may conclude the design process.

5.3. Evaluation with Students and Experts

5.3.1. Usability Test with Students

The evaluation was first conducted with nine graduate students (at Delft University of Technology) for the usability tests of the tools of the framework. The students were divided into three groups, and each group participated in a design workshop where they were given a design task.

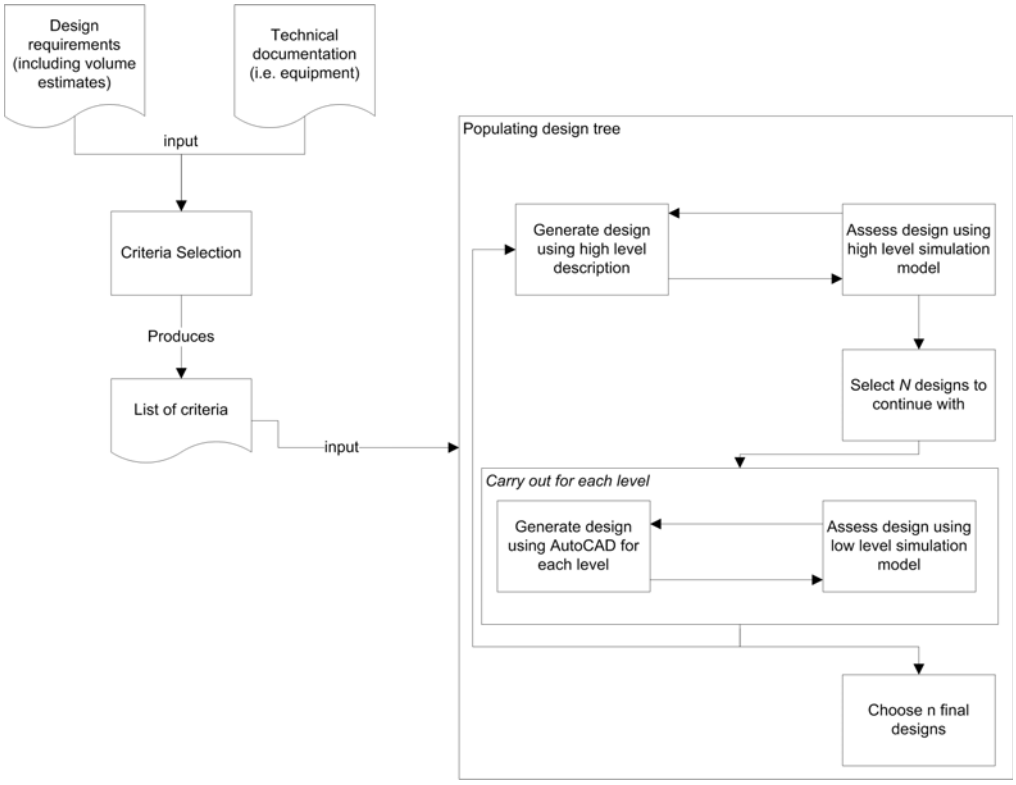


Figure 3: The Design process Using the SBD Framework (Fumarola 2011)

Each student was assigned a different role with specific design goals. Semi-structured group interviews were given to the participants after each workshop.

The students indicated that the tool guided them in their design effort and helped them in structuring design decisions. They gave a higher score to the support by statistics than by animation. As novice users, they had problems in using the AutoCAD design environment. Because conventions are strictly required for the CAD designs in order to transform them into simulation models, the errors had to be corrected first to further proceed with the design process. To solve this problem, formal CAD model checking mechanisms have to be added into the simulation environment. The students also suggested that a more extensive set of statistics could be added to provide more detailed views on the performances of the designs.

The students' design processes were given to an expert (in automated container terminals) to ask for opinions. The expert noticed that the students' design process complied with good design process practices in which the first level designs were extensive, and from those one or two designs were chosen to be continued. The expert further commented that in current practices, unfortunately, limited by resources and costs, only a few design alternatives (typically two) were compared. The comparison of a larger number of high level designs is appreciated by the expert as it allowes the exploration for non-trivial designs.

5.3.2. Effectiveness Test with Experts

Semi-structured interviews were opted for the expert evaluation. Eight domain and methodological experts participated in the interviews. The questions focused on several aspects: the interviewee's level of expertise; the perceived problems they have encountered in their organization; the proposed framework, and possible adoption of the framework in their current organization.

The experts stated that simulation models are frequently used to support decision making in logistics management. However, in the design process, because designers do not always understand the model (partially or fully), they are reluctant to use them. When the designers are presented with models that do not conform to their expectations, they will disregard the models and continue the design process without simulation. Several interviewees pointed out that the introduction of SBD (that uses simulation models) *would* therefore meet some resistance. The resistance is due to "fear of the unknown" instead of any intrinsic quality of a design method.

On the other hand, they also pointed out that simulation models *should* be used earlier and oftener in the design process. An earlier feedback loop is mostly missing in current design approaches. And modern logistics systems are often too complex to comprehend without the support of simulation models. The proposed framework has the potential to solve these problems.

They acknowledged that the SDB framework, particularly the tree structure in organizing different designs, can help to structure the design process and to alter among different design abstractions. However, the actors in the design process should be fully aware of the type of questions that can be answered in every level of the process.

Issues		Usability Workshop	Expert Evaluation
1	*Need a structured design process*	The approach helped participants follow a structured design process.	The approach is perceived useful in collaboration and comparing results in a structure way.
2	*Manage different system designs*	Participants were capable of comparing different designs.	Quantitative comparison of alternatives helps collaboration between different actors.
3	*Document the process*	(not evaluated)	Storing a design process can be very useful for future reference.
4	*Compare different designs*	Participants used the comparison to make design decisions.	Quantitative comparison of alternatives helps collaboration between different actors.
5	*Mutually consistent conceptual and detailed designs*	(not evaluated)	Attention should be given to the choice between analytical models (high level) and simulation models (detailed design).
6	*Involvement of actors throughout the process*	Each participant was involved in the design process.	Collaboration is required throughout a design process.
7	*Predict the influence of design decisions*	The results from the simulation models were used to make comparisons.	KPIs help streamline the discussion between several actors.
8	*Understandable presentation*	The 3D and statistical output were used to compare designs.	(not evaluated)
9	*Interactive and iterative process*	Participants were able to design collaboratively.	Iteration is required in a design process.

Table 1: Evaluation of the Method with Issues Identified in Section 2.3

They acknowledged that the SDB framework, particularly the tree structure in organizing different designs, can help to structure the design process and to alter among different design abstractions. However, the actors in the design process should be fully aware of the type of questions that can be answered in every level of the process.

Using simulation to guide the design process also has benefits from a collaboration perspective. The proposed framework could support a collaborative design process. One domain expert phrased this as such: when you start to consider various dimensions of a design, finding a common communication mechanism among the actors is very important; this framework could achieve this goal. Many projects are not about finding the optimal solution but to let the actors understand the problem and make decisions that are good enough to suffice the agreed-upon requirements. A documented and structured design process can serve as a knowledge management system. The proposed framework has the potential to be used for large-scale design project in other domain, particularly engineering projects. To provide an overview, a brief summary of the evaluation is listed in Table 1.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we presented a SBD framework. In a multi-actor design environment, pure hard systems approaches showed limitations in involving different actors. Some common issues were discussed. The proposed framework is composed of four constructs that address the identified issues. It guides users to follow a tree-like design refinement that specifies and evaluates the designs in more detail at each level. The evaluation showed that the framework and the supporting tools suit well in a multi-actor environment, and they allow interactive and structured exploration of the design solution space.

Further research can be conducted to post-process the simulation results to meet the needs of individual actors from multiple perspectives, and to extend the simulation component library accordingly.

REFERENCES

Ackoff, R. L. (1979). The future of operational research is past. *The Journal of the Operational Research Society 30*, 93–104.

Checkland, P. (1999). *Systems Thinking, Systems Practice: Includes a 30-Year Retrospective*. John Wiley & Sons.

den Hengst, M., G.-J. de Vreede, and R. Maghnouji (2007). Using soft OR principles for collaborative simulation: a case study in the Dutch airline industry. *Journal of the Operational Research Society 58*, 669–682.

Dwarakanatha, S. and K. M. Wallacea (1995). Decision-making in engineering design: observations from design experiments. *Journal of Engineering Design 6*(3), 191–206.

Fumarola, M. (2011). *Multiple Worlds: A multi-actor simulation-based design method for logistics systems*. Ph. D. thesis, Delft University of Technology, The Netherlands.

Fumarola, M., Y. Huang, C. Tekinay, and M. D. Seck (2010). Simulation-based design for infrastructure system simulation. In G. K. Janssen, K. Ramaekers, and A. Caris (Eds.), *Proceedings of The 2010 European Simulation and Modelling Conference*, Belgium, pp. 288–293. Eurosis-ETI.

Fumarola, M., M. D. Seck, and A. Verbraeck (2010). A DEVS component library for simulation-based design of automated container terminals. In *SIMUTools '10: Proceedings of the 3rd International ICST Conference on Simulation Tools and Techniques*, Belgium, pp. 1–7. ICST.

Girod, M., A. C. Elliott, N. D. Burns, and I. C. Wright (2003). Decision making in conceptual engineering design: An empirical investigation. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 217*(9), 1215–1228.

Goel, V. (1995). *Sketches of Thought*. MIT Press.

Hoover, S. P. and J. R. Rinderle (1991). Models and abstractions in design. *Design Studies 12*, 237–245.

Huang, Y., M. D. Seck, and A. Verbraeck (2011). From data to simulation models: Component-based model generation with a data-driven approach. In S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu (Eds.), *Proceedings of the 2011 Winter Simulation Conference*, Phoenix, AZ, pp. 3724–3734. IEEE, Piscataway, NJ.

Law, A. M. (2007). *Simulation Modeling and Analysis* (4th ed.). McGraw-Hill.

Mingers, J. and J. Rosenhead (2004). Problem structuring methods in action. *European Journal of Operational Research 152*, 530–554.

Munro, I. and J. Minger (2002). The use of multimethodology in practice: results of a survey of practitioners. *Journal of the Operational Research Society 53*, 369–378.

Paredis, C., A. Diaz-Calderon, R. Sinha, and P. Khosla (2001). Composable models for simulation-based design. *Engineering with Computers 17*(2), 112–128.

Robinson, S. (2001). Soft with a hard centre: Discrete-event simulation in facilitation. *The Journal of the Operational Research Society 52*, 905–915.

Saanen, Y. A. (2003). *An approach for designing robotized marine container terminals*. Ph. D. thesis, Delft University of Technology, The Netherlands.

Shannon, R. E. (1975). *Systems simulation: the art and science*. Prentice Hall, Inc.

Simon, H. A. (1996). *The Sciences of the Artificial* (3rd ed.). MIT Press.

Sol, H. G. (1982). *Simulation in Information Systems Development*. Ph. D. thesis, University of Groningen, The Netherlands.

Ullman, D. G., T. G. Dietterich, and L. A. Stauffer. (1998). A model of the mechanical design process based on empirical data. *Artificial Intelligence for Engineering, Design, Analysis and Manufacturing 2*(1), 33–52.

Zeigler, B. P., H. Praehofer, and T. G. Kim (2000). *Theory of Modeling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems* (2nd ed.). Elsevier/Academic Press.

OPTIMAL ALLOCATION OF ECONOMIC RESOURCES USING THE AHP ABSOLUTE MODEL

Fabio De Felice[a], Antonella Petrillo[b], Michele Tricarico[c]

[a] Department of Civil and Mechanical Engineering - University of Cassino, Italy
[b] Department of Civil and Mechanical Engineering - University of Cassino, Italy
[c] Horseracing Italian Agency, Italy

[a]defelice@unicas.it, [b]a.petrillo@unicas.it

ABSTRACT
One often wonders why more people in organizations do not rush today to use a formal decision-making approach to make their complex decisions. A strange thing about people is that they value money and other valuable resources over their own loosely defined and not well-organized subjective value systems. This paper places special emphasis on the measurement of intangible criteria and on their incorporation into the allocation process through a proper decision making approach. The purpose of decision-making is to help people make decisions according to their own understanding. In this paper, a well know decision-making method, the Analytic Hierarchy Process (AHP) is applied to identify a quality model to evaluate Italian racecourses performances based upon the criteria: Quality organization of Racing, Infrastructure and Equipment, Attractiveness and Management Skill. The main conclusion is that the AHP model adopted can manage all the information of the real-world problem.

Keywords: Analytic Hierarchy Process, DSS, Performances, Racecourse

1. INTRODUCTION

In today's global economy, characterized by a dynamic and volatile environment, many researchers stress the importance of international location factors (Badri and Davis, 1995). Some of the issues associated with global expansion and location include multiple political, economic, legal, social, and cultural environments. Location-allocation decisions involve a substantial capital investment and result in long-term constraints on production and distribution of goods (Strebel, 2003; Seeley, 2002). These problems are complex and, like most real world problems depend upon a number of tangible and intangible factors which are unique to each problem (De Felice and Petrillo, 2010 a). The complexity stems from a multitude of quantitative and qualitative factors influencing location choices as well as the intrinsic difficulty of making numerous trade-offs among those factors (De Felice and Petrillo, 2012). One analytical approach often suggested for solving such a complex problem is the Analytic Hierarchy Process (AHP) introduced by Saaty (Saaty, 1980). The AHP enables the decision maker to structure a complex problem in the form of a simple hierarchy and to evaluate a large number of quantitative and qualitative factors in a systematic manner under conflicting multiple criteria. It is developed and designed to solve complex problems involving multiple criteria. It is a highly flexible decision methodology that can be applied in a wide variety of situations (De Felice, Petrillo and Silvestri, 2012). There are two types of measurement involved in the AHP, *absolute* and *relative*. The first requires a standard with which to compare elements, but mostly alternatives at the bottom of the hierarchy. The process leads to absolute preservation in the rank of the alternatives no matter how many are introduced. The second is based on paired comparisons among the elements of a set with respect to a common attribute. This process is essential for comparing intangible attributes for which there are no agreed upon measures. At the level of alternatives new elements (i.e. alternatives) do introduce new information generated by the changing number in the set and by their measurement which essentially rescales the criteria and hence can lead to reversals of previous rank orders. Absolute measurement is used on standardized problems whereas relative measurement is used in new learning situations (Saaty, 2005). Absolute method is typically used in a decision situation, which involve selecting one (or more) decision alternatives from several "candidate" decision alternatives on the basis of multiple decision criteria of a competing or conflicting nature (McCarthy, 2000). In this paper, we have developed a case study on racecourse performance appraisal *using AHP absolute model*. Though AHP has been applied in numerous real settings, but there isn't evidence that AHP has been applied in racecourse performance evaluation (De Felice and Petrillo, 2010 b). This paper attempts to fill up the gap. The aim of our paper is to explain, through a real case study, the uses of multi-criteria prioritization in resource allocation, and in particular the use of absolute measurement in the optimal assignment of economic resources.

2. THE PROGRESSION OF THE PLANNING MODEL

A feature of absolute measurement AHP is that the scale for each lowest level criterion consists of indicator categories (e.g. A, B, C, etc.). Thus, the alternatives

consist of the these categories or grades. Absolute measurement AHP requires a pairwise comparison procedure between indicator categories (for each lowest level criterion) to establish the relative weights for these categories using eigenvector approach (Park and Lim, 1999). In other words in absolute measurement the properties of an element are compared or "rated" against a standard (Leskinen, 2000). In this method an element is compared against an ideal property; i.e. a "memory" of that property (Saaty *et. al*, 2003). Generally, only the final alternatives of choice are measured absolutely. For example, students applying for admission are rated on grades, letters of recommendation and standardized test scores. A student's final rating is the weighted sum of the ratings on the various criteria (De Felice and Petrillo, 2011).
Here below and in Figure 1 are the steps of absolute measurement process adopted:

- **Step 1:** Identify the criteria, subcriteria and alternatives (to be evaluated) for evaluation and put them into the AHP hierarchy.
- **Step 2:** Calculate the weights of the decision criteria and subcriteria by the relative measurement of AHP, i.e., construct the pairwise comparison matrix for all the criteria and compute the normalized principal right eigenvector of the matrix. This vector gives the weights of the criteria.
- **Step 3:** Divide each subcriteria into several intensities or grades. Set priorities on the intensities by comparing them pairwise under each subcriteria. Multiply these priorities by the priority of the parent subcriteria.
- **Step 4:** Take one alternative at a time and measure its performance intensity under each subcriteria.

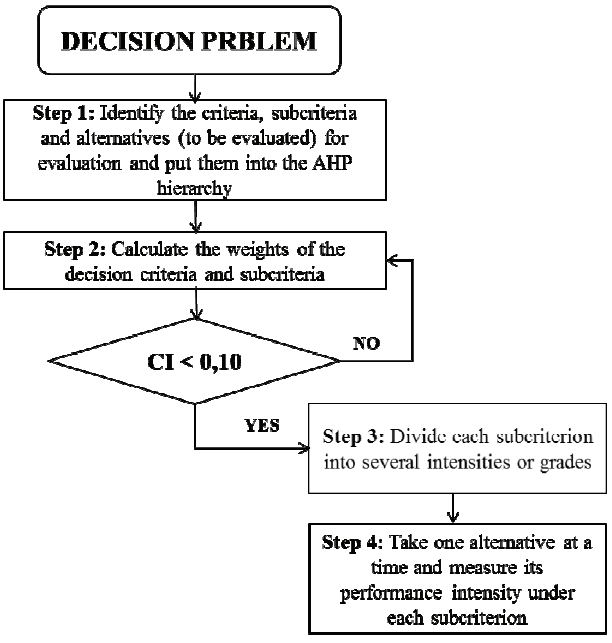


Figure 1: Methodological approach

3. THE CASE STUDY

In this paragraph we will analyze the AHP model adopted in order to rank racecourses quality performance.

3.1. STEP 1: Identify the criteria, subcriteria and alternatives

We developed the following AHP Model to determine the criteria and subcriteria weights. In figure 1 is shown AHP Model (see Appendix A).

Regards the alternatives we selected 40 different Italian racecourses.

Here below in the following tables (Table from 1 to 5) is the description of criteria and subcriteria.

Table 1: Criteria - Description

Criteria	Description
C1	The ability to organize spectacular and corrected races.
C2	The availability of appropriate facilities and equipment maintained in good condition
C3	The ability to attract and retain customers
C4	The adoption of policies that introduce the culture associated with the horse culture of corporate values (integrity, ethics, competitiveness, investment)

Table 2: SubCriteria C1 - Description

SubCriteria C1	Description
C1.1	Evaluation of the ability to organize and plan races with a reasonable number of participants by offering an enjoyable spectacle
C1.1	Assessment of the ability to encourage the creation of a field starters balanced in order to ensure an enjoyable and profitable spectacle from the point of view of the bets
C1.3	Technical evaluation of horses winning
C1.4	Evaluating the timeliness of races (deadlines and procedures, in the absence of accidents)
C1.5	Disciplinary reports for each racecourse

Table 3: SubCriteria C2 - Description

SubCriteria C2	Description
C2.1	Areas dedicated to horse racing (sum of all areas of race tracks); parameters to consider are the type (sand, grass, synthetic), size (length x width)
C2.2	Area devoted to public aims such as parterres, bars, parks, restaurants, or other structures with free access for public
C2.3	Structure of the relevance of the hippodrome, independent and external to it, continuously and exclusively dedicated to training, including trails and picnic areas (parameter indicative of the right size: relationship to runways. / N.box)
C2.4	Areas which include services for owners / operators such as surgery, dining room, lunchroom, etc.. Areas which include services such as veterinary clinic for horses, garage available, etc.
C2.5	Number and type of racecourse facilities support activities, such as runway lighting system, timekeeping system, TV system
C2.6	Indicator characterizing the value of the plant

Table 4: SubCriteria C3 - Description

SubCriteria C3	Description
C3.1	Evaluation of the ability to make equestrian events and horse shows, contemporary and otherwise, who play a role call for the competitive event. The initiatives should be compatible and complementary to the races, including through the promotion of culture horseracing courses (promotion)
C3.2	Evaluation of plant capacity to attract bettors
C3.3	Evaluation of attractiveness on the betting market
C3.4	Number and types of services appropriate and welcoming to the public
C3.5	Importance of the plant economy tradition of horse racing
C3.6	Evaluation of the ability of attracting the public

Table 5: SubCriteria C4 - Description

SubCriteria C4	Description
C4.1	Certification of financial statements
C4.2	Achievement of certification by recognized organizations: Quality ISO 9000, ISO 14001 Environment, OHSAS 18001 Safety
C4.3	Production of a document demonstration of the ability to generate social values in the local context, linked to the economic value of the (animal protection, employment, etc.)
C4.4	Assessment of financial strength
C4.5	Statement of annual expenditure
C4.6	Company's ability to have other forms of financing including sponsorship and related activities

3.2. STEP 2: Calculate the weights of the decision criteria and subcriteria

In this phase were developed pairwise comparison matrices to determine the criteria and subcriteria weights. In Appendix B are shown the pairwise comparison matrices (figure 2, 3, 4, 5 and 6). In the AHP paired comparisons are made with judgments using numerical values taken from the AHP absolute fundamental scale of 1-9. A scale of relative values is derived from all these paired comparisons and it also belongs to an absolute scale that is invariant under the identity transformation like the system of real numbers . After all pairwise comparison the consistency index (CI) of the derived weights was calculated by Equation (1):

$$CI \frac{\lambda_{max} - n}{n-1} \quad (1)$$

In general, if CI is less than 0.10, satisfaction of judgments may be derived. Here below in the following figures (figure 7, 8, 9, 10 and 11) are shown the weights derived from pairwise comparison for each criteria and subcriteria.

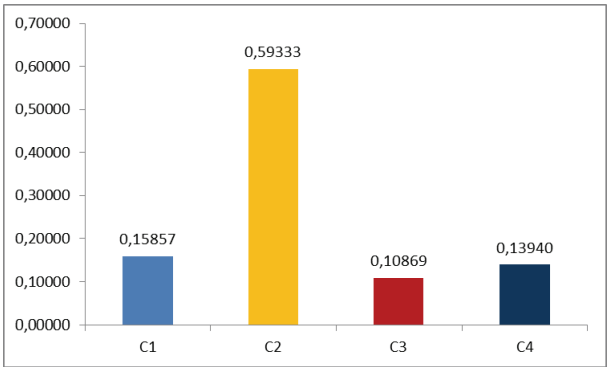


Figure 7: Weights for each criteria

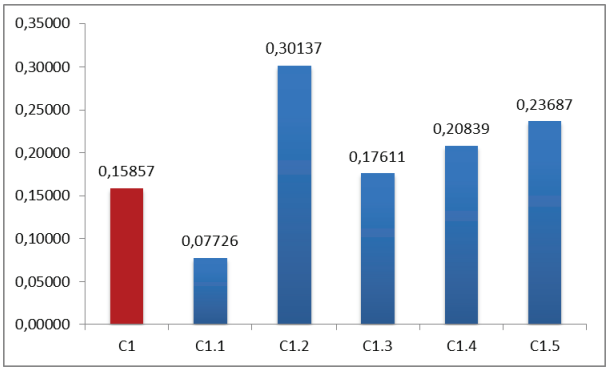


Figure 8: Weights for subcriteria C1

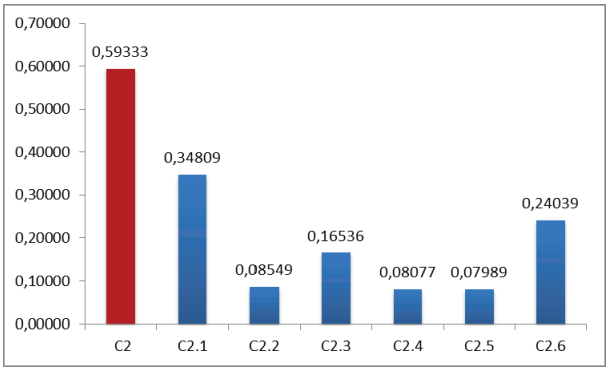


Figure 9: Weights for subcriteria C2

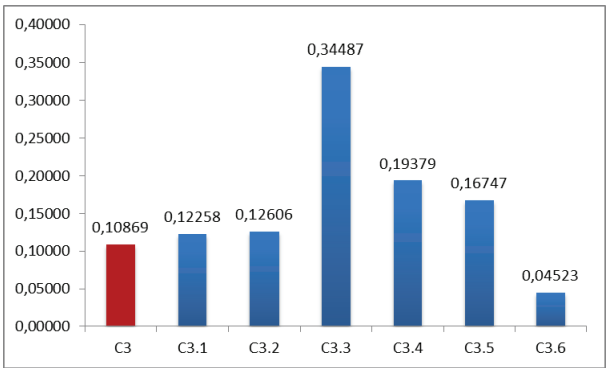


Figure 10: Weights for subcriteria C3

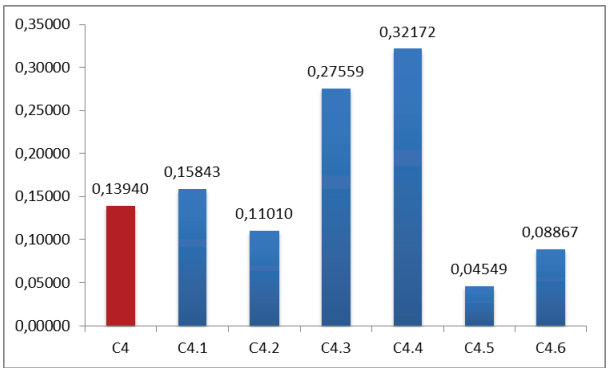


Figure 11: Weights for subcriteria C4

In appendix C are show weights for all criteria and subcriteria.

3.3. STEP 3: Divide each subcriteria into several intensities or grades

In this step each subcriteria are further subdivided into a level for intensities. Each criterion has ratings listed under it. An example would be to take a criterion of cost and list under it “very high”, “high”, “average”, and “low”. These are ratings that are then prioritized to determine their relative importance. The type and number of ratings for each criterion may be different. An intensity is a range of variation of a criterion that enables one to distinguish the quality of an alternative for that criterion. An intensity may be expressed as a numerical range of values if the criterion is measurable or in qualitative terms (Saaty, et al. 2007; Rafikul and Mohd Rasad, 2005).

For example, the evaluation criteria for “ C1.1 - Average of the horses left for race” have the following intensities (see Table 6):

Table 6: Example of criteria intensity

Score	from	to
1	0	<7
3	7	<8
5	8	<10
7	10	<12
9	12	up

We set priorities for the criteria by comparing them in pairs. We then pairwise compare the intensities according to priority with respect to their parent criterion C1 (Table 7). The priorities of the intensities are divided by the largest intensity for each criterion and subcriteria to put it in the ideal mode

Table 7: Comparing Intensity of C1.1

	0-<7	7-<8	8-<10	10-<12	12-up	priority
0-<7	1	3	5	7	9	0,493
7-<8	1/3	1	3	5	7	0,255
8-<10	1/5	1/3	1	9	7	0,167
10-<12	1/7	1/5	1/9	1	9	0,061
12-up	1/9	1/7	1/7	1/9	1	0,024

Table 7 gives a comparison of the intensities for C1.1. The other intensities are similarly compared.

Table 8: The ideal intensity Mode

	Priorities Weighted by C1 and C1.1	Divide by largest value
0-<7	(*) 0,006038	0,495597
7-<8	0,012184	1
8-<10	0,004659	0,382407
10-<12	0,002032	0,166781
12-up	0,000897	0,073596
(*) 0,493x 0.158x0.0772 =0,006038		

Table 8 gives the ideal intensity mode for C1 and C1.1. The other ideal intensities mode are similarly obtained.

3.4. STEP 4: Take one alternative at a time and measure its performance intensity under each subcriteria

In this step, finally, we rate each alternatives by assigning the intensity rating that applies to them under each criterion (Table 9). The scores of these intensities are each weighted by the priority of its criterion and summed to derive a total ratio scale score for the alternative.

Table 9: Ranking

Name	Ideals	Normals
AB	0.501236	0.023535
AC	0.465246	0.021845
AD	0.728787	0.034220
AE	0.255768	0.012009
AF	0.463271	0.021753
AG	0.429542	0.020169
AH	0.634037	0.029771
AI	0.490323	0.023023
AJ	0.404723	0.019003
AK	0.456050	0.021414
AL	0.187590	0.008808
AM	0.375334	0.017624
AN	0.524765	0.024640
AO	0.560826	0.026333
AP	0.601429	0.028240
AQ	0.604732	0.028395
AR	0.785092	0.036863
AS	1.000.000	0.046954
AT	0.610738	0.028677
AU	0.462112	0.021698
AV	0.603960	0.028359
AW	0.520404	0.024435
AX	0.648255	0.030438
AZ	0.556822	0.026145
BA	0.444911	0.020891
BB	0.909919	0.042725
BC	0.382625	0.017966
BD	0.241810	0.011354
BE	0.879245	0.041284
BF	0.624901	0.029342
BG	0.684889	0.032159
BH	0.460543	0.021625
BI	0.501866	0.023565
BJ	0.444563	0.020874
BK	0.460436	0.021619
BL	0.252033	0.011834
BM	0.499912	0.023473
TRT	0.459184	0.021561
TVO	0.569757	0.026753
VAS	0.609651	0.028626

4. CONCLUSIONS

Economic allocation resources is more complex and risky due to uncertainty and volatility of international environments. The global location-allocation decision process involves qualitative as well as quantitative factors. The decision-makers can no longer ignore the influence of highly judgmental and sensitive factors such as the political situation, global competition and survival, government regulations, and economic factors. In this context our aim is to develop a flexible decision model in order to cope with the changes.

On the other hand the use the use of absolute (rather than relative) scales for scoring the alternatives provides the following advantages:

- The addition of a new alternative doesn't require new pairwise comparisons with all other alternatives;
- There is no potential for rank reversal with addition or deletion of the alternative;
- As an alternative is added or deleted or as its score changes, the scores of all other alternatives remain the same.

We can conclude that this approach can be used whenever it is possible to set priorities for intensities of criteria; people can usually do this when they have sufficient experience with given operation. In addition, one can use this approach to rate many alternatives but then choose the top few and perform paired comparisons on them directly with respect to the criteria by deleting the intensities from hierarchy.

REFERENCES

Badri, M., D. Davis, 1995. Decision support models for the location of firms in industrial sites, *International Journal of Operations and Production Management,* 15 (1), 50-62.

De Felice, F., Petrillo, A., Silvestri, A., 2012. Multi-criteria risk analysis to improve safety in manufacturing systems. *International Journal of Production Researc*h. ISSN 0020–7543 print/ISSN 1366–588X online. pp. 1-16.

De Felice, F., Petrillo, A., 2012. Hierarchical model to optimize performance in logistics policies: multi attribute analysis. *The 8th International Strategic Management Conference. Elsevier Procedia Social and Behavioral Sciences*.

De Felice, F., Petrillo, A., 2011. Methodological Approach for Performing Human Reliability and Error Analysis in Railway Transportation System. *International Journal of Engineering and Technology,* Vol.3 (5), 2011, 341-353

De Felice, F., Petrillo, A., 2010 a. A multiple choice decision analysis: an integrated QFD – AHP model for the assessment of customer needs.

International Journal of Engineering, Science and Technology, Vol. 2, No. 9, 2010, pp. 25-38.

De Felice, F., Petrillo, A., 2010 b. A new multicriteria methodology based on Analytic Hierarchy Process: the "Expert" AHP. *International Journal of Management Science and Engineering Management*, 5(6): 439-445, 2010.

Leskinen P., Measurement Scales and Scale Independence in the Analytic Hierarchy Process. *Journal of Multi-criteria decision Analysis*, 2000.

Kyung S. Park, Chee Hwan Lim, 1999. A structured methodology for comparative evaluation of user interface designs using usability criteria and measures. *International Journal of Industrial Ergonomics*, 23, 379-389.

McCarthy, J., 2000. How to Conduct Productive Performance Appraisals. *Journal of Property Management*, 22-25.

Rafikul, I., Mohd Rasad S.. Employee performance evaluation by AHP: A case study. *ISAHP 2005*, Honolulu, Hawaii, July 8-10, 2005.

Saaty, T.L., 1980. *The Analytic Hierarchy Process*. Third ed. McGraw-Hill, New York.

Saaty, T.L., 2005. *Theory and Applications of the Analytic Network Process: Decision Making with Benefits, Opportunities, Costs, and Risks*. RWS Publications, 4922 Ellsworth Ave., Pittsburgh, PA, 2005, p. 15213.

Saaty, T.L., Peniwati, K., Shang, Jen S., 2007. The analytic hierarchy process and human resource allocation: Half the story. *Mathematical and Computer Modellin*g 46, 1041–1053.

Evans, W.A., 1994. Approaches to intelligent information retrieval. *Information Processing and Management*, 7 (2), 147–168.

Saaty, T.L., Vargas, L.G., Dellmann, K. (2003). The allocation of intangible resources: the analytic hierarchy process and linear programming, Socio-Economic Planning Sciences 37 (2003) 169–184.

Seeley R., Stephens T., Tate P.: *Essentials of Anatomy and Physiology*. McGraw-Hill, 2002.

Strebel B. J., *The Manager's Guide to Effective Meetings*. McGraw-Hill, 2003.

AUTHORS BIOGRAPHY

Fabio De Felice, Professor at the Faculty of Engineering of the University of Cassino (Italy), board member of several international organizations and responsible for scientific research and training in industrial plants.

The scientific activity developed through studies and researches on problems concerning industrial plant engineering. Such activity ranges over all fields from improvement of quality in productive processes to the simulation of industrial plants, from support multicriteria techniques to decisions (Analytic Hierarchy Process, Analytic Network Process), to RAMS Analysis and Human Reliability Analysis. The main university courses in which he is involved are: Safety of Industrial Plants, Industrial Production Management, Industrial Simulation, Human Reliability Analysis.

He is author of several books, and papers in international journals and conference proceedings. He has also been a member of several editorial boards. In addition he is founder of the AHP Academy - International Association for the promotion of multi-criteria decision making methods.

Antonella Petrillo, degree in Mechanical Engineering, now PhD at the Faculty of Engineering of University of Cassino where she conducts research activities on Multi-criteria decision analysis (MCDA), industrial plant, safety, supply chain and quality management at the Department of Civil and Mechanical Engineering.

Michele Tricarico, degree in Electronical Engineering at University of Rome "Tor Vergata" is PMP certified Employee at Horseracing Italian Agency. He was Project Manager at SAP and he was Consultant for several company such a Siemens, Accenture, etc.

APPENDIX A

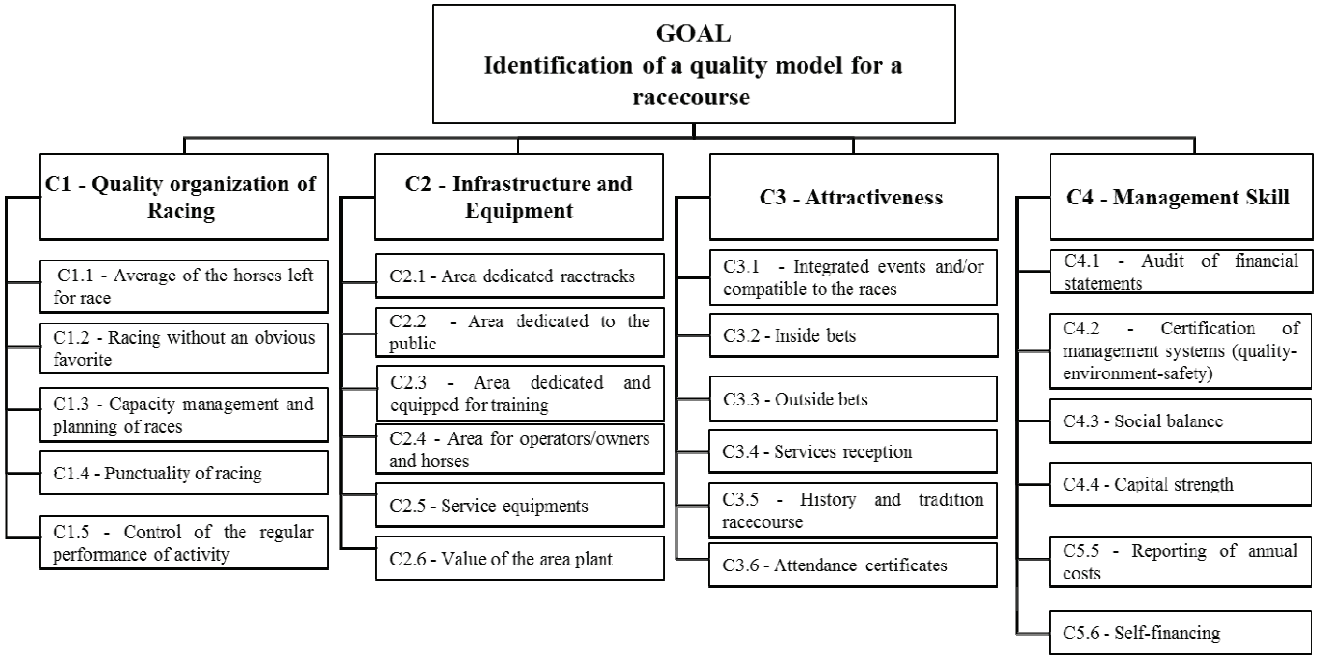


Figure 1: AHP Model

APPENDIX B

C1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1**
C1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1**
C1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1**
C2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2**
C2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2**
C3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3**
CI = 0.07193																		

Figure 2: Criteria - Pairwise comparision

C1.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1.2**
C1.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1.3**
C1.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1.4**
C1.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1.5**
C1.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1.3**
C1.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1.4**
C1.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1.5**
C1.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1.4**
C1.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1.5**
C1.4	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C1.5**
CI = 0.06506																		

Figure 3: Subcriteria C1 - Pairwise comparision

C2.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.2**
C2.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.3**
C2.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.4**
C2.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.5**
C2.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.6**
C2.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.3**
C2.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.4**
C2.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.5**
C2.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.6**
C2.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.4**
C2.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.5**
C2.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.6**
C2.4	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.5**
C2.4	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.6**
C2.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C2.6**
CI = 0.09894																		

Figure 4: Subcriteria C2 - Pairwise comparision

C3.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.2**
C3.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.3**
C3.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.4**
C3.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.5**
C3.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.6**
C3.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.3**
C3.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.4**
C3.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.5**
C3.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.6**
C3.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.4**
C3.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.5**
C3.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.6**
C4.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.5**
C4.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.6**
C3.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C3.6**
CI = 0.09694																		

Figure 5: Subcriteria C3 - Pairwise comparision

C4.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.2**
C4.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.3**
C4.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.4**
C4.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.5**
C4.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.6**
C4.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.3**
C4.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.4**
C4.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.5**
C4.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.6**
C4.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.4**
C4.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.5**
C4.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.6**
C4.4	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.5**
C4.4	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.6**
C4.5	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	**C4.6**
CI = 0.09755																		

Figure 6: Subcriteria C4 - Pairwise comparision

APPENDIX C

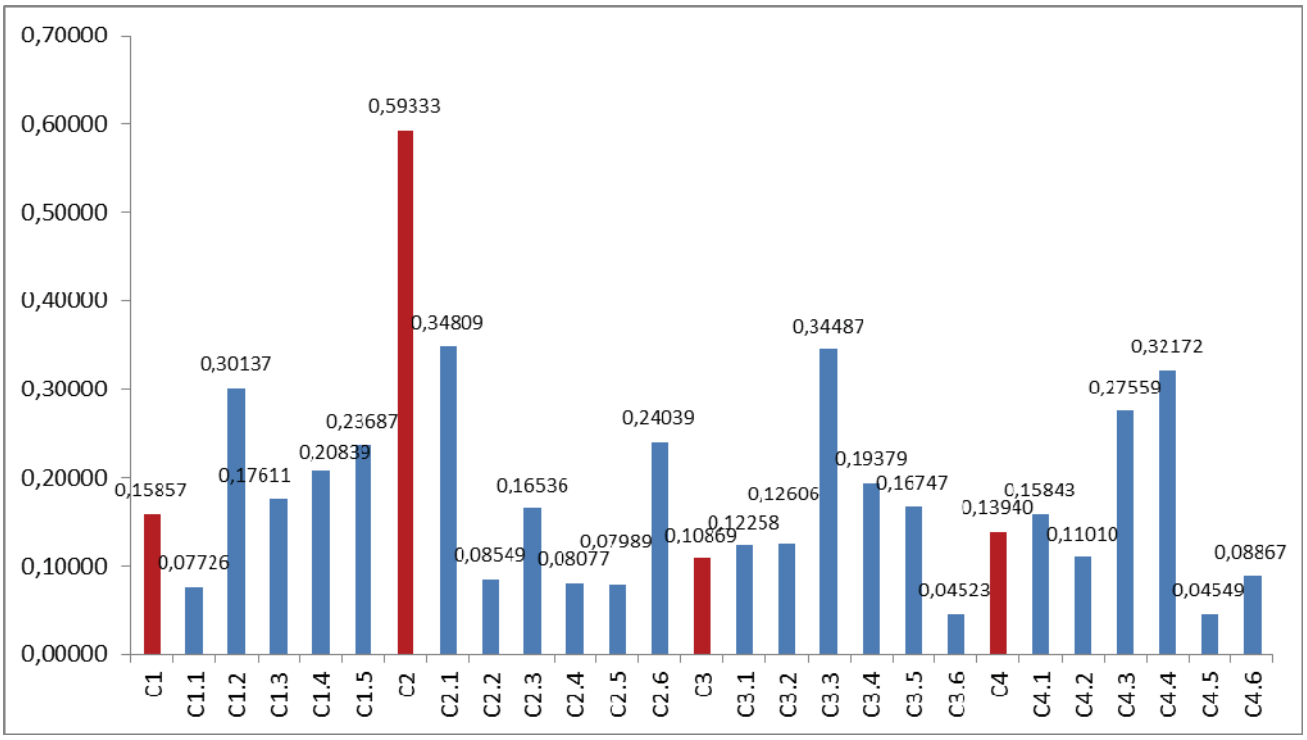


Figure 12: Weights for all criteria and subcriteria

ULTRA-FAST REGISTRATION OF 2D ELECTRON MICROSCOPY IMAGES

Santiago García[(a)], Julio Kovacs [(b)] and Pablo Chacón [(a)]

[(a)] Department of Biological Physical Chemistry. Rocasolano Physical Chemistry Institute, CSIC, Serrano 119, Madrid 28006, Spain.

[(b)] QG Research, Los Angeles, CA

[(a)]pablo@chaconlab.org; sgarcia@iqfr.csic.es

ABSTRACT

Registration of noisy images is a crucial problem in multiple fields. In single-particle electron microscopy a huge number of two-dimensional (2D) image projections of biological molecules must be aligned multiple times to obtain a three-dimensional (3D) reconstruction. The 3D reconstruction quality is substantially determined by the accuracy and speed of the alignment method. Here, we implemented and validated a real-space correlation method based on Fourier-Bessel functions to perform the matching with a single 3D Fast Fourier transform. Our results show that we achieved faster and more accurate 2D alignments than state-of-the-art methods available in current image EM processing packages. The parallel-CUDA implementation achieves speedups ranging from 23 to 300 folds, depending on the angular sampling used. The achieved over-performance provides an excellent background for developing advanced single particle analysis algorithms towards higher resolutions. Besides, the method presented here is fully applicable to any 2D image rigid-body registration problem.

Keywords: image processing, 2D registration, electron microscopy, fast rotational matching, Bessel functions

1. INTRODUCTION

Single particle electron microscopy (EM) has become the most powerful technique in structural biology for studying large macromolecules and their assemblies. At near physiological conditions, the 3D structure of relevant biological complexes can be determined from sub-nanometer to even near atomic resolution (Frank, 2006). Using a small amount of purified sample, researchers process and average the collected 2D EM data into a 3D reconstruction. In the first step, numerous single molecule images are obtained from multiple EM measures. These images correspond to electronic density 2D projections of the molecule in different random orientations. Because such projections are extremely noisy images, they are aligned and classified by similarity to reduce the signal-to-noise ratio (SNR). In fact, single particle 3D reconstruction is an iterative alignment and classification procedure, where strong image averages produced by classification are used as reference images for the subsequent refinement steps. Therefore, the alignment, which ultimately determines the 3D reconstruction quality, is repeated multiple times. Such alignment typically maximizes a cross-correlation function (simple scalar product of the EM electron density values stored in the 2D images) between experimental noisy images and the reference images. Moreover, the computation time for 3D reconstruction increases with the number of images, becoming a bottleneck for high-resolution studies. Approximately 10^6 image projections are needed to target high-resolution. In this context, the complete process can take even days in a multiprocessor cluster. In summary, the alignment is a critical step that largely controls the efficiency and accuracy of the 3D EM reconstruction.

Current EM image processing packages use different alignment kernels to perform 2D registration efficiently, as described elsewhere (Joyeux and Penczek, 2002). The most popular approaches are the self-correlation method (SCF) and the resampling to polar coordinates (RPC) method. The latter resamples a fixed image into polar coordinate space with respect to several locations of the other image. By means of the Fourier convolution theorem, the rotational angle between projections is determined from a 1D fast Fourier Transform (FFT), whereas the two translational parameters (i.e., x and y shifts) are discretely scanned. In contrast, SCF is a Fourier method that decouples rotation and translation. The rotational angle is computed in the same way as in the RPC method, but using a mutual-correlation function instead the standard density cross-correlation. The translational parameters are obtained by a 2D FFT that greatly speeds up the alignment. Although SCF is much more efficient than RPC, it is less robust against noise (Cong et al., 2003; Joyeux and Penczek, 2002). To improve its accuracy, researchers have added a post local refinement step. This SCF plus refined protocol is the standard fastest alignment procedure. Alternatively, the fast rotational matching method (FRM2D) maintains the accuracy of RPC and is still competitive (2x slower) with SCF (Cong et al., 2003; Cong et al., 2005). In FRM2D, the alignment problem was recast into one translational and two rotational degrees of freedom. Instead of fixing one

image while rotating and translating the other, both

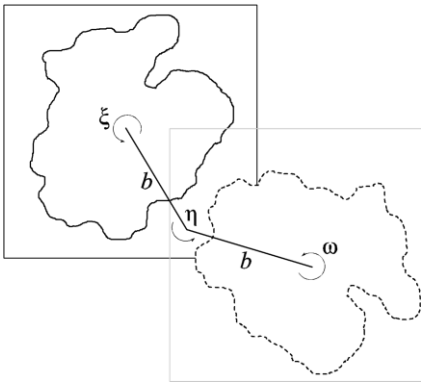


Figure 1. Matching setup of FBM method. The image on the left side is fixed, and the image on the right is moved by the rotations of ξ, η and ω to match the image on the left. *b* is a fixed value. In our implementation, *b* was fixed to 3.5 pixels.

images are rotated and one of them is translated along the axis formed between the images centers. This method accelerates the estimation of the two rotations by FFTs, while the remaining translational parameter is systematically explored (Cong et al., 2003; Cong et al., 2005). FRM2D has been implemented in the standard package EMAN and has been successfully used in challenging high-resolution 3D reconstructions (Cong et al., 2010). This method can also be considered as a 2D version of the 3D fast rotational matching procedure that we employed to predict protein interactions (Garzon et al., 2009) and to fit atomic structures into low-resolution EM density maps (Garzon et al., 2007).

More recently, Kovacs and collaborators (Kovacs et al., 2007) have outlined a new real-space correlation-based method to perform the 2D alignment step, known as the Fast Bessel Matching (FBM) method. In FBM, the matching problem is recast into three angular parameters that can be estimated by a single 3D Fourier transform. To speed up all the rotational degrees of freedom, the correlation function was expressed in terms of the Fourier-Bessel transform of the image projections. Theoretical estimates from FBM showed a much lower complexity than current alignment 2D algorithms (Kovacs et al., 2007). These results suggested FBM as the optimal real-space method for matching EM noisy images. However, this hypothesis has not yet been confirmed.

In this article, we present the first FBM implementation and its adaptation to High-Performance Computing (HPC) systems. The parallelization can be performed directly by farming the graphics processing unit (GPU) processors with different alignments. The main advantage is its relative low storage needs and its simplicity when implemented on the GPU. Next, we summarize the methodology employed. Then, we confirm its robustness and accuracy. Finally, we describe and test the GPU implementation.

2. METHODS

The 2D image alignment in single particle analysis should be considered as a template matching problem. The templates are either average images obtained using clustering or generated as 2D projections from 3D reference structure. The experimental EM projections are compared with the template images to find most similar alignment parameters. Although these registration parameters are typically defined by two translations and one rotation between a fixed and a moving image, here registration is defined by the three angles as depicted in Figure 1. With this set up, FBM allows the direct estimation of all cross-correlation values between two images as a function of the rotation angles ξ, ω and η. We provide a brief summary of FBM, which has been described in detail elsewhere (Kovacs et al., 2007). First, the template image Fourier-Bessel transform can be written as:

$$F_m(x) = \int_0^\infty \hat{f}_m(u)\, J_m(ux)\, u\, du \tag{1}$$

where $\hat{f}_m(u)$ corresponds to the Fourier transform of a given image *f* sampled in polar coordinates, with *u* being the fixed radius. The term $J_m(ux)$ corresponds to a Bessel function of the first kind with order *m*. The Fourier-Bessel transforms $G_m(x)$ of the other image to be matched is obtained as in Eq. (1). Note that both Fourier-Bessel transforms are computed using *2B* angular samples, and only for $|m| \leq B$, where *B* is the bandwidth. To compute the correlations with respect to the three angular variables, we can use:

$$C(\xi,\eta,\omega) = 2\pi \sum_{m,h,m'} e^{\mathrm{i}\left(m\xi + h(\eta+\epsilon) + m'\omega\right)} \times \int_0^\infty J_{m-h}(bx)\, J_{h-m'}(bx) \times F_m(x)\, \overline{G_{m'}(x)}\, x\, dx \tag{2}$$

where ϵ is a shift angle added to η to avoid duplicated values in the solutions. In our implementation ϵ was fixed to $\pi/(2k)$. Applying the following variable change in Eq. (2), we obtain:

$$h = h_1 + m' \quad m = m_1 + h_1 + m' \qquad \eta' = \xi + \eta \quad \omega' = \xi + \eta + \omega \tag{3}$$

This equation yields a correlation $C'(\xi,\eta',\omega')$, whose Fourier transform is:

$$\hat{T}(m_1,h_1,m') = 2\pi e^{\mathrm{i}(h_1+m')\epsilon} \times \int_0^\infty F_{m_1+h_1+m'}(x)\, \overline{G_{m'}(x)} \times J_{m_1}(bx)\, J_{h_1}(bx)\, x\, dx \tag{4}$$

By taking the inverse Fourier Transform of this equation, we can compute directly the matching correlation values for all angular triplets. In other words, we can recover the best matching angle solutions by simply find maximal values from the correlation 3D matrix. Efficiency has been achieved for two reasons: *i)* all of the registration procedures are reduced to a single inverse FFT and *ii)* the integral Eq. (4) involved simple operations between pre-calculated terms. In fact, all image Fourier-Bessel transforms are calculated at once, as are the Bessel functions. FBM was implemented in both CPU and GPU following this simple pseudo-code:

```
Precompute all image Fourier-Bessel transforms.
For each EXP image{
    Load EXP image
    For each REF image{
        Load REF image
        Calculate correlation REF / EXP
        Search best correlations
        Store the corresponding angles
    }
    Search best matching for this EXP
    Store best matching
}
```

2.1. CUDA Implementation

The graphical device FBM procedure was performed in three steps: *i*) compute the integral of Eq. (4) *ii*) compute the inverse Fourier Transform *iii*) search the highest correlation and store its alignment values. To compute the 3D FFT, we employed NVIDIA's cuFFT library. In particular, we obtained the best performance employing cufftPlanMany interface with the native compatibility configuration. However, this configuration needs an extra temporary correlation matrix, doubling the memory required. Thus, the FFTs number and hence the 2D alignments computed in batch are limited by the available GPU memory. For example, in a GTX 470 card we were only able to process packages of 9280, 1120, and 320 FFTs, for bandwidths 64, 128 and 192, respectively. Note that bandwidths are

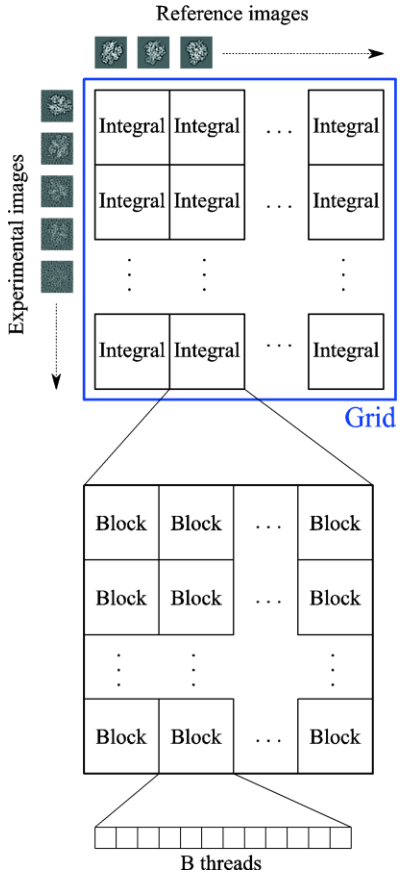


Figure 2. Grid layout of the main GPU kernel that computes the integral of Eq. (4). See details on the text.

multiple of 32 in agreement with NVIDIA's architecture guidelines. Constrained by FFT memory needs and by the bandwidth, we have designed two kernels: one to compute the integrals and other one to find the best matching results. The overall grid layout integral kernel is schematized in Figure 2. At the hierarchy top level, a grid is organized as a 2D array. The number of reference and experimental images computed in batch determined the dimensions of such 2D computational grid. At low level, the threads are organized in bandwidth size blocks to compute a single row of a given 3D matrix integral. Once the kernel has computed all the integrals, the inverse Fourier transforms are calculated by means of cuFFT. The resulting matrices containing all the correlation values are subsequently searched to find the highest match scores. The search is a quite serial process and a non-optimal task for the graphical device since correlation matrix transference to RAM is prohibitive. For the search kernel, each GPU block is devoted to process a single experimental image against all the references. Every thread in the block searches on its corresponding correlation matrix for the angular triplet (ξ, η and ω) with the highest value. This kernel ends transferring to the RAM the best reference registration and its parameters. After all the experimental images are processed, finally, on CPU the best matches are selected and stored. For this particular implementation, most of the computing time is spent the integral kernel (~75%) whereas FFT and correlation search consumed the remaining 10% and 15%, respectively.

2.2. Benchmark

To conduct the validation and comparison tests, we generated a matching benchmark from 2D projections of RNA polymerase II. This important macromolecule, which catalyzes the DNA transcription, has been characterized by EM in several conformational states (see, for example, Opalka et al., 2003). The simulated data have been obtained from the atomic structure (Protein Data Bank ID: 3M3Y) by using the single particle analysis software Xmipp (Marabini et al., 1996; Scheres et al., 2008). First, the electron density 3D atomic structure is projected into a 128 x 128 x 128 voxel density map with a 1.5 Å/voxel sampling rate by using the convert_pdb command. Then, a Gaussian low-pass filter was applied to simulate a 15 Å resolution

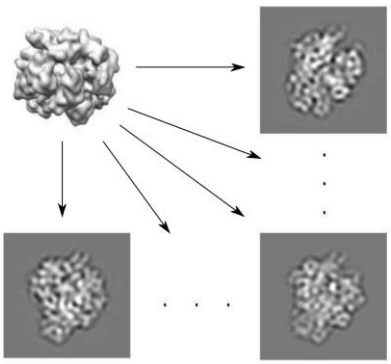


Figure 3. RNA polymerase EM map (grey surface) and some illustrative template projections (2D images).

map (fourier_filter command). From this simulated map, 80 random 2D projections of 128 x 128 pixels were created. Each projection corresponds to a given molecule orientation and, in principle, conforms to a different 2D shape (see Figure 3). To mimic real EM data, we also have simulated the microscope effect on these ideal projections using Xmipp phantom_simulate _microscope tool. This tool allowed us to add the contrast transfer function (CTF), such as defocus, astigmatism and lens aberration, and to simulate real noise with a given signal-to-noise ratio (SNR).

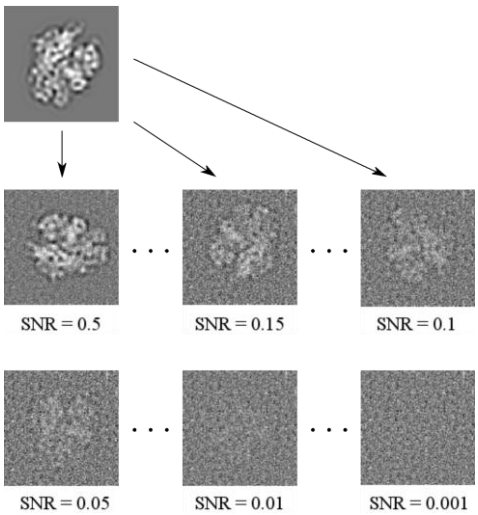


Figure 4. Illustrative examples of noise levels added to the random projections

Fifty different particle orientations were obtained by applying random rotations and translations to each reference image for generating a benchmark stack of 4000 image projections. The rotation shift was fixed to any angle between 0 and 360 degrees and, the translational shift was limited below 7 pixels. To simulate real data we also included noise to the images. We employed 19 different noise levels ranging from 0.5 to 0.001 of SNR, defined as $\sigma^2_{signal} / \sigma^2_{noise}$. Figure 4 illustrates the noise effect over the benchmark images. As can be observed, the macromolecule shape is only perceived at high SNR values. After applying all noise levels to every randomly oriented images, we had a total of 76000 experimental-like 2D projections.

The matching test will consist in recovering the original references of the whole experimental set by aligning them to the reference/template images. In principle, the highest cross-correlation values will correspond to the correct matches. The matching accuracy was measured by:

$$d\left|\sin(\Delta\phi/2)\right| + \sqrt{\Delta x^2 + \Delta y^2} \qquad (5)$$

where d is the particle diameter in pixels, $\Delta\phi$ is the relative angle misalignment, and x and y are the translational parameters (Joyeux and Penczek, 2002). The first term corresponds to the radial error, and the second corresponds to the translational error between reference and aligned images.

2.3. Technical details

These tests were conducted on a machine with an Intel i7 950 for the CPU and two NVIDIA GeForce GPUs, a GTX 470 and a GTX 680. The GTX 470 is a Fermi architecture card that has 448 cores along 14 multiprocessors, running at 1.22 GHz, and 1280 MB of RAM. The GTX 680 card has the latest Kepler architecture. It has 1536 cores along 8 multiprocessors running at a speed of 1.06 GHz and 2 GB of RAM available. To execute the FBM algorithm on the graphical device, we used CUDA 5.0 and NVIDIA driver version 302.59. The employ of the latest version of CUDA was mandatory to exploit the GTX 680 card capabilities.

3. RESULTS

3.1. CPU implementation

Because FBM has been never implemented, we first check its performance in a single CPU. To this end, we performed the matching test of the 4000 images against the 80 reference ones (see Methods) at three different bandwidths: 64, 128 and 192. These bandwidths correspond to an angular sampling of 2.8, 1.4 and 0.94 degrees, respectively. We also conducted the test at different noise levels to mimic the experimental conditions and test the method. The accuracy results of all 76000 matching experiments and their corresponding averages are shown in Figure 5. As it can be seen, the FBM method maintained subpixel accuracy (solid

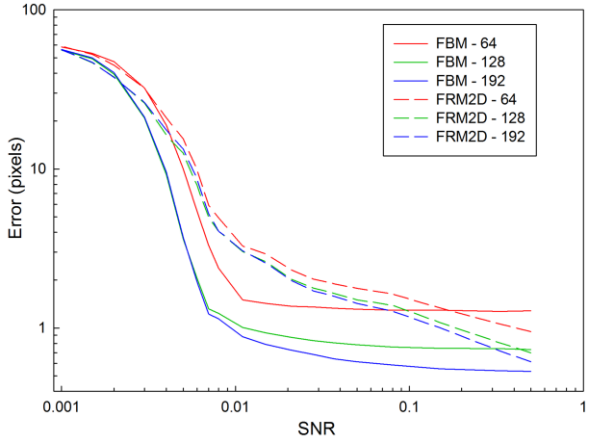


Figure 5. Average pixel error for FBM (solid lines) and FRM (dashed lines) for bandwidths 64 (red), 128 (green) and 192 (blue).

lines), even for SNR values close to 0.01. Around this region, alignments started to fail, and from this point, the error shown a fast accuracy loss. As expected, the accuracy improves as the angular sampling decreases (Figure 5). However, the differences between the 128 (green line) and 192 (blue line) bandwidths were not quite significant. The timing results are summarized in Table 1. FBM took less than 4 minutes with a bandwidth of 64 to match the 80 templates against 4000 experimental like images and less than an hour for 128. At the highest bandwidth FBM almost 5 hours were used to match the whole stack.

Table 1. Execution time in seconds for matching a stack of 4000 experimental test images against 80 reference images.

	CPU		FBM-GPU	
Bandwidth	FRM2D	FBM	GTX 470	GTX 680
64	2720	238	13.4	8.88
128	11293	3122	185	76
192	23939	16006	1014	443

For comparative purposes, we repeated the validation with FRM2D alignment method. This method exhibits better accuracy than the fastest SCF protocol at relatively low SNR ranges while keeps a good performance (Cong et al., 2005). We went a step further and, as suggested by Kovacs *et al.* (Kovacs et al., 2007), we optimized the FRM2D algorithm by reducing the bandwidth of an angular variable (Cong et al., 2003). This improvements result in a speed-up of one order of magnitude equaling the SCF performance. Therefore, in terms accuracy and efficiency, our FRM2D implementation is likely the best current reference method. In our tests, FBM clearly surpasses the accuracy of our optimized FRM2D version throughout the whole noise range (Figure 5). Even the FBM (solid lines) with the crudest sampling (64) is significantly more accurate than the FRM2D using thinner samplings (dashed lines) with SNR below 0.1. For example, at 128 the FBM error curve is always below that of FRM2D. The over-performance is more evident at the somewhat lower SNR levels. For example, at 0.01 of SNR, FBM matches with an error around 1 pixel, whereas FRM2D misses many alignments with an error larger than 3 pixels. From this SNR and under, the error recorded for both methods increases drastically, until both methods reach the maximum average error of 57 pixels at an SNR of 0.001. Fortunately, the typical EM experimental SNR ranges from 0.01 to 0.1.

The timing results are summarized in Table 1. For FRM2D, 45 minutes or up to 3 hours are needed for bandwidths of 64 and 128, respectively. On CPU, FBM provides speedups ranging from 1.5 to 11 times greater than FRM2D in the same conditions. These values are slightly better than the theoretical expectations, which estimated speedup gains around 2-5 fold (Kovacs et al., 2007). In summary, the obtained results confirm the superior accuracy and efficiency of FBM relative to the optimized FRM2D and by extension to current state of the art 2D alignment methods.

3.2. GPU implementation

Once the FBM was tested and its over-performance was demonstrated, we proceeded to implement a parallel version on the GPU by using CUDA. Parallelization was straightforward because the alignments of each reference image are independent tasks (see Method section for a detailed description).

To match each 4000 image stack, the CUDA version using GTX470 took 13.4 seconds, 3 minutes, or 17 minutes for bandwidths 64, 128 or 192, respectively. The executing times were between 66% and 40% smaller with newer and faster GTX 680 card. Thus, depending on the graphic device our CUDA implementation provides maximal speedups from 41x to 148x relative to the FBM-CPU version (Table 2). The speedup is not linear with the bandwidth. In the case of GTX 680 at 64, the card resources are not completely used and the speedup was limited to 27 fold. At bandwidth 192 the speedup decreases in both GPU cards. In this situation, the algorithm is likely to be saturating the device resources reducing the overall efficiency. However, in practical situations, angular samplings below 1º are not used because such accuracy is hardly achieved with noisy images. In summary, compared to the CPU, the GPU-FBM provides excellent performance. For example, it only took less than 10 minutes to match the whole 76.000 experimental image set, including all tested noise levels, with the fastest card at 128. More than 6.5 hours were needed for the CPU version in the same conditions.

If we now compare with currently used algorithms in CPU, such as FRM2D, using a GTX 470 card we found 203-, 61- and 23.6-fold speedups, for 64, 128 and 192 bandwidths, respectively. As expected, the speedups were substantially bigger for the GTX 680 ranging from 54 to 306. We did not implement the FRM2D GPU version because of its high memory requirements already pointed out in (Cong et al., 2003). Finally, it is important to mention that the CUDA version maintains the CPU accuracy in all cases tested.

Table 2. GPU speedups relative to CPU implementations

	GTX 470		GTX 680	
Bandwidth	FBM	FRM2D	FBM	FRM2D
64	17.7x	203.0x	26.8x	306.3x
128	17.8x	61x	41.1x	148.6x
192	15.8x	23.6x	36.1x	54.0x

4. Conclusions

In this paper, we implemented and validated a novel real-space and correlation-based 2D image registration algorithm. By means of Fourier-Bessel functions and a suitable recasting of the matching problem, we reduced the alignment process to calculate a single 3D FFT.

To verify the FBM robustness, we performed efficiency and accuracy tests with simulated RNA polymerase II images over a wide range of noise levels. The method maintains subpixel resolution at experimental-like noise levels. The GPU implementation boosts the efficiency between 16 and 41-fold with respect to the single CPU version. Moreover, compared with FRM2D, which is currently available on the *de facto* standard EM data processing package EMAN, FBM stands a significant improvement. In fact, for two different graphic devices we obtained speedups ranging from 23 to 300 folds relatively to our optimized FRM2D version. More importantly, FBM is considerably more accurate at experimental like noise conditions.

Based on the obtained results, our CUDA-FBM should be a sensible choice for image 2D alignment. It

will be particularly useful in the upcoming EM high-throughput scenario, where high-resolution structures and their huge number of projections have to be processed. In fact, we already had successful results with real experimental data using this novel approach. Our approach could complement recent CUDA developments in the EM image processing field (Castaño-Díez et al., 2010; Li et al., 2010; Schmeisser et al., 2009; Tagare et al., 2010) and other related Bessel based approximations for 3D reconstruction (Estrozi and Navaza, 2010).

Finally, our CUDA-FBM alignment kernel is a general method that can be useful in any application where the registration of multiple noisy images is required.

ACKNOWLEDGMENTS

This study was supported by BFU2009-09552, CAM-S2010/BMD-2353, CTQ2012-35873 and by the Human Frontier Science Program - RGP0039/2008. We thank P. Chys and JR Lopez-Blanco for their suggestions.

REFERENCES

Castaño-Díez, D., Scheffer, M., Al-Amoudi, A., Frangakis, A.S., 2010. *Alignator: A GPU powered software package for robust fiducial-less alignment of cryo tilt-series.* Journal of Structural Biology 170, 117-126.

Cong, Y., Kovacs, J.A., Wriggers, W., 2003. *2D fast rotational matching for image processing of biophysical data.* Journal of Structural Biology 144, 51-60.

Cong, Y., Jiang, W., Birmanns, S., Zhou, Z.H., Chiu, W., Wriggers, W., 2005. *Fast rotational matching of single-particle images.* Journal of Structural Biology 152, 104-112.

Cong, Y., Baker, M.L., Jakana, J., Woolford, D., Miller, E.J., Reissmann, S., Kumar, R.N., Redding-Johanson, A.M., Batth, T.S., Mukhopadhyay, A., Ludtke, S.J., Frydman, J., Chiu, W., 2010. *4.0-Å resolution cryo-EM structure of the mammalian chaperonin TRiC/CCT reveals its unique subunit arrangement.* Proceedings of the National Academy of Sciences of the United States of America 107, 4967-4972.

Estrozi, L.F., Navaza, J., 2010. *Ab initio high-resolution single-particle 3D reconstructions: The symmetry adapted functions way.* Journal of Structural Biology 172, 253-260.

Frank, J., 2006. *Three-Dimensional Electron Microscopy of Macromolecular Assemblies.* Oxford University Press, New York.

Garzon, J.I., Kovacs, J., Abagyan, R., Chacon, P., 2007. *ADP_EM: fast exhaustive multi-resolution docking for high-throughput coverage.* Bioinformatics 23, 427-433.

Garzon, J.I., Lopez-Blanco, J.R., Pons, C., Kovacs, J., Abagyan, R., Fernandez-Recio, J., Chacon, P., 2009. *FRODOCK: a new approach for fast rotational protein-protein docking.* Bioinformatics 25, 2544-2551.

Joyeux, L., Penczek, P.A., 2002. *Efficiency of 2D alignment methods.* Ultramicroscopy 92, 33-46.

Kovacs, J.A., Abagyan, R., Yeager, M., 2007. *Fast bessel matching.* Journal of Computational and Theoretical Nanoscience 4, 84-95.

Li, X., Grigorieff, N., Cheng, Y., 2010. GPU-enabled *FREALIGN: Accelerating single particle 3D reconstruction and refinement in Fourier space on graphics processors.* Journal of Structural Biology 172, 407-412.

Marabini, R., Masegosa, I.M., San Martín, M.C., Marco, S., Fernández, J.J., De La Fraga, L.G., Vaquerizo, C., Carazo, J.M., 1996. *Xmipp: An image processing package for electron microscopy.* Journal of Structural Biology 116, 237-240.

Opalka, N., Chlenov, M., Chacon, P., Rice, W.J., Wriggers, W., Darst, S.A., 2003. *Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase.* Cell 114, 335-345.

Scheres, S.H.W., Núñez-Ramírez, R., Sorzano, C.O.S., Carazo, J.M., Marabini, R., 2008. *Image processing for electron microscopy single-particle analysis using XMIPP.* Nature Protocols 3, 977-990.

Schmeisser, M., Heisen, B.C., Luettich, M., Busche, B., Hauer, F., Koske, T., Knauber, K.H., Stark, H., 2009. *Parallel, distributed and GPU computing technologies in single-particle electron microscopy.* Acta Crystallographica Section D: Biological Crystallography 65, 659-671.

Tagare, H.D., Barthel, A., Sigworth, F.J., 2010. *An adaptive Expectation-Maximization algorithm with GPU implementation for electron cryo-microscopy.* Journal of Structural Biology 171, 256-265.

GPU-ACCELERATED MODELLING OF BIOLOGICAL MEMBRANES ION-TRANSPORT

Adam Gorecki[(a,*)] , Krzysztof Dołowy[(b)]

[(a)]Warsaw University of Life Sciences WULS-SGGW, Nowoursynowska St. 159, 02-787 Warsaw, Poland
[(b)]Warsaw University of Life Sciences WULS-SGGW, Nowoursynowska St. 159, 02-787 Warsaw, Poland
[(a)]adam_gorecki@sggw.pl ,
[(b)]krzysztof_dolowy@sggw.pl
* the corresponding author

ABSTRACT

The ion-transport modeling through biological membranes is important for understanding of many life processes. The transmembrane potential and ion concentrations in the stationary state can be measured in *in-vivo* experiments. They can be also simulated within membrane models. Here we consider a basic model of ion transport that describes the time evolution of ion concentrations and potentials through a set of nonlinear ordinary differential equations.

To reduce the computation time we have developed a GPU-optimized application for simulation of the ion-flows through a membrane starting from an ensemble of initial conditions. The application is written in CUDA programming language and runs on NVIDIA TESLA family of numerical accelerators. The calculation speed can be increased over 10000 times compared with a sequential program running on a PC.

Keywords: biological membranes, electrochemistry, differential equations integration, CUDA, TESLA

1. INTRODUCTION

Every living cell has to exchange energy and mass with its environment in a selective way. An animal cell is surrounded by a lipid bilayer, called as a biological membrane (Stryer 1981). The membrane includes special proteins (channels, transporters and pumps) which enable selective transmission of ions. The activity of these transmission devices is controlled by different factors such as ion concentration, electric fields or presence of specific molecules.

Studies on ion-transport through biological membranes are crucial for understanding the etiology of many diseases. Abnormal ion transport is the cause of many serious health problems and is responsible for toxicity of many chemicals. For example, the malfunctioning of CTFR channel that controls transport of chlorine and bicarbonate ions in bronchial (lungs) epithelial tissue causes *cystic fibrosis* (Gadsby, Vergani and Csanády 2006). Problems in ion transfer through nervous cells are observed in some types of epilepsy (Scheffer and Berkovic 1997). The mechanisms of many poisons or venoms such charybdotoxin (Goldstein and Miller 1993) or iberiotoxin (Candia, Garcia and Latorre 1992**)** is based on blocking the activity of important membrane channels. A recent review of important protein channels has been presented in (Toczylowska-Maminska and Dolowy 2012).

The time evolution of membrane potentials and ion concentrations can be modeled using differential equations (Falkenberg and Jakobsson 2010, Sohma et all. 2000). The variables of the models (potentials and concentrations) can be also observed in *in-vivo* experiments. The comparison between simulations and experiments allows for optimization of model parameters to make them more realistic. In this paper we present results of simulations of transmembrane potential generated by epithelial cell monolayer. The calculations are based on phenomenological model with a reduced number of parameters.

Epithelial cells are external cells of organs contacting with external environment. An epithelial cell membrane consists of two parts:

- *basolateral* – contacting with the internal cells of the organ,
- *apical* – contacting with an external environment .

Both parts of the cell membrane are covered by a solvent which in biological system is a water solution of ions at physiological concentrations. The activity of membrane proteins and their selectivity for specific ions defines the stationary state of the system. The bilayer is polarized with membrane electric potential (so called resting potentials), which stops effective ion currents. This potential can be measured using special electrodes.

The most important ions contributing to the membrane transport are:

- potassium K+
- sodium Na+
- chlorine Cl-
- bicarbonate HCO_3-

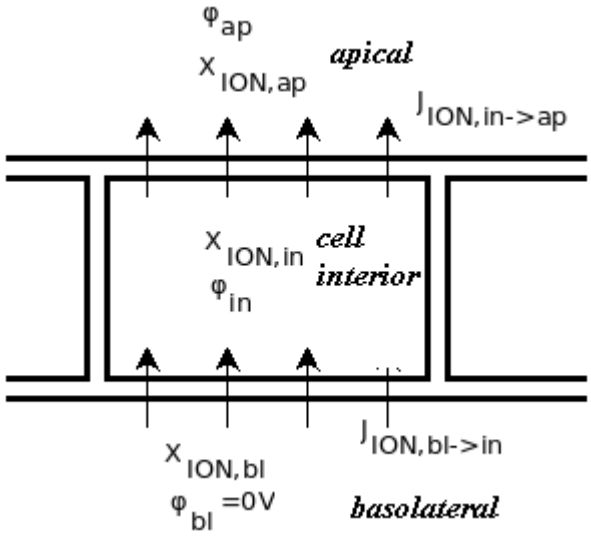


Figure 1. The considered model of ion transport in epithelial cell monolayer. The space around a cell is divided into apical(*ap*), basolateral (*bl*) and cell interior (*in*) domains

2. THE SIMULATED SYSTEM

The model of epithelial tissue considered below is illustrated on Figure 1. The cell and its neighborhood are divided into three geometrical regions: apical (*ap*), the cell interior (*in*), and basolateral (*bl*).
We assume that:

- the concentrations of ions in basolateral and apical regions are constant in time,
- the basolateral area has reference electric potential φ_{bl}=0 V.

Within our model the state of the system is fully represented by the following variables:

- the apical potential φ_{ap},
- the cell interior potential φ_{in},
- the ion concentrations in the cell interior $X_{ION,in}$, where *ION* = K+, Na+, Cl-, HCO_3-

The apical potential is equivalent to transmembrane potential because we have assumed φ_{bl}=0 V.
Potentials φ_{ap}, φ_{in} characterize the state of two capacitors created by apical and basolateral sides of the membrane. The currents charging these 'capacitors' are related to the total flows of ions of different types.

There are many models of ion flows described in literature, for example (Falkenberg and Jakobsson 2010, Sohma, Gray, Imai and Argent, 2000). These models describe time-evolution with ordinary differential equations. The equations include many parameters characterizing specific ion-channel activity. If a model contains too many parameters than usually its difficult to find their realistic values.
We proposed our model with a reduced number of parameters:

$$\partial_t \varphi_{ap} = \frac{\sum_{IONS} J_{ION}(in \to ap)}{C_{ap}} + \frac{\sum_{IONS} J_{ION}(bl \to in)}{C_{bl}}$$

$$\partial_t \varphi_{bl} = \frac{\sum_{IONS} J_{ION}(bl \to in)}{C_{bl}}$$

$$\partial_t X_{ION,in} = \frac{J_{ION}(in \to ap) - J_{ION}(bl \to in)}{F \cdot z_{ION} \cdot Vol}$$

where:

$J_{ION}(bl \to in)$ - the electric current (positive charges flow) corresponding to flow of ions type *ION* from basolateral area to the cell interior,

$J_{ION}(in \to ap)$ - the electric current (positive charges flow) corresponding to flow of ions type *ION* from basolateral area to the cell interior,

C_{bl} - the basolateral membrane capacity,

C_{ap} - the apical membrane capacity,

z_{ION} - the sign of ion type ION,

Vol - the volume of the cell,

F - the Faraday constant.

The currents of ions are given by equations:

$$J_{ION}(bl \to in) = G_{ION,bl} \cdot (-\varphi_{in} - \varphi_{Nernst}(T, z_{ION}, X_{ION,bl}, X_{ION,in})),$$

$$J_{ION}(in \to ap) = G_{ION,ap} \cdot (\varphi_{ap} - \varphi_{in} - \varphi_{Nernst}(T, z_{ION}, X_{ION,in}, X_{ION,ap})),$$

where:

$\varphi_{Nernst}(T, z_{ION}, X_{ION,src}, X_{ION,trg})$ is the Nernst equilibrium potential for *ION* and flow direction of *src* → *trg*,

$G_{ION,ap}$, $G_{ION,bl}$ are effective electric permeabilities (reciprocal resistance) of apical and basolaterar side of membrane, respectively.

The Nernst resting potential (Wright 2004) is defined as

$$\varphi_{Nernst}(T, z_{ION}, X_{ION,src}, X_{ION,trg}) =$$

$$= -\frac{RT}{Fz_{ION}} \cdot \ln\left(\frac{X_{ION,src}}{X_{ION,trg}}\right),$$

where T is absolute temperature of solvent.
When the potential difference between the sides of membrane *src*, *trg* is equal to Nernst potential, there is no effective current of ion of type *ION*.

3. IMPLEMENTATION

The simulation program is written in CUDA C language and designed to work on NVIDIA TESLA family graphical accelerators (NVIDIA company 2012).

The general idea of our approach is to perform the same operations on different data in parallel – it is so called Single Instruction Multiple Data approach. For our applications (the same algorithm, many data sets, small amount of required operational memory) we need many instances of simple scalar calculations. NVIDIA GPU accelerator is perfectly suited for such task, because we can perform separate simulations on different cores. We expect the maximum speedup if the number of separate tasks does not exceed the maximum number of threads allowed to run in parallel. In our application the GPU accelerator works as a computer farm executing separate instances of the same program, so we have not used advanced CUDA environment features, such as dedicated numerical libraries or texture processing.

To integrate the differential equations we use the forth-order Runge-Kutta integration scheme.
The host process creates GPU device processes performing simulations starting from different initial conditions (for example ion concentrations) or different membrane parameters (see Figure 2). This approach can give linear speedup as a result of parallel calculations on separate threads with no shared memory.

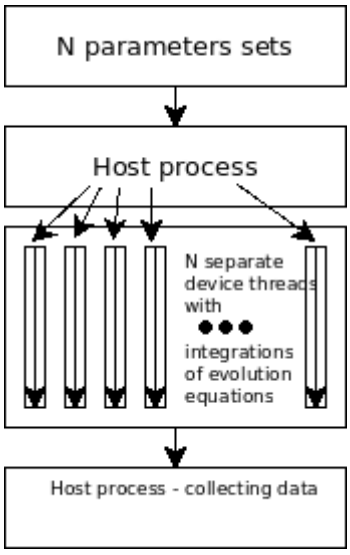


Figure 2. GPU optimization of N instances of simulations of the system. The host process reads initial data and parameters, distributes these sets in GPU device memory, and launches apropriate number of GPU device threads. Each GPU thread performs simulation for one parameters set.

The host process collects the data. Usually we do not need the whole evolution of the system, but only the values of variables when the system has converged to the stationary state.

4. RESULTS

The example results illustrating transmembrane potential φ_{ap} are shown on Figure 3-6.

The contour plots show transmembrane potential value in volts[V] as a function of two variables: concentrations of potassium K+ and sodium Na+ ions at apical side of the cell. The ion concentrations is given in milimoles per litre [mM].

The concentrations of other ions in other areas, the membrane permeabilities $G_{ION,ap}$, $G_{ION,bl}$, membrane capacities C_{bl}, C_{ap}, and the temperature T are the same for all calculations.

The parameters used in simulations are given in Table 1.

The calculations of transmembrane potential were performed for a grid of 9 x 15 = 135 parameter sets with:

- K+ ion concentation $X_{K+,ap}$changing from 1 mM to 9 mM with step 1 mM,
- Na+ ion concentration $X_{Na+,ap}$ changing from 40 mM to 180 mM with step 10 mM.

The average time of calculations on NVIDIA TESLA C870 graphics accelerator was about **0.175 seconds** of total time, comparing to about **40 minutes** in for a scalar calculations using program compiled by Free Pascal Compiler on AMD Athlon 2 GHz.

So we have a speedup **greater than 10000 !**
Such speedup it is probably related to very good computational performance and communication with local memory of NVIDIA GPU cores compared to conventional CPU.

Figure 3 shows the stationary value of the transmembrane potential for the normal membrane conductivity for all ions. As we can see, the relationship between the potential and K+, Na+ concentrations is strongly nonlinear. Small differences in K+ ion concentrations have an important influence on the potential. The transmembrane potential function $\varphi_{ap}(X_{K+,ap}, X_{Na+,ap})$ does not factorize as a product of two functions depending on single ion concentrations.

Figure 4 shows the stationary value of the transmembrane potential when K+ ion transport is block at the apical side of a membrane. As we can expect, the membrane potential depends of Na+ concentrations only. The values of transmembrane potential are different than in the normal case described previously because the calculation starts from different initial membrane potential φ_{ap} . We assumed that membrane capacities were not charged before evolution started.

Figure 5 shows the stationary value of the transmembrane potential for the normal membrane

conductivity for all ions as a function of Na+ and Cl- concentrations. The same potential as a function of concentration of HCO_3- and Cl- ions is illustrated in Figure 6. As we can see, the contribution of those ions for transmembrane potential is less significant than the influence of potassium ion.

CONCLUSIONS

We have developed a tool for modeling membrane ion flows working on NVIDIA graphics accelerators. The program increases the speed of calculations over 10000 times if compared with a our previous approach running on a scalar CPU.

Our GPU program may be very useful for membrane model parameterization and its experimental verification. Thanks to its speed we can optimize model parameters like membrane permabilities and capacities using a large number of experimental data. Moreover the code can be easily modified to other forms of differentials equations.

The model reported above threats the membrane as a single entity without focusing on particular ion channels. It can be easily generalized for a specific type membrane by modification of $J_{ION}(src \to trg)$ terms. We can consider different current-voltage characteristics of specific protein channels by selecting an appropriate $J_{ION}(src \to trg)$ term form.

The speed achieved on GPU accelerators seems to be sufficient for non-local model of membrane transport. In such models channels of different types are spatially distributed on the membranes and the equilibration process involves local currents flowing inside the cell.

In future we also plan to develop a user-friendly graphical interface, which will make the program simpler for scientist and students who are not computational experts.

ACKNOWLEDGMENTS

The work was supported by Ministry of Science and Higher Education Grant No 1828/B/PO1/2010/39.

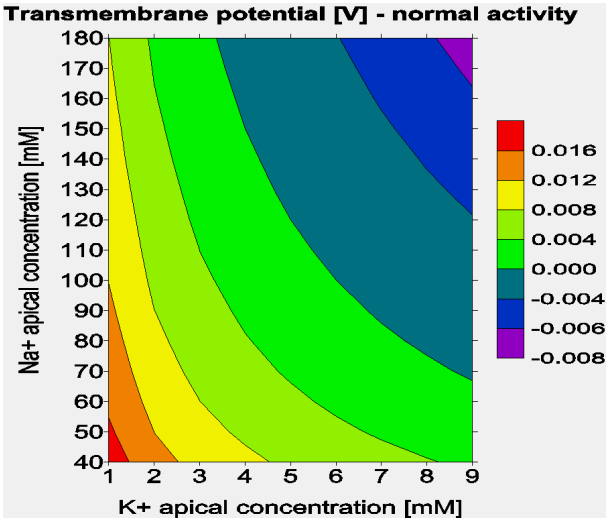


Figure 3. Example results from membrane flows simulation . The transmembrane potential (potential between basolateral side and apical side) as a function of variables - concentrations of K+ and Na+ ions at apical side. The normal transport of all ions is assumed.

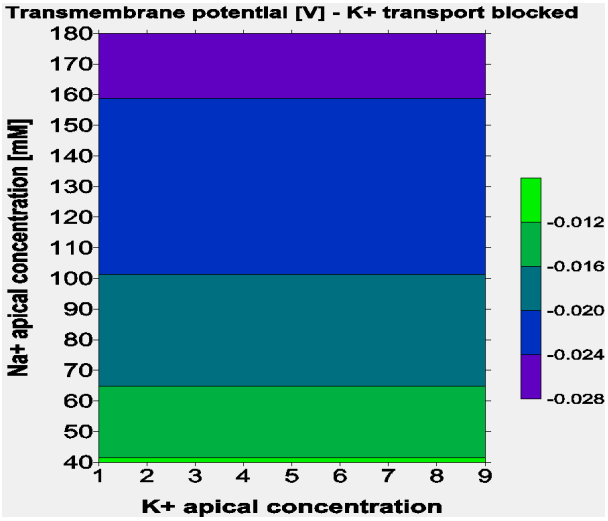


Figure 4. As Figure 3, but for the blocked transport of K+ ions.

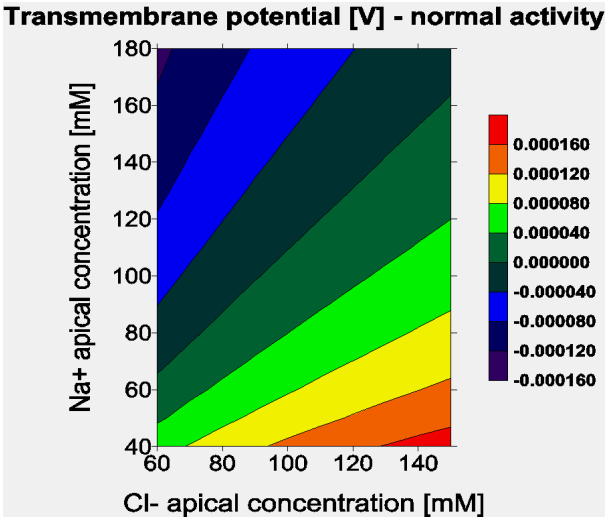


Figure 5. The transmembrane potential as a function of Na+ and Cl- apical concentrations. K+ and HCO_3- concentrations remains constant and equal to $X_{K+,ap}$ =5mM, $X_{HCO3-,ap}$ =24 mM respectively.

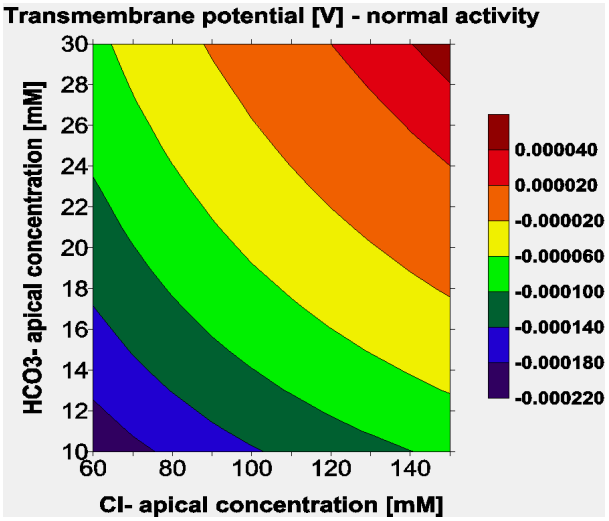


Figure 6. The transmembrane potential as a function of HCO3- and Cl- apical concentrations. K+ and Na+ concentrations remains constant and equal to $X_{K+,ap}$ =5mM, $X_{Na-,ap}$ =140 mM respectively.

Table 1: Membrane flow model parameters used for example results

Parameter	Value	Unit
Ion concentrations at apical side		
$X_{K+,ap}$	1-9, step 1	mM
$X_{Na+,ap}$	40-180, step 10	mM
$X_{Cl-,ap}$	110	mM
$X_{HCO3-,ap}$	24	mM
Ion concentrations at basolateral side		
$X_{K+,bl}$	5	mM
$X_{Na+,bl}$	120	mM
$X_{Cl-,bl}$	110	mM
$X_{HCO3-,bl}$	24	mM
Temperature		
T	300	K
Apical side permeability (reciprocal resistance)		
$G_{K+,ap}$ - Figure 3	$5 \cdot 10^{-4}$	Ω^{-1}
$G_{K+,ap}$ - Figure 4	0	Ω^{-1}
$G_{Na+,ap}$	$5 \cdot 10^{-4}$	Ω^{-1}
$G_{Cl-,ap}$	$5 \cdot 10^{-4}$	Ω^{-1}
$G_{HCO3-,ap}$	$5 \cdot 10^{-4}$	Ω^{-1}
Basolateral side permeability (reciprocal resistance)		
$G_{K+,bl}$	$5 \cdot 10^{-3}$	Ω^{-1}
$G_{Na+,bl}$	$5 \cdot 10^{-3}$	Ω^{-1}
$G_{Cl-,bl}$	$5 \cdot 10^{-3}$	Ω^{-1}
$G_{HCO3-,bl}$	$5 \cdot 10^{-3}$	Ω^{-1}
Apical side capacity		
C_{ap}	10^{-6}	F
Basolateral side capacity		
C_{bl}	10^{-5}	F
Cell volume		
Vol	10^{-9}	m^3

REFERENCES

Toczylowska-Maminska, R., Dolowy, K., 2012.
Ion transporting proteins of human bronchial epithelium.
Journal of Cellular Biochemistry 113:426-432.

Stryer, L., 1981.
Biochemistry, New York: W. H. Freeman,

NVIDIA corporation, 2012.
CUDA C Programming Guide
Available from:
http://developer.NVIDIA.com/NVIDIA-gpu-computing-documentation
Accesed: Jul 11, 2012

Falkenberg, C. V. , Jakobsson, E., 2010.
A Biophysical Model for Integration of Electrical, Osmotic, and pH Regulation in the Human Bronchial Epithelium.
Biophysical Journal 98:1476–1485.

Sohma, Y., Gray, M.A., Imai, Y. , Argent, B.E., 2000.
HCO_3^- Transport in a Mathematical Model of the Pancreatic Ductal Epithelium
Journal of Membrane Biology 176:77–100.

Gadsby, D.C., Vergani, P., Csanády, L., 2006.
The ABC protein turned chloride channel whose failure causes cystic fibrosis.
Nature 7083: 477–83.

Scheffer, I., Berkovic, S., 1997.
Generalized epilepsy with febrile seizures plus. A genetic disorder with heterogeneous clinical phenotypes.
Brain 120: 479–90.

Goldstein, S.A., Miller, C., 1993.
Mechanism of charybdotoxin block of a voltage gated K+ channel.
Biophysical Journal 65:1613–1619.

Candia, S., Garcia, M.L., Latorre, R., 1992.
Mode of action of iberiotoxin, a potent blocker of the large conductance Ca(2+)-activated K+ channel.
Biophysical Journal 63:583–590.

Wright, S.H., 2004.
Generation of resting membrane potential.
Advances in Physiology Education 28:139-142.

DESCRIPTION AND OPTIMIZATION OF THE STRUCTURE OF HORIZIONTALLY HOMOGENEOUS PARALLEL AND DISTRIBUTED PROCESSING SYSTEMS

Tiit Riismaa

Institute of Cybernetics at Tallinn University of Technology

Akadeemia tee 21, Tallinn, 12618

tiitr@ioc.ee

ABSTRACT

A method of description and optimization of the structure of horizontally homogeneous multi-level parallel and distributed processing systems is presented. The set of feasible structures for such class of systems is defined. The description of this set is constructed in terms of the graph theory. For representation, the feasible set of structures, a condition for adjacency matrixes of adjacent levels is derived. For the reduced statement, two types of variable parameters are defined: for the level size and for the relations of adjacent levels. The formalism considered here, enables to state the structure optimization problem as a two-phase mutually dependent discrete optimization problem and to construct some classes of effective solution methods. Modelling and optimization of the structure of multi-level processing system illustrates the considered approach.

Keywords: Parallel and distributed simulation; mathematical programming; multi-level processing; multi-level selection procedure; ordering of non-ordered sets.

1. INTRODUCTION

Large-scale problems can be decomposed in many different ways (Mesarovic, Macko and Takahara 1970). The current approach for describing and optimizing the structure of hierarchical systems is based on a multi-level partitioning of given finite set in which the qualities of the system may depend on the partitioning (Riismaa, Randvee and Vain 2003). Examples of problems of this class are aggregation problems, structuring of decision-making systems, database structuring, multiple distribution or centralization problems, multi-level tournament systems, multi-level distribution systems and optimal clustering problems.

In a multi-level distribution system each element is a supplier for some lower level elements and a customer for one higher-level element. The zero-level elements are only customers and the unique top-level element is only a supplier. The choice of optimal number of suppliers-customers on each level is a mathematically complicated problem.

The multi-level tournament system (Laslier 1997) is a relatively simple special case of a multi-level processing system. To consider a tournament system, the number of games (pair-wise comparisons) is a quadratic function of the number of participants. This is a very quickly increasing function. If the number of participants was large, the number of games is very large. This is a reason why the multi-level approach is useful for the selection of the winner. From the tournaments of the first level, the winners are distributed between the tournaments of the next level. The second level tournaments´ winners are going to the third level, until the winner is selected. Suppose the goal is to minimize the number of all games. If the price for all games is the same, the solution of the problem is well known. Each tournament has two participants and only one game is played. If the prices of games for different levels are different or constraints to the number of levels are active, a relatively complicated nonlinear integer programming problem arises.

The assembling problem as well as a broad class of design and implementation problems, such as component selection in production systems, reconfiguration of manufacturing structures, optimization of the hierarchy of decision making systems, multi-level aggregation, creation and cancellation of levels, etc. can be mathematically stated as a multi-level selection problem (Riismaa, Randvee and Vain 2003).

In this paper a method of description and optimization of the structure of multi-level parallel and distributed processing systems is presented. The set of feasible structures for such class of systems is defined. The description of this set is constructed in terms of the graph theory. For representation the feasible set of structures a condition for adjacency matrixes of adjacent levels is derived.

The general problem of optimal multi-level paralleling procedure is presented. This problem is

stated as a problem of selecting the feasible structure which corresponds to the minimum of total loss.

An important special case is considered, where the connection cost between the adjacent levels is the property of the supreme level: each row of the connection cost matrices between the adjacent levels consists of equal elements. It means that for each item on the next level the connection with all items on the previous level have the same costs. For this reduced statement two types of variable parameters are defined. Free variables of the inner minimization are used to describe the connections between the adjacent levels. Free variables of the outer minimization are used for the representation of the number of elements at each level.

For horizontally homogeneous hierarchies the inner minimization problem (to find optimal connection between adjacent levels) is solved analytically.

2. THE PROBLEM OF OPTIMAL MULTI-LEVEL PARALLEL AND DISTRIBUTED PROCESSING SYSTEM

Consider all *s*-levels hierarchies, where nodes on level *i* are selected from the given nonempty and disjoint sets and all selected nodes are connected with the selected nodes on adjacent levels. All oriented trees of this kind form the feasible set of hierarchies (Riismaa 1993, 2003; Riismaa, Randvee and Vain 2003).

The illustration of this formalism is given in Fig.1 (Riismaa 2003).

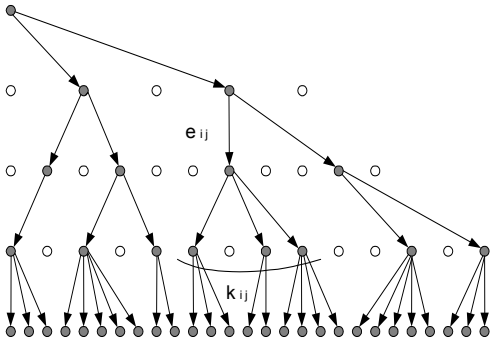


Figure 1 - Feasible set of structures

Suppose $m_i \times m_{i-1}$ matrix $Y_i = \left(y^i_{jr}\right)$ is an adjacent matrix of levels *i* and $i-1$ $(i = 1,...,s)$ where

$$y^i_{jr} = \begin{cases} 1, \, j-th\ element\ on\ level\ i\ connected \\ \quad with\ r-th\ element\ on\ level\ i-1 \\ 0, otherwise \end{cases}.$$

Suppose m_0 is the number of 0-level elements (level of object).

Theorem 1. All hierarchies with adjacency matrixes of adjacent levels $\{Y_1,...,Y_s\}$ from the described set of hierarchies satisfy the condition

$$Y_s \cdot ... \cdot Y_1 = (\underbrace{1,...,1}_{m_0}) \tag{1}$$

The assertion of this theorem is determined directly.

The general optimization problem is stated as a problem of selecting the feasible structure which corresponds to the minimum of total loss given in the separable-additive form:

$$\min \left\{ \sum_{i=1}^{s} \sum_{j=1}^{m_i} h_{ij} \left(\sum_{r=1}^{m_{i-1}} d^i_{jr} y^i_{jr} \right) \middle| \, Y_s \cdot ... \cdot Y_1 = (\underbrace{1,...,1}_{m_0}) \right\}$$

over $Y_1,..., Y_s$. (2)

Here $h_{ij}(\cdot)$ is an increasing loss function of *j*-th element on *i*-th level and d^i_{jr} is the element of $m_i \times m_{i-1}$ matrix D_i for the cost of connection between the *i*-th and (*i*-1)-th level.

The meaning of functions $h_{ij}(k)$ depends on the type of the particular system.

In this paper we suppose additionally that

$h_{ij}(0) = 0 (i = 1,..,s; j = 1,...,m_i)$ and $h_{ij}(k) (i = 1,...,s; j = 1,...,m_i)$ are increasing functions.

By the optimization of the structure of multi-level tournament system, the loss inside the *j*-th tournament on *i*-th level is

$$h_{ij}(k_{ij}) = d^i_j k_{ij} (k_{ij} - 1),$$

where k_{ij} is the number of participants of *j*-th tournament of *i*-th league.

By complexity optimization of hierarchically connected subsystems, the loss inside the *j*-th set of partitioning on *i*-th level may be defined as follows:

$$h_{ij}(k) = \sum_{q=1}^{k} a_{ijq} \frac{k!}{q!(k-q)!}.$$

In this case, the value of the function $h_{ij}(k)$ describes the number of all nonempty subsystems inside the *j*-th set of partitioning on *i*-th level.

Mathematically, this problem is an integer programming problem with a non-continuous objective function and with a finite feasible set.

3. REDUCED PROBLEM OF OPTIMAL PARALLEL AND DISTRIBUTED PROCESSING SYSTEM

Here an important special case is considered, where the connection cost between the adjacent levels is the property of the supreme level: each row of the connection cost matrices between the adjacent levels consists of equal elements. It means that for each item on the next level the connection with all items on the previous level have the same costs:

$$d^i_{jr} = d^i_j \ (i = 1,...,s; j = 1,...,m_i; r = 1,...,m_{i-1}).$$

Now there is a possibility to change the variables and to represent the problem so that

$$d^i_{jr} = 1; \quad i = 1,...s; \ j = 1,...,m_i; \ \ r = 1,...,m_{i-1}.$$

Now the total loss depends only on sums $\sum_{r=1}^{m_{i-1}} y^i_{jr} = k_{ij.}$, where k_{ij} is the number of edges beginning in the *j*-th node on *i*-th level.

In terms of tournament theory, the goal function doesn't depend how to distribute the winners on previous level between tournaments on the next level. But the goal function depends only how large are the tournaments. In terms of graph theory, the goal function doesn't depend what nodes connect but depends how many nodes to connect. Shortly, if additionally to change the variable, each connection between adjacent levels has the same cost.

Recognize also that $\sum_{j=1}^{m_i} k_{ij} = p_{i-1}, \ \ i = 1,...,s$, where p_i is the number of nodes on *i*-th level. If to suppose additionally that $h_{i1}(k) \le \cdots \le h_{im_i}(k)$ for each integer k, the general problem (2) transforms into the two mutually dependent phases:

$$\min\left\{\sum_{i=1}^{s} g_i(p_{i-1}, p_i) \,\middle|\, (p_1,..., p_{s-1},1) \in W^s\right\} \quad (3)$$

$$\text{over} \quad p_1,..., p_{s-1}$$

where

$$g_i(p_{i-1}, p_i) = \min\left\{\sum_{j=1}^{p_i} h_{ij}(k_{ij}) \,\middle|\, \sum_{j=1}^{p_i} k_{ij} = p_{i-1}\right\}$$

$$\text{over} \quad k_{i1},..., k_{ip_i}$$

$$W^s = \{(p_1,..., p_{s-1},1) \,|\, 1 \le p_i \le p_{i-1}\}. \quad (4)$$

Free variables of the inner minimization (4) are used to describe the connections between the adjacent levels. Free variables of the outer minimization (3) are used for the representation of the number of elements at each level.

For solving problem (3), (4) double-cycle recursive optimization algorithms are constructed (Riismaa 2011). The inner cycle increases the number of elements inside of the current level by one unit, and outer cycle on each step increases the number of levels by one unit. On the each iteration, the one-parameter integer-programming problem must be solved.

4. ANALYTICAL METHOD OF SOLVING REDUCED PROBLEM FOR HORIZONTAL HOMOGENEOUS HIERARHIES

The hierarchy is called horizontal homogeneous if

$$h_{ij}(k) = h_i(k) \ \ (j = 1,...,m_i; i = 1,...,s). \quad (5)$$

Now (3), (4) transforms to

$$\min\left\{\sum_{i=1}^{s} g_i(p_{i-1}, p_i) \,\middle|\, (p_1,..., p_{s-1},1) \in W^s\right\} \quad (6)$$

$$\text{over} \quad p_1,..., p_{s-1}$$

where

$$g_i(p_{i-1}, p_i) = \min\left\{\sum_{j=1}^{p_i} h_i(k_{ij}) \,\middle|\, \sum_{j=1}^{p_i} k_{ij} = p_{i-1}\right\}$$

$$\text{over} \quad k_{i1},..., k_{ip_i} \quad (7)$$

$$W^s = \{(p_1,..., p_{s-1},1) \,|\, 1 \le p_i \le p_{i-1}\}.$$

This statement has some advantages from the point of view of the optimization technique. It is possible to adapt effective methods of the convex programming for solving outlined special cases.

The function $f : X \to R$, $X \subset R^n$, is called discrete-convex (Riismaa 1993; Murota 2003) if for all $x_i \in X (i = 1,...,n+1); \lambda_i \ge 0 (i = 1,...,n+1)$ and $\sum_{i=1}^{n+1} \lambda_i = 1; \ \sum_{i=1}^{n+1} \lambda_i x_i \in X$ holds

$$f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) \le \sum_{i=1}^{n+1} \lambda_i f(x_i).$$

Suppose additionally, that $h_i(k) \ (i = 1,...,s)$ in (7) are discrete-convex functions, and $h_i(0) = 0 (i = 1,...,s)$.

Then $\sum_{i=1}^{s} g_i(p_{i-1}, p_i)$ in (6) is a discrete-convex function (Riismaa 2011).

Now it is possible to solve the inner minimization problem (3) (to find the optimal connections between the adjacent levels) analytically:

$$g_i(p_{i-1}, p_i) = \\ = \left(p_i \cdot \left(\left[\frac{p_{i-1}}{p_i} \right] + 1 \right) - p_{i-1} \right) \cdot h_i \left(\left[\frac{p_{i-1}}{p_i} \right] \right) + \\ + \left(p_{i-1} - p_i \cdot \left[\frac{p_{i-1}}{p_i} \right] \right) \cdot h_i \left(\left[\frac{p_{i-1}}{p_i} \right] + 1 \right) \qquad (8)$$

$(i = 1,..., s-1)$,

$g_s(p_{s-1}, 1) = h_s(p_{s-1})$, $[p]$ - integer part of p.

Denote $p_i \cdot \left(\left[\frac{p_{i-1}}{p_i} \right] + 1 \right) - p_{i-1} = p_i^*$ and

$p_{i-1} - p_i \cdot \left[\frac{p_{i-1}}{p_i} \right] = p_i^{**}$.

Recall

$$k_{ij} = \left[\frac{p_{i-1}}{p_i} \right] \left(\mathrm{j} = 1,..., \mathrm{p_i^*} \right),$$

$$k_{ij} = \left[\frac{p_{i-1}}{p_i} \right] + 1 \quad \left(j = p_i^* + 1,..., p_i^* + p_i^{**} \right)$$

$(i = 1,..., s)$

Certainly $p_i^* + p_i^{**} = p_i$.

To complete the solving of problem (3), (4) it is enough to use (8) for outer optimization problem (3):

$$\min \left\{ \sum_{i=1}^{s} \left(\begin{array}{l} \left(p_i \cdot \left(\left[\frac{p_{i-1}}{p_i} \right] + 1 \right) - p_{i-1} \right) \cdot h_i \left(\left[\frac{p_{i-1}}{p_i} \right] \right) + \\ \left(p_{i-1} - p_i \cdot \left[\frac{p_{i-1}}{p_i} \right] \right) \cdot h_i \left(\left[\frac{p_{i-1}}{p_i} \right] + 1 \right) \end{array} \right) \middle| \; (p_1,..., p_{s-1}, 1) \in W^s \right\}$$

over $p_1,..., p_{s-1}$.(9)

This problem can be solved with method of recursive optimization (Riismaa 2011).

If $h_i(\cdot)(i = 1,..., s)$ are discrete-convex functions, the problem (9) is a discrete-convex programming problem and can be solved with method of recursive optimization or with method of local searching (Riismaa 2003).

Is possible to approximate the functions (8) with

$$g_i(z_{i-1}, z_i) = z_i h_i \left(\frac{z_{i-1}}{z_i} \right) (i = 1,..., s), z_s = 1.$$

Here $z_i \; (i = 1,..., s)$ are not integer and

$k_{ij} = k_i = \frac{z_{i-1}}{z_i} (i = 1,..., s)$ are not integer.

With this approximation and with (5) the problem (3) – (4) transforms to

$$\min_{z_1,..., z_{s-1}} \left\{ \sum_{i=1}^{s} z_i h_i \left(\frac{z_{i-1}}{z_i} \right) \middle| \begin{array}{l} z_0 \geq z_1 \geq ... \geq z_{s-1} \geq 1; \\ z_i \in R^+ (i = 1,..., s-1) \end{array} \right\}$$

(10)

If $h_i(\cdot)(i = 1,..., s)$ are convex functions, then problem (10) is convex programming problem.

Unfortunately there is difficult to estimate this approximation error.

5. ILLUSTRATIVE EXAMPLE: OPTIMIZATION THE STRUCTURE OF MULTI-LEVEL PARALLEL AND DISTRIBUTED PROCESSING SYSTEM

Consider the processing of *n* parts (Riismaa 2011). In case of one processing unit, the overall processing and waiting time for all n parts is proportional to n^2 and is a quickly increasing function. For this reason, the hierarchical system of processing can be suitable. From zero-level (level of object) the parts will be distributed between p_1 first-level processing units and processed (aggregated, packed etc.) by these units. After that, the parts will be distributed between p_2 second-level processing units and processed further and so on. From p_{s-1} (s - 1)-level, the units will be sent to the unique s-level unit and processed finally. The cost of processing and waiting on level i is approximately

$$g_i(p_{i-1}, p_i) = \left(d_i \, l_{i-1} p_{i-1} / p_i \right)^2 p_i + a_i p_i$$
$(i = 1,..., s)$.

Here l_i is the number of aggregates produced by one robot on level i (a number of boxes for packing unit), d_i is a loss unit inside level i, and a_i is the cost of *i*-th level processing unit. The variable parameters are the number of processing units on each level $p_i(i = 1,..., s)$.

The goal is to minimize the total loss (processing time, waiting time, the cost of processing units) over all levels:

$$\min \sum_{i=1}^{s} ((d_i l_{i-1})^2 \left(\begin{array}{l} \left(p_i \left(\left[\frac{p_{i-1}}{p_i} \right] + 1 \right) - p_{i-1} \right) \left[\frac{p_{i-1}}{p_i} \right]^2 + \\ + \left(p_{i-1} - p_i \left[\frac{p_{i-1}}{p_i} \right] \right) \left(\left[\frac{p_{i-1}}{p_i} \right] + 1 \right)^2 \end{array} \right) + a_i p_i)$$

over natural $p_i (i = 1,...,s)$. Here $[p]$ is the integer part of p.

6. CONCLUSION

Many discrete or finite hierarchical structuring problems can be formulated mathematically as a multi-level partitioning procedure of a finite set of nonempty subsets. This partitioning procedure is considered as a hierarchy, nodes of hierarchy correspond to the subsets of partitioning and the relation of containing of subsets defines the arcs of the hierarchy. The feasible set of structures is a set of hierarchies (oriented trees) corresponding to the full set of multi-level partitioning of given finite set.

In mathematical modeling, the choice of variables is important problem. For some special cases here two types of variables are defined. The variables of inner minimization are used to describe the connections between the adjacent levels. The variables of outer minimization are used for the presentation of number of elements on each level. The formalism considered here, enables to state the structure optimization problem as a two-phase mutually dependent discrete optimization problem and to construct some classes of effective solution methods.

Examples of problems of this class are aggregation problems, structuring of decision-making systems, database structuring, multi-level tournament systems, multi-level distribution systems.

The approach is illustrated by a multi-level production system example.

REFERENCES

Laslier, J.- F., 1997. *Tournament solutions and majority voting*. Springer, Berlin: Springer.

Mesarovic, M. D., Macko, D. and Takahara, Y., 1970. Theory of hierarchical multi-level systems. New York: Academic Press.

Murota, K., 2003. *Discrete convex analysis*. Philadelphia: SIAM.

Riismaa, T., 1993. Description and Optimization of Hierarchical Systems. *Automation and Remote Control*, 12, 146 - 152

Riismaa, T., 2003. Optimization of the Structure of Fuzzy Multi-Level Decision-Making System. *Proceedings of International Conference on Modelling and Simulation of Business Systems*, pp. 31 – 34. May 13 – 14, 2003 Vilnius, (Lithuania).

Riismaa, T., Randvee, I. and Vain, J. 2003. Optimization of the Structure of Multi-Level Parallel Assembling. *Proceedings of 7th IFAC Workshop on Intelligent Manufacturing Systems*, pp. 235 – 238. April 6 –8, Budapest, (Hungary).

Riismaa, T., 2011. Recursive algorithm for optimization the structure of complicated logistic, manufacturing and processing environments. *Proceedings of 13rd International Conference on Harbour, Maritime & Multimodal Logistic Modeling nd Simulation*, pp. 99 – 104. September 12 – 14, Rome (Italy).

AUTHORS BIOGRAPHY

Tiit Riismaa is graduated from the Tartu University in 1972 as a mathematician and received a Ph.D. in 1986 in mathematical cybernetics at the Computing Centre of Russian Academy of Science. Now he is the researcher for the Institute of Cybernetics at Tallinn University of Technology. His fields of research include mathematical theory of systems, hierarchical systems, optimization of structures, mathematical programming.

MODELLING AND SIMULATION OF DIRECT STEAM INJECTION FOR TOMATO CONCENTRATE STERILIZATION

Paolo Casoli[(a)], Gabriele Copelli[(b)]

[(a)] Department of Industrial Engineering, University of Parma, Viale G.P. Usberti 181/A, 43124 Parma Italy
[(b)] Interdepartmental Center SITEIA.Parma, University of Parma, Viale G.P. Usberti 181/A, 43124

[(a)]paolo.casoli@unipr.it, [(b)]gabriele.copelli@unipr.it

ABSTRACT

Direct steam injection (DSI) is a sterilization technique which is often used for high viscosity fluid food when the preservation of the quality characteristics and energy efficiency are the priority.

In this work an apparatus for the sterilization of tomato concentrate has been analyzed by means of multidimensional CFD (Computational Fluid Dynamics) models, in order to optimize the quality and safety of the treated food.

A multidimensional two-phase model of steam injection inside a non-newtonian pseudoplastic fluid was adopted to evaluate the thermal history of the product and the steam consumption during the target process.

Subsequently CFD analysis has been extended to examine the effects of the different process parameters (sterilization temperature, steam flow rate, radial and axial temperature profiles, nozzle geometry) on the resulting product.

Result obtained are in agreement with available data acquired in industrial plant.

Keywords: Steam injection, non-Newtonian flow, CFD, Multiphase flow, Design optimization

1. INTRODUCTION

Fluid food products in the agri-food industry are commonly subjected to thermal treatments to ensure safety and improve quality. Process parameters need to be accurately selected and monitored in order to effectively sterilize the product while at the same time avoiding over-processing, which would negatively affect product quality (Abdul Ghani et al 2001). A compromise is therefore required between safety, taste preservation and energy efficiency.

Among the different thermal treatments available, continuous direct steam injection is widely used to quickly raise the temperature of a process media, either for pure heating or for sterilization of the product. The basic idea is to heat the liquid flow by injection of superheated steam from several nozzles, in order to reach homogeneous heating. The main benefit of the direct contact condensation process is the high heat transfer rates and the low product fouling compared with other methods such as heat exchangers. For these reasons, steam injection is used in various applications across in the food industry, such as the sterilization of milk, fruit juices and puree.

In this work, an apparatus for the sterilization of the tomato concentrate has been analyzed by means of multidimensional CFD (Computational Fluid Dynamics) models. CFD has been applied in recent years to model the sterilization problem in order to gain a better understanding of the process and optimize quality while at the same time guaranteeing the safety of the product (Debonis and Ruocco 2009; Marsh 2006).

While the overall energy balance for the process can be easily calculated from the process parameters (steam properties and product flow rate), a numerical model is required for the evaluation of temperature history and distribution Moreover, the available literature on direct steam injection is mostly limited to the case of steam injection in a stagnant liquid, typically water (Sagar 2006; Sachin 2010).

Modeling and simulation of fluid process allow: (a) the optimization of heat transfer in terms of energy efficiency, equipment design, product safety and quality retention; (b) better understanding of the heat transfer process, which helps to control the process and manage deviations, thus reducing production costs and improving quality and safety of the product (Norton 2006).

The rheological behavior of tomato concentrate is non-Newtonian. Due to the high viscosity of the product, enhanced heat transfer surface such as embossed or corrugated pipes are usually employed to ensure a good overall heat transfer coefficient; in this

case CFD simulations may help to understand the effects of pipe shape on temperature distribution.

2. MODEL SET-UP

2.1. Geometrical model

A 3D CAD model of the exchanger was defined: the geometry consisted of a horizontal pipe with a series of equally-spaced radially injectors placed in the first half section (injection section). The second half section (mixing section) presents several hemispherical bosses, whose function is supposed to improve the mixing effect and the heat transfer coefficient and avoid flow stratification.

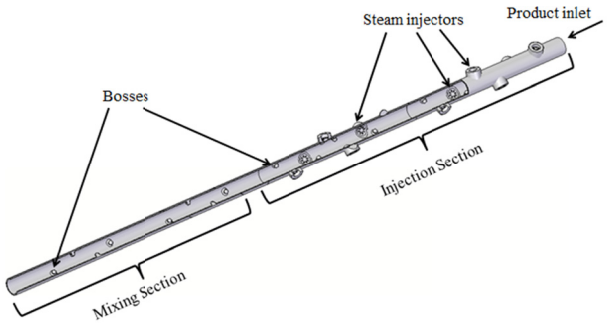


Figure 1: Pipe and injectors geometry

The inner diameter of the pipe is 51 mm; the total length of the exchanger is 2.1 m.

2.2. Rheological Model

Tomato concentrate has a pseudoplastic non-newtonian flow behavior that follows the power-law model (Rozzi 2007):

$$\mu = K\dot{\gamma}^{n-1}, \tag{1}$$

where μ is the dynamic viscosity, K the viscosity consistency, $\dot{\gamma}$ the shear strain rate and n the flow index.

Flow of pseudoplastic fluids is characterized by the generalized Reynolds number

$$Re' = \frac{\rho W^{2-n} D^n}{K' 8^{n-1}} \tag{2}$$

where

$$K' = K\left(\frac{3n+1}{4n}\right)^n. \tag{3}$$

Note that *Re'* is a generalization of the Reynolds number. *Re'* tends to Re when *n* is close to 1.

An exponential dependence on temperature was assumed for *K*, while for *n* a linear dependence on temperature reciprocal has been adopted, according to Trifirò (2001), as follows:

$$K = K_0 \cdot \frac{K_T \cdot 1000/T}{1000} \tag{4}$$

$$n = n_0 + n_T \cdot 1000/T, \tag{5}$$

where n_0 and n_T are parameters depending on the product, while k_0 and n_0 are the viscosity consistency and the flow index at the reference temperature.

The values used for the rheological parameter are reported in Table 1; they have been obtained from typical Hot-break tomato concentrate.

Table 1: Rheological constant of Hot-Break concentrate

K_0	0.3314
K_T	7.8814
n_0	1.114
n_T	-0.3152

In Figure 2 the variation of the rheological parameters with temperature is reported; Equation (1) describes a significant decrease of dynamic viscosity with temperature.

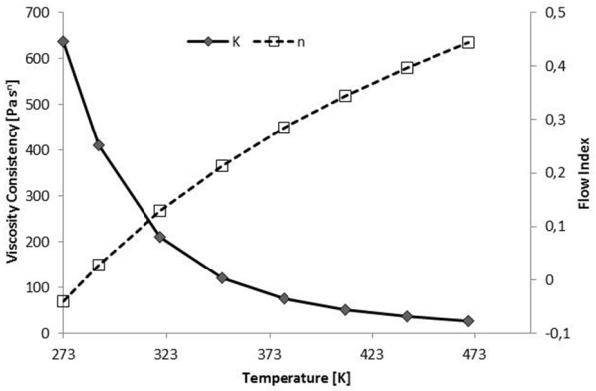


Figure 2: Temperature variation of rheological parameter for tomato concentrate

Rheological and physical properties of water steam are assumed from tabulated values depending on temperature and pressure.

2.3. Numerical model

An Eulerian-Eulerian homogeneous two-phase model was used to describe the flow of tomato concentrate and steam entering the exchanger. Tomato concentrate has been modeled as a continuous phase, while saturated water vapor was modeled as a dispersed phase. In the Eulerian multiphase model the continuous fluid forms a continuous connected region, while the dispersed fluid is present in discrete regions which can be not connected.

The Eulerian-Eulerian model was chosen as each phase is treated separately and it is applicable for wide range of volume fractions (Frank 2007). The phases can move at different velocities, but the model assumes local equilibrium over short spatial length scales, resulting in a strong coupling between phases. A particle model was chosen to describe the interaction between the two phases; the dispersed phase droplets

are assumed to be spherical. All phases are assumed incompressible, accordingly with low kinetic energy; physical properties of all phases are assumed to be temperature dependent.

The flow of tomato concentrate, due to its high viscosity, is laminar; generalized Reynolds number *Re'* is always lower than 20.

The phases are not generally in thermal equilibrium, due to temperature differences across phase boundaries, therefore heat is transferred across phase interfaces via interphase transfer terms (Brennen, 2005). Heat transfer across a phase boundary is described in terms of a heat transfer coefficient. Only heat transfer between phases is taken into account, walls are assumed adiabatic. Two resistance heat transfer model was used to describe combined heat and mass transfer due to steam condensation. Ranz-Marshall correlation was used to evaluate the heat transfer coefficient of the continuous phase (Pechenko 2010); Nusselt number is evaluated through the following equation:

$$Nu = 2 + 0.6Re^{0.5}\,Pr^{0.3} \qquad (6)$$

Inter-phase mass transfer has been tracked by using the thermal-phase change model. Latent heat is not directly specified, but is obtained indirectly as the difference between the enthalpies of the two phases.

The governing equations are the unsteady Navier-Stokes equations in their conservation form. For a multicomponent fluid, scalar transport equations are solved with respect to velocity, pressure, temperature and other quantities of the fluid. An additional equation must be solved to determine how the components of the fluids are transported within the fluid. Each component has its own equation for mass conservation:

$$\frac{\partial \alpha_i \rho_i}{\partial t} + \nabla \cdot \left(\alpha_i \rho_i \boldsymbol{U_{ij}}\right) = \Gamma_i, \qquad (7)$$

which is solved for each phase *i*; U_{ij} is the mass-averaged velocity of fluid component *i*:

$$U_{ij} = \frac{1}{\rho_m}\sum\left(\rho_i U_{ij}\right) \qquad (8)$$

and

$$\rho_m = \sum(\alpha_i \rho_i), \qquad (9)$$

where α_i is the volume fraction of phase *i*. The term Γi in Equation (7) represent the mass source per unit volume into phase *i* due to interphase mass transfer.

The following general form is used to model interphase drag force acting on phase *i* due to phase *j*:

$$M_i = c_{ij}\left(\boldsymbol{U_j} - \boldsymbol{U}_i\right), \qquad (10)$$

where c_{ij} is the drag coefficient. For the particle model (spherical particle), the drag coefficient can be obtained in terms of dimensionless drag coefficient C_D as follows:

$$c_{ij} = \frac{C_D}{8} A_{ij} \rho_i \left|\boldsymbol{U_j} - \boldsymbol{U}_i\right|. \qquad (11)$$

Energy equation is also modified adding the contribution S_E due to steam condensation inside the continuous phase, depending on Γi in Eq. (7)

$$S_E = \Gamma_i \Delta h \qquad (12)$$

2.4. Initial and boundary conditions

Temperature and flow rate have been set for tomato concentrate at the exchanger inlet. A constant temperature of 75°C was set for tomato concentrate at the inlet of all simulations.

During normal operation the exchanger is under a constant head of 5 bar. The absolute level of pressure at the exchanger outlet was assumed to be 5 bar. Steam pressure and temperature have been fixed at the injectors inlet accordingly to normal operation settings of the exchanger. All walls are assumed adiabatic, with imposed no slip condition.

2.5. Simulation Details

An unstructured tetrahedral grid of 1.3×10^6 nodes was employed; meshes were created with ICEM CFD Software. Grid density has been increased near walls and at injector-pipe junctions. To improve stability and reduce the overall number of elements the hexa-core option was activated.

The commercial code Ansys CFX© was employed for all simulations. A coupled solver was used to solve governing equations. A high resolution discretization scheme was used for the continuity, energy and momentum equations, while the upwind discretization scheme was employed for the volume fraction equations. Tolerance was set for all variables at 1×10^{-3}.

3. RESULTS AND DISCUSSION

3.1. Injector detail simulations

A first set of simulations were carried out to estimate the flow characteristic of the injectors using a simplified computational domain. As shown in Figure 3, the computational domain consist of a 200 mm pipe section with a single radial injector. The junction between the main pipe and the injector is obtained with six radial slots, through which steam enters in the exchanger.

Different simulations were carried out with imposed tomato flow rates ranging from 1500 l/h to 20000 l/h; all simulations were performed with vapor mass flow from 0.002 kg/s to 0.04 kg/s.

Vapor penetration inside tomato concentrate is heavily dependent on both tomato and vapor flow rate (Figures 4 and 5).

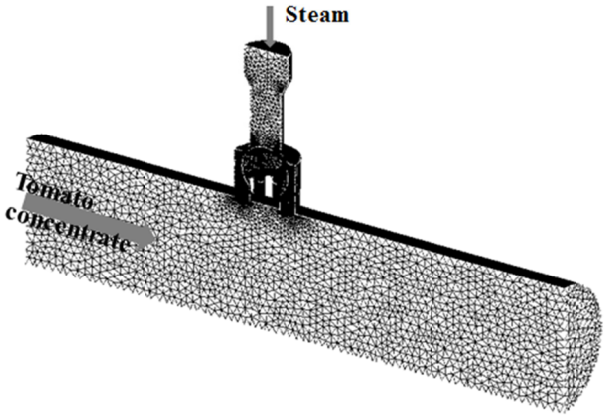


Figure 3: Simulation grid around injector

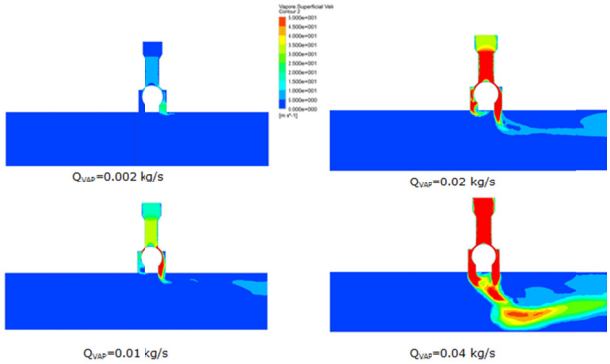


Figure 4: Dispersed phase (steam) superficial velocity contours at different steam flow rates (tomato flow rate 1500 l/h)

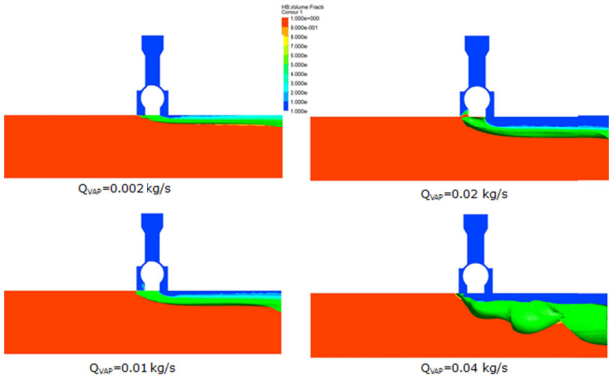


Figure 5: Tomato volume fraction contours and iso-surface (volume fraction = 0.5, green) at different steam flow rate (tomato flow rate 20000 l/h)

Steam penetration inside tomato concentrate is extremely low for high product flow rates. A well distributed radial pattern of injector is critical to avoid over-processing of fluid near the walls.

Simulations allowed to compute the flow characteristic of the injector depending on tomato flow rate, as shown on figure 6 influence on flow characteristic of tomato flow rate is relevant only for high values of steam mass flow rates, typically over 0.02 kg/s.

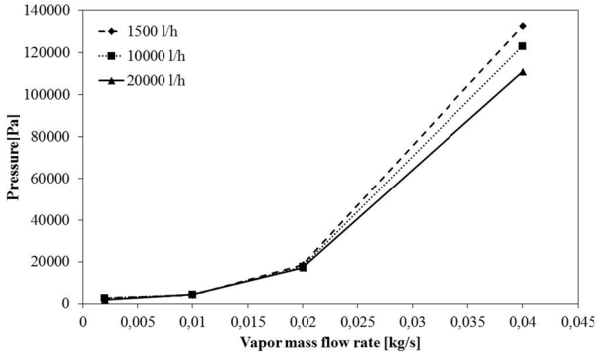


Figure 6: Injector flow characteristic (pressure vs mass flow rate) at different product flow rate

Knowledge of the injector characteristic allows the control of the amount of steam injected inside the exchanger during the sterilization process.

3.2. Injection simulations

Simulations of the complete geometry were carried out to investigate the effects of steam flow rate on the temperature history of product. Constant steam mass flow rates at the injector inlet were imposed. Since not all the injectors are active during the sterilization, the effects of different number of active injectors were investigated. Figure 7 shows the simulation results of different active injectors working with same steam mass flow rate of 0.02 kg/s. Inlet temperature of tomato and steam are respectively 75°C and 158°C. Steam temperature is referred to saturated conditions at average exchanger absolute pressure (5 bar). Tomato concentrate volumetric flow rate at inlet is fixed for all the simulations at 15000 liters per hour.

Due to high temperature difference and a low value of steam flow rate, the steam is completely condensed before reaching the exchanger outlet. Figure 8 shows a contour representation of tomato volume fraction on an axial cut plane. A complete condensation which occurs near the end of the injection section of the exchanger, can assure a gradual temperature rise and the mixing section reduces temperature difference when the product exit the exchanger.

In the industrial practice it is extremely difficult and expensive to measure and adjust the steam flow rate. Steam flow rate is usually set indirectly by adjusting the pressure in the manifold before the injectors. Due to high pressure losses within the pipe, constant steam pressure at the injectors inlet normally produce different steam flow rate in the exchanger.

Injection with constant pressure at the injector inlets was simulated to investigate the effect of steam flow rate on the temperature distribution at the product outlet. Constant pressure boundary condition was applied to eight equally- spaced injectors along the exchanger; the remaining four injectors were modeled using a wall condition. The values imposed are referred to exchanger outlet absolute pressure.

Figure 9 shows the tomato volume fraction at different steam relative pressure at the injectors. In all

the simulations the pressure imposed at the injector inlet of the first injectors is lower than tomato pressure inside the exchanger; an automatic wall condition is set by the solver; due to high pressure difference on the last injectors, steam flow rates are extremely high and non-condensed steam exits through the exchanger outlet.

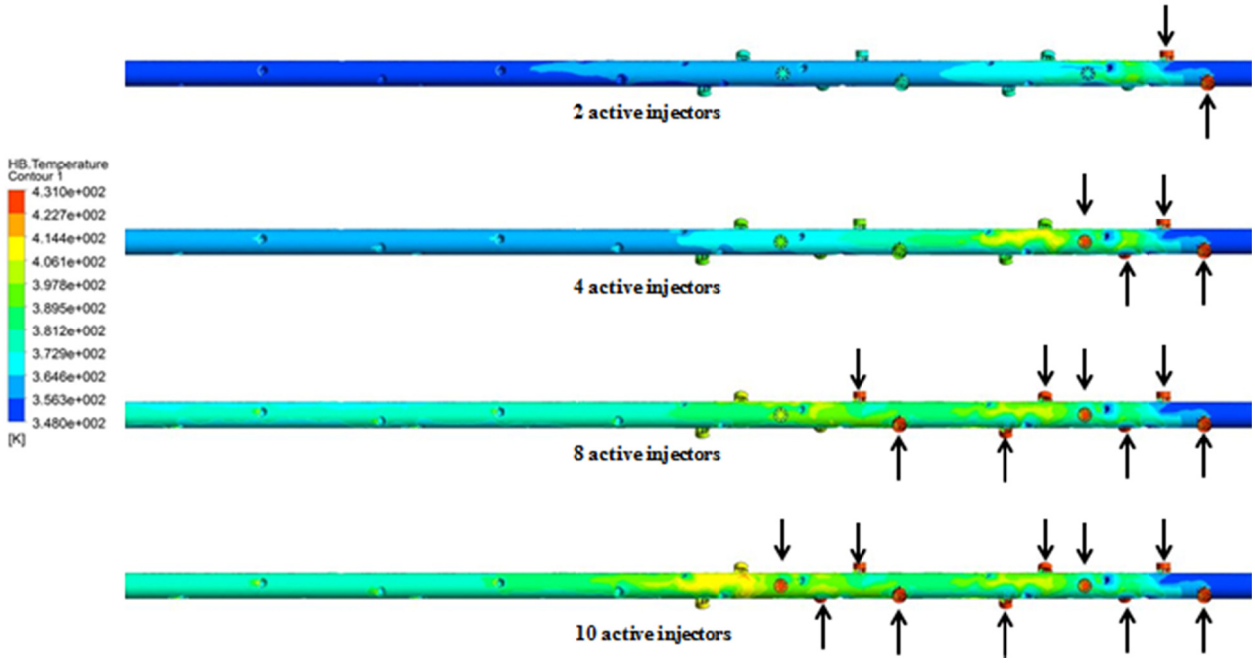


Figure 7: Tomato Temperature contours with 2, 4, 8, 10 active injectors

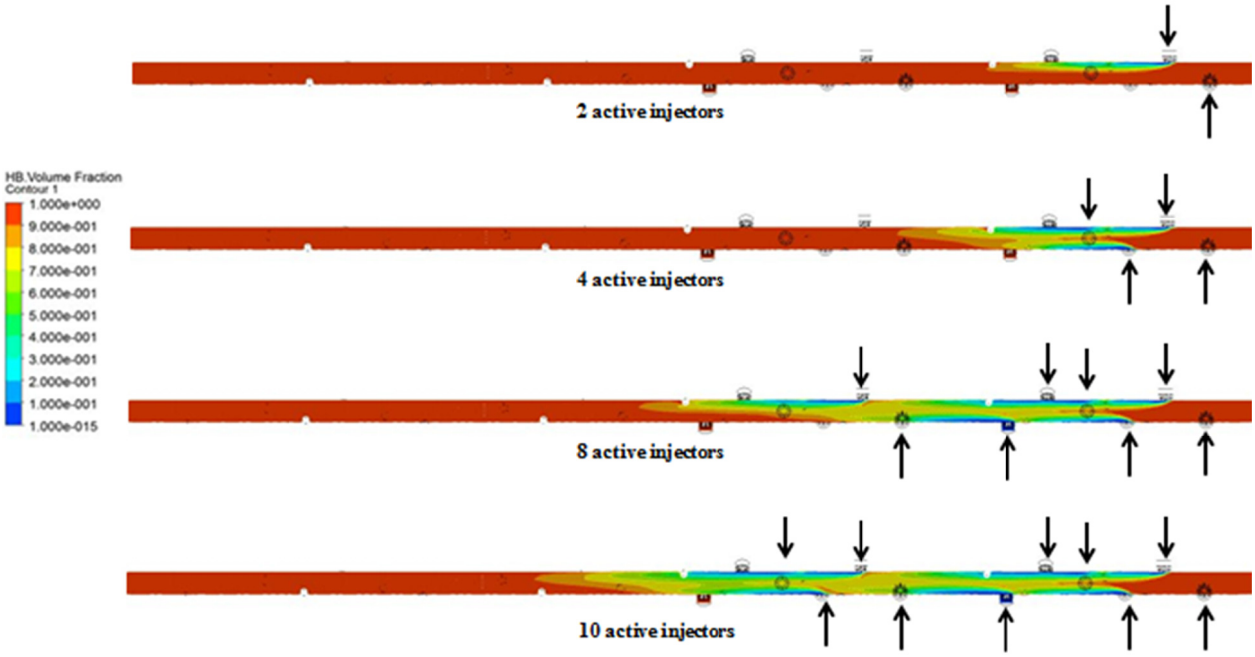


Figure 8: Tomato volume fraction contours with 2, 4, 8, 10 active injectors

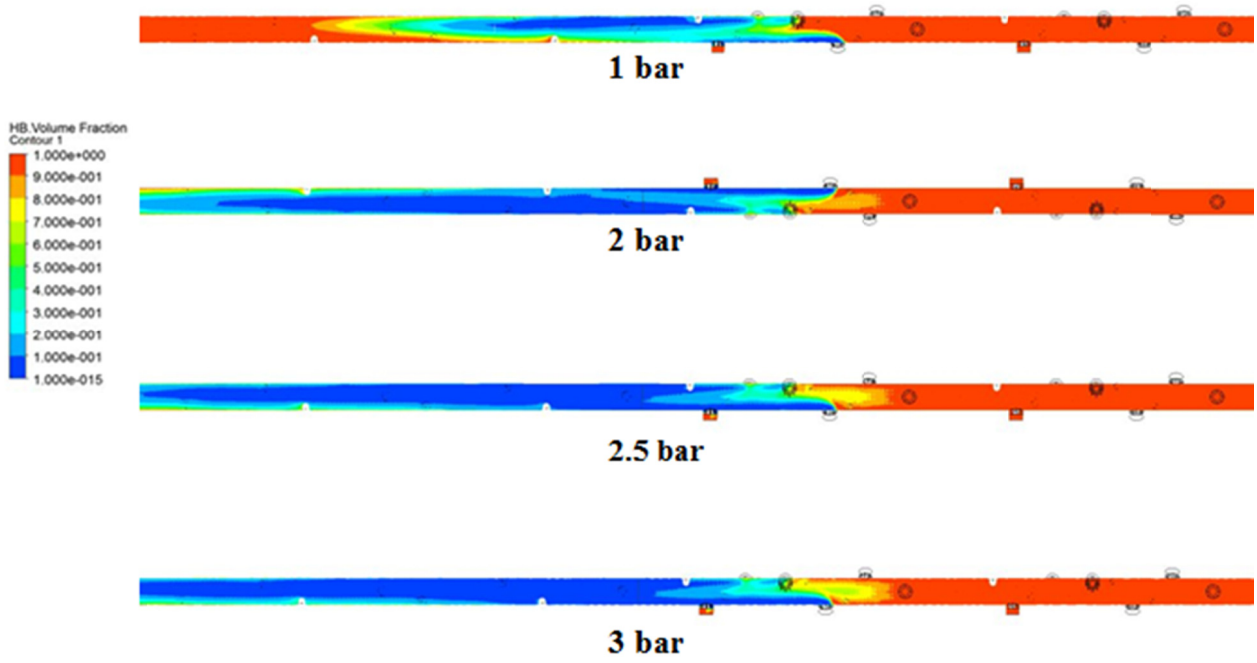


Figure 9: Tomato volume fraction contours with pressure value set to 1, 2, 2.5, 3 bar on injectors inlet

The mass flow rate averaged temperature of tomato on the exchanger outlet, with maximum and minimum temperature on the same section have been computed. Figure 10 shows the temperature values for each simulation with different injection pressure.

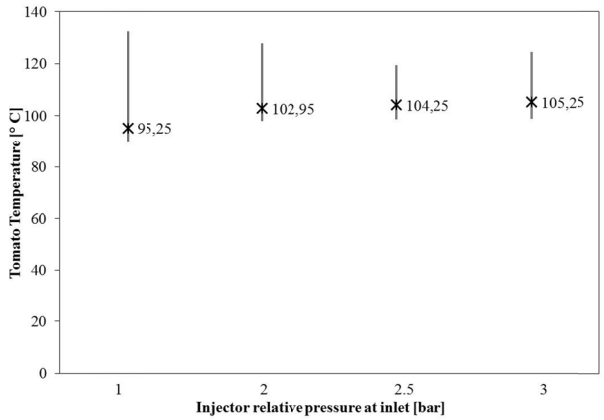


Figure 10: Mass flow rate averaged temperature (tomato) on the exchanger outlet. Lines show maximum and minimum values on the same section.

Temperature differences on outlet sections are extremely high; maximum values over 125°C can result in degradation of the taste of the product. Average temperature around 105°C on the exchanger outlet is the typical set point for sterilization processes.

As shown in Figure 10, for injection pressure higher than 2 bar the average temperature remains constant at around 105°C.

Considering same operation parameters of the exchanger, values obtained with CFD simulations of average outlet temperature are in agreement with available data acquired in industrial plant.

A last set of simulation has been performed to analyze different pressure distributions on the injectors. For confidential reasons the pressure distributions on the injectors are not reported. .

The simulation results 1 are shown in Figures 11 and 12. Despite high global steam flow rate, complete condensation is reached before the exchanger outlet.

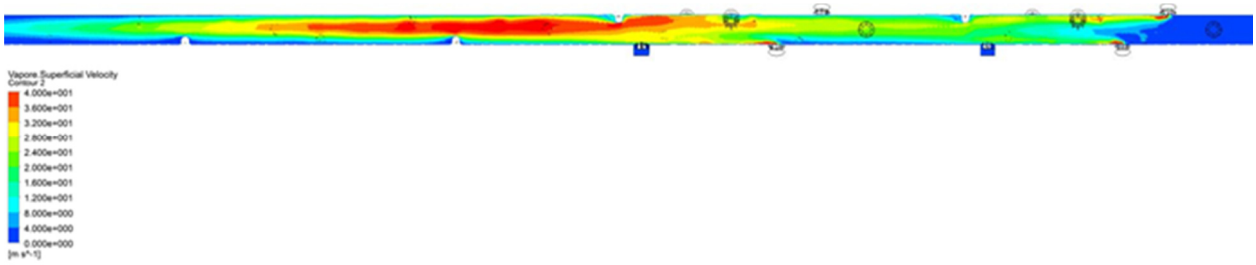


Figure 11: Steam superficial velocity contour, axial plane .

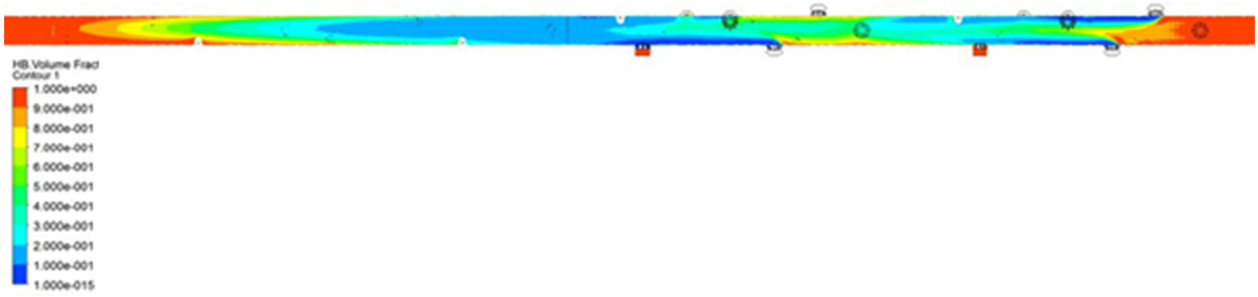


Figure 12: Tomato concentrate volume fraction contour, axial plane

Figure 13 shows that the temperature difference for three different pressure patterns imposed on the injector inlets is always lower than 5°C.

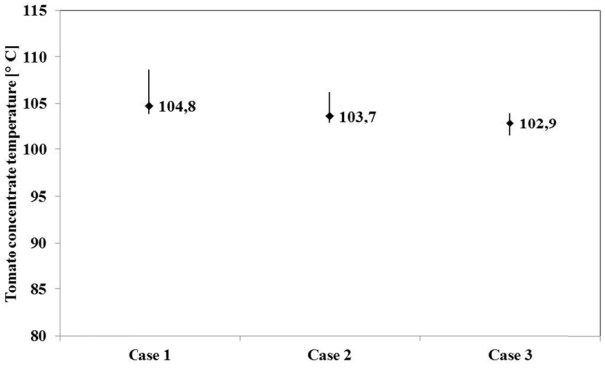


Figure 13: Mass flow rate averaged temperature (tomato) on the exchanger outlet. Lines show maximum and minimum values on the same section.

Compared to Figure 10, the decrease in the temperature difference on the exchanger outlet section is more than 4 times lower. Therefore a suitable pressure distributions on the injectors could reduce the temperature difference at a maximum value of 5°C.

As previously remarked, avoiding high temperature values ensures better product quality and avoids degradation of food taste. Decreasing steam pressure values on injectors inlet allowed to reach average product temperature on exchanger outlet comparable to constant pressure simulation; as result the sterilization is achieved with more uniform temperature at outlet.

4. CONCLUSIONS

In this work an apparatus for the sterilization of tomato concentrate was analyzed by means of multidimensional CFD models, in order to optimize quality and safety of the treated food.

A flow domain was created from the exchanger geometry and a flow model was built by using the ANSYS CFX® commercial code. A multiphase approach was adopted and a rheological model of tomato concentrate was created.

First simulations were carried out to investigate the flow characteristics of steam injectors. The knowledge of the injector characteristic allows to precisely control the amount of steam injected inside the exchanger during sterilization process.

In the second part of the work, simulations of the complete geometry were carried out to investigate the effects of steam flow rate on the temperature history of product. Different boundary conditions at the injector inlets were used to investigate the effects of the different process parameters. High differences in the temperature of the product at the exchanger outlet were found with constant pressure applied to the injector inlets.

Better results were reached by applying a suitable pressure distributions on the injector inlets: starting from temperature differences higher than 25°C on the outlet section of the exchanger, the simulation allowed to identify better injection settings and temperature differences lower than 5°C were obtained.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the active support of this research by Ing. A. Rossi Impianti Industriali S.r.l.

REFERENCES

Brennen, E. B., 2005. *Foundamental of multiphase flow*. Cambridge University Press.

Maria Valeria De Bonis, Gianpaolo Ruocco 2009. Heat and mass transfer modeling during continuous flow processing of fluid food bydirect steam injection *International Communications in Heat and Mass Transfer*, 37 (2010), 239-244.

S. Rozzi, R. Massini, G. Paciello, G. Pagliarini, S. Rainieri, A. Trifiro 2006. Heat treatment of fluid foods in a shell and tube heat exchanger: Comparison between smooth and helically corrugated wall tubes *Journal of Food Engineering 79 (2007)* 249–254.

Curtis Marsh, Denis Withers 2006. CFD modeling of direct contact steam injection. *Fifth International Conference on CFD in the Process Industries CSIRO, Melbourne, Australia.*

Abdul Ghani, A.G., Farid, M.M., Chen, X.D., Richards, P., 2001. Thermal sterillization of canned food in a 3-D pouch using computational fluid dynamics. *Journal of Food Engineering* 48 (2), 147-156.

Sagar S. Gulawani, Jyeshtharaj B. Joshi, Manish S.Shah, Chaganti S. RamaPrasad, Daya S. Shukla 2006. CFD analysis of flow pattern and heat

transfer in direct contact steam condensation. *Chemical Engineering Science 61* (2006) 5204 – 5220.

Sachin K. Dahikar, Mayur J. Sathe, Jyeshtharaj B. Joshi, 2010. Investigation of flow and temperature patterns in direct contact condensation using PIV, PLIF and CFD. *Chemical Engineering Science 65* (2010) 4606-4620.

Pecenko A. 2010. *Numerical simulation methods for phase-transitional flow*. Thesis (PhD). Eindhoven University.

Frank T., 2007. Simulation of flashing and steam condensation in subcooled liquid using ANSYS CFX. *5th FZD & ANSYS MPF Workshop,* April 2007.

Trifirò, A., Gherardi, S., Castaldo, D., 1991. Use of rheological parameters for pressure losses calculation in continuous heat exchangers for tomato paste. *Journal of Food Engineering 23* (1991) 233.

Norton, T., Sun Da-Wen 2005 Computational Fluid Dynamics (CFD) – an effective and efficient design and analysis tool for the food industry: a review *Trends in Food Science & Technology 17* (2006) 600-620.

ANSYS CFX-Solver Theory Guide. ANSYS Ltd, 2010.

AUTHORS BIOGRAPHY

Paolo Casoli is Associate Professor of fluid machinery for food industries at the Faculty of Engineering of the University of Parma

Gabriele Copelli is a Research Fellow of the Centro Interdipartimentale SITEIA.PARMA

$_E$SBMTOOLS: PYTHON TOOLS FOR ENRICHED STRUCTURE BASED MODELING

Benjamin Lutz[a,b] , Claude Sinner[a,b], Geertje Heuermann[a,b], Abhinav Verma[b], Alexander Schug[b,*]

[a] Department of Physics, Karlsruhe Institute of Technology, Karlsruhe, Germany

[b]Steinbuch Centre for Computing, Karlsruhe Institute of Technology, Karlsruhe, Germany

[*] alexander.schug@kit.edu

ABSTRACT

Biomolecular simulations provide a computational microscope to dynamically visualize biomolecular systems with atomic resolution. Given the advances in speed of computational resources, simulations can complement experiments and help understand the relationship of folding, structure, and function of proteins, structured RNA, or DNA. Structure-based models (SBM) provide a computationally inexpensive tool to study the folding and structural assembly of (macro)biomolecules. Their theoretical foundations are energy landscape theory and the principle of minimal frustration. Here, we present $_E$SBMTools: python tools that assist to setup and analyze structure-based simulations of proteins and nucleic acids, both at the C_α and all-atom level. The tools interface with GROMACS and support its standard output formats. Information from other sources like bioinformatics or experimental data can be added as enrichments. One example would be docking protein complexes out of the composing individual known proteins plus bioinformatically derived information of the inter-protein interface contacts.

Keywords: protein folding, protein structure prediction, structure-based model, setup and analysis

1. INTRODUCTION

In the last few decades, important progress has been achieved in the field of biomolecular sciences. Experiments can explore these systems in high detail (Sadqi, Fushman et al. 2006) even on the single molecule level (Mickler, Dima et al. 2007; Gambin, VanDelinder et al. 2011). The resolution of experiments, however, is still often limited by technical constraints. Simulation techniques based on Monte Carlo (Schug and Wenzel 2004; Schug, Herges et al. 2005; Verma, Schug et al. 2006; Verma and Wenzel 2009; Perez-Sanchez and Wenzel 2011) or Molecular Dynamics (MD) (Adcock and McCammon 2006; Lee, Hsin et al. 2009) have similarly advanced and can complement experiments (Lange, Lakomek et al. 2008; Gambin, Schug et al. 2009). The advances in MD result from improved hardware performance, utilizing new architectures like Cell (Olivier, Prins et al. 2007), GPU or specialized supercomputers (Shaw, Maragakis et al. 2010), better force fields (Lindorff-Larsen, Piana et al. 2010) and more effective simulation algorithms (Hess, Kutzner et al. 2008). Taken together, this enables the detailed exploration of dynamical properties for a biomolecular system. This step is crucial for understanding its function at an atomic resolution. Deeper insights in the dynamics of proteins and structured RNA result from, for example, the investigation of conformational changes in the systems of interest (Okazaki, Koga et al. 2006; Schug, Whitford et al. 2007) and challenges as complex as protein folding become accessible (Kussell, Shimada et al. 2002; Thirumalai, Klimov et al. 2002; Onuchic and Wolynes 2004; Thirumalai and Hyeon 2005; Dill, Ozkan et al. 2008; Schug and Onuchic 2010). As many dynamic processes occur on slow μs to ms timescales, however, MD simulations struggle to reach these scales given a time step of 1-2fs.

In order to reach the desired timescales, coarse-graining has proven to be a reliable approach by reducing the complexity of the simulated system (Klimov and Thirumalai 1998). This can occur on multiple scales. For example, coarse-graining the biomolecular system to a C_α level reduces each amino acid to a single bead (Clementi, Nymeyer et al. 2000). This approach reduces the number of particles in the biophysical system by around two orders of magnitude compared to an all-atom system with explicit water molecules. The price paid is reduced insight into, for example, details of side-chain interactions or the influence of base-pairing and stacking interactions in RNA or DNA.

The structure-based model (SBM) approach is based on energy landscape theory and the principle of minimal frustration for protein folding and structured RNA (Onuchic and Wolynes 2004; Schug and Onuchic 2010). Accordingly, biomolecular folding occurs on a funneled energy landscape with its free-energy minimum in the native fold (Bryngelson, Onuchic et al. 1995). The SBM-Hamiltonian (Clementi, Nymeyer et al. 2000; Lammert, Schug et al. 2009; Whitford, Noel et al. 2009; Whitford, Schug et al. 2009) is directly based on this native fold. A crucial part of these force fields are interactions of amino acid pair contacts described as a contact map, i.e., of a matrix of spatially close interacting amino acids (Noel, Whitford et al. 2012).

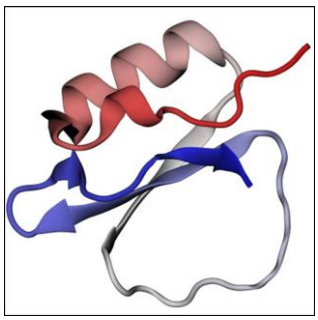


Figure 2: Cartoon representation of CI-2 (PDB-code 2CI2)

Typically, the physicochemical details of each individual interaction are condensed into simple terms like Lennard-Jones or Gaussian interactions(Lammert, Schug et al. 2009) further reducing computational complexity. SBM have been parameterized for different resolutions from the C_α level (Clementi, Nymeyer et al. 2000), over the $C_\alpha C_\beta$ (Oliveira, Schug et al. 2008)to the all-atom level (Whitford, Noel et al. 2009). In spite of their simplified energetics, they have shown good agreement with experimental measurements (Clementi, Jennings et al. 2001; Schug, Whitford et al. 2007; Gambin, Schug et al. 2009; Schug, Weigt et al. 2009; Baxter, Jennings et al. 2012).

When enriched with bioinformatic information (Weigt, White et al. 2009) as additional distance constraints and more detailed biophysical force fields, even accurate predictions of three-dimensional structures of protein complexes (Schug, Weigt et al. 2009), globular proteins (Sulkowska, Morcos et al. 2012), trans-membrane proteins (Hopf, Colwell et al. 2012) or active conformations (Dago, Schug et al. 2012) have been made. Similarly, integrating distance constraints from experimental measurements like FRET or EM density maps could be included.

The main purpose of the present tool collection is to facilitate the scriptable setup of huge systems in SBM simulations for the GROMACS (Hess, Kutzner et al. 2008) software package and enhancing these simulation with information from other sources. This reduces the effort for a single simulation run. The tools can be included in automated workflows for a wide range of biophysical investigations. The tools include routines that run post processing protocols of standard analysis procedures, like contact map analysis, Q value generation, Phi value or RMSD evaluation. In this paper we discuss the methods that SBM are based on, the implementation of pre and post processing functionality of ESBMTools, and give an overview over several exemplary scenarios that apply their functionality.

2. THEORETICAL BACKGROUND

Molecular dynamics simulation technique solves Newtonian equations of motion via numerically integrating over time. The central characterization of a system of interest is introduced by a potential for the equations of motion. We describe the underlying potentials for the all-atom and the C_α case. Furthermore, we discuss the C_α method as a coarse-grained approach in the context of protein simulation. The contact map as the substantial ingredient to SBM is described afterwards.

2.1. Structure-based Potential

The most basic information that characterizes a molecular dynamics simulation is aggregated in its potential from which the force field for the Newtonian equations of motion is derived. The all-atom formulation of the structure-based potential (Whitford, Noel et al. 2009) reads as

$$\begin{aligned} V_{AA} = & \sum_{\text{bonds}} K_b(r - r_0)^2 + \sum_{\text{angles}} K_a(\theta - \theta_0)^2 \\ & + \sum_{\text{impropers}} K_i(\chi - \chi_0)^2 + \sum_{\text{dihedrals}} K_d f_d(\phi) \\ & + \sum_{\text{contacts}} K_c\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - 2\left(\frac{\sigma_{ij}}{r}\right)^{6}\right] + \sum_{\text{non contacts}} K_{nc}\left(\frac{\sigma_{nc}}{r}\right)^{12}, \end{aligned} \tag{1}$$

where the dihedral or torsional angle potential is given by

$$f_d(\phi) = [1 - \cos(\phi - \phi_0)] + \frac{1}{2}[1 - \cos(3(\phi - \phi_0))], \tag{2}$$

and K_b, K_a, K_i, K_d, K_c and K_{nc} are the corresponding force constants. $r_0, \theta_0, \chi_0, \phi_0$ and σ_{ij} are taken from the native structure. Accordingly, the potential has its minimum at the native conformation.

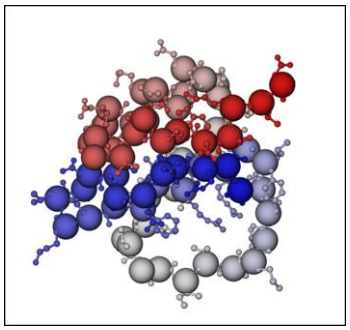


Figure 1 All-atom and C_α representation of CI-2

2.2. Coarse-graining

We present an approach of coarse-graining for protein systems that reduces each amino acid of a protein to a single bead at the position of the protein's C_α atom. This reduction decreases the number of particles in the computational system. Removing an explicit water representation reduces the number of atoms by one order of magnitude and reducing an all-atom (non-Hydrogen) representation to the C_α level results in another order of magnitude less atoms. Therefore, the approach results in computationally far less demanding simulations. The structure-based potential for proteins in a coarse-grained C_α formulation (Clementi, Nymeyer et al. 2000) reads as

$$V_{C\alpha} = \sum_{\text{bonds}} K_b(r - r_0)^2 + \sum_{\text{angles}} K_a(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} K_d f_d(\phi) + \sum_{\text{contacts}} K_c\left[5\left(\frac{\sigma_{ij}}{r}\right)^{12} - 6\left(\frac{\sigma_{ij}}{r}\right)^{10}\right] + \sum_{\text{non contacts}} K_{nc}\left(\frac{\sigma_{nc}}{r}\right)^{12}, \quad (3)$$

where $f_d(\phi)$ is the dihedral potential as defined in Eq. 2 and structural information is incorporated as in Eq. 1, correspondingly. Relative to Eq. 1, the potential in coarse-grained formulation exhibits two changes: The potential lacks terms for improper dihedral angles since the C_α formulation loses the possibility to model planarity within the sidechain. Secondly, the Lennard-Jones potential is changed from its standard 6-12 formulation to a 10-12 formulation. Equivalent Hamiltonians can be found using Gaussian potentials (Lammert, Schug et al. 2009).

It should be emphasized that the negligence of structural detail in the C_α method, makes it insufficient to describe nucleic acid chains. Their base-base interactions (pairing and stacking) play a crucial role in nucleic acid strands, both for structural and dynamic investigations, and cannot be neglected (Thirumalai and Hyeon 2005; Whitford, Schug et al. 2009).

2.3. Contact Map

The information of bonded interactions (bonds, angles and dihedrals) is complemented by contact information (Noel, Whitford et al. 2012). This information is aggregated in the contact map of a biomolecular structure. In its simplest form, a contact between two atoms is formed if the distance σ_{ij} between the two of them is below a certain threshold (typically 4-5Å). Typically, a minimal distance in sequence can be required in case of proteins, while nucleic acids need to be able to form contacts between neighboring residues as stacking interactions. Contact information is represented by repulsive and attractive terms of a Lennard-Jones potential, as denoted in Eq. 1 and 3. All other possible pairings of atoms are assigned to a repulsive Lennard-Jones term that is characterized by the exclusion radius σ_{nc}.

3. SIMULATION SETUP

The simulation setup consists of the standardized generation and, as the case may be, customizable manipulation of the coordinate and topology file of a biomolecular system. The generation is based on a PDB conform structure file and several user defined parameters. The native conformation is taken from the structure file and combined with an XML based topology that defines bonded interactions to create the following input files for a GROMACS(Hess, Kutzner et al. 2008) (4.5.4) simulation. The $_E$SBMTools package is written in python 2.7 and requires biophython (Cock, Antao et al. 2009) and scientific python. It can be downloaded from sourceforge. Examples are included.

3.1. Coordinate File

The coordinate file for a GROMACS simulation contains the atoms represented by their names, residue names and atom types in combination with their individual Cartesian coordinates. The coordinates represent the native conformation which introduces the conformation of minimal potential energy in the SBM. These coordinates are generated automatically from a PDB conform data file of the biomolecular system of interest. In case of coarse-graining, the only atom type that is present in the coordinate file is the C_α type. The corresponding coordinates are again taken from the PDB structure file.

3.2. Topology File

The topology file introduces the biomolecular system by a list of its atoms. Referring to the list of atoms, the topology file also contains the force constants and geometrical quantities of equilibrium of the SBM potential. The geometrical quantities are calculated from the system's coordinates. The particular geometrical associations of bonded interactions are defined in an XML file which can be adapted to user defined scenarios. Topologies for amino acids and nucleic acids are provided with $_E$SBMTools, but the topologies are easily expandable by the user's own defined topological rules, if, e.g., ligands are needed. The contact map as part of the potential is also included in the topology file.

3.3. Look-up Table

In case of a coarse-grained simulation, the Lennard-Jones potential terms is formulated with powers of 10-12 instead of 6-12. GROMACS offers the introduction of such modifications on the standard potential expressions. To this end, the user has to provide a look-up table for GROMACS to specify the desired modifications. $_E$SBMTools generates a file (table.xvg) that contains such a look-up table in the format required by GROMACS.

3.4. Configuration File

A standard configuration file (md.mdp) for a GROMACS simulation can be created by $_E$SBMTools in order to be equipped with a complete set of required files for a molecular dynamics simulation. The creation of this configuration file is also customizable by, e.g., setting the number of integration steps, the temperature or generation seeds for random events.

3.5. Input Modification

Existing input files can be modified with several functions in $_E$SBMTools. The manipulation of existing contacts and the introduction of new contacts, atoms, bonds, angles or dihedral angles is a desirable feature in the course of SBM simulation. This approach enables the user to set up a heterogeneous potential or a combination of two separately generated SBM systems.

3.5.1. Contact Map Modification

The contact map of a SBM plays a crucial role in introducing tertiary structural elements in the biomolecular system. Contact map extension to an existing protein-protein interface can model, for example, structural transitions correctly by adding a second local minimum. To this end, ESBMTools offers a variety of functions that facilitate the access to a given contact map and the scriptable generation of new entries and modification of existing entries. Externally generated contact maps can be read in, random subsets of contact maps can be generated, contact strengths be modified and the force constant in the Lennard-Jones terms for user-defined ranges of contact partners can be modified.

3.5.2. Merging & Enrichment

In order to combine two existing SBM systems, it is necessary to merge their coordinate and topology files. ESBMTools provides functionality to generate a single SBM for the combination of two existing systems to allow, for example, simulating protein complex formation. The two systems stay isolated from each other in terms of interactions by a plain merging procedure. This functionality can be complemented by introducing additional inter-molecular contacts from non-structural sources. Examples are distance contraints derived from bioinformatic information(Weigt, White et al. 2009) like in MAGMA (Schug, Weigt et al. 2010) to predict a coarse-grained protein complex model, which can be afterwards relaxed in more detailed biophysical force fields to accurately predict three-dimensional structures of protein complexes(Schug, Weigt et al. 2009), active conformations(Dago, Schug et al. 2012) or globular proteins(Sulkowska, Morcos et al. 2012). Other possibilities would be including distance constraints from experimental measurements like FRET, small angle X-ray scattering (Jamros, Oliveira et al. 2010) or cryo EM (Whitford, Ahmed et al. 2011).

4. SIMULATION ANALYSIS

Simulation analysis is mainly based on the trajectory as outcome of a molecular dynamics simulation. ESBMTools provides interfaces to several GROMACS evaluation extensions that process simulation data. The output of these extensions can be read in for further processing. We present a variety of possible post processing scenarios that can be conducted in the following.

4.1. RMSD

The root mean square deviation (RMSD) of a trajectory in each dumped frame describes the average deviation of the current structure in comparison to a reference structure, often the native structure. These values are a measure for fluctuations in the course of a simulation. This analysis can serve as a rough estimate for a folded or unfolded state of a biomolecular state.

4.2. Q Value Evaluation

The Q value evaluation is a standard analysis for SBM simulations(Cho, Levy et al. 2006) where the Q value is defined for each simulation frame as the fraction of formed native contacts. By this means, the Q value itself can serve as a reaction coordinate that represents a mapping on the folding progress in time if the Q value is monotonically increasing over time. ESBMTools provides several functions that filter and bin Q value trajectories for requested ranges of involved atoms or residues in order to analyze, e.g., folding characteristics in biomolecular systems.

4.2.1. Contact Maps

Contact maps are a standard visualization technique of the present contacts not just in a native structure but also along a simulated trajectory. To this end, ESBMTools offers the functionality to plot contact maps from a given topology as well as a user-defined set of contact maps from a trajectory. These contact maps along a trajectory can be also arranged to a movie by standard movie encoding software.

4.2.2. State Population

States are visualized as a histogram of Q (or other reaction coordinates) (Cho, Levy et al. 2006) timedirectly from a trajectory.

4.2.3. Φ Value

The **Φ-**value analysis investigates the stabilizing influence of amino acids on the transition state in a two-state folding protein(Fersht 1995). This analysis originates from an experimental approach. The experimental phi value analysis can be translated in a computational analogue that is based on the probabilities of occupancy of the three possible states folded, unfolded and transition state.

5. EXAMPLES

We illustrate the variety of preparation and investigation related procedures provided by ESBMTools.

ESBMTools generates all necessary files for a molecular dynamics simulation with GROMACS. The contact map is calculated from the native state(see Fig. 3).

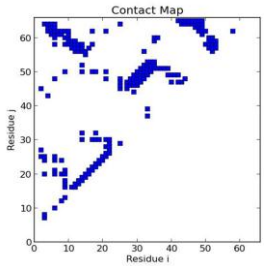


Figure 3: C_α based Contact map of CI-2

After running a simulation, user-defined frames from the trajectory of a denaturing protein (Fig. 4) can be

plotted in order to investigate the dynamics of folding paths by the order of opening structural elements.

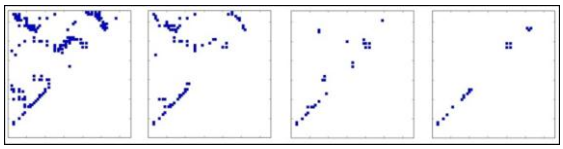


Figure 5: Contact maps of denaturing CI-2. The leftmost contact map is in the folded state, while the contacts maps to the right are found at lower Q-values (i.e. more unfolded).

A statistically relevant analysis is the investigation of temporal progress of the Q value as the number of formed native contacts (Fig. 5, Top). This value identifies the nativeness of the trajectory at every timestep. A histogram of the Q value distribution (Fig. 5, Bottom) embodies the frequency distribution of the system's state variable. Therefore, it serves as a starting point from which the free energy can be calculated.

As part of a more specific analysis, $_{E}$SBMTools filters the trajectory's Q values for user-defined ranges of involved contact partners. The investigation of melting characteristics in the context of RNA hairpins (Fig. 6) can be based on a Q value filtering for base pair contacts in combination with averaging over several hundreds of trajectories.

Our last example illustrates the usage of $_{E}$SBMTools for a **Φ**-value analysis. Fig. 7 shows the **Φ** values calculated from a simulation based on a SBM simulation run and evaluated by our toolset.

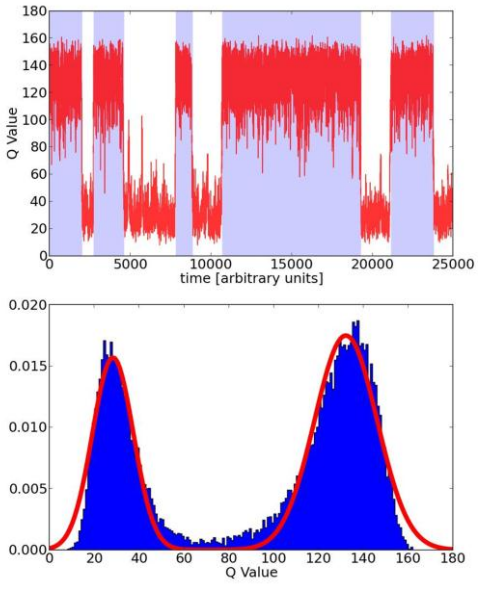


Figure 4: Top: Temporal progress of the Q value for CI-2 close to the folding temperature that indicates folded (light blue) and unfolded (white) domains. Bottom: Histogram of the population of Q values over a trajectory.

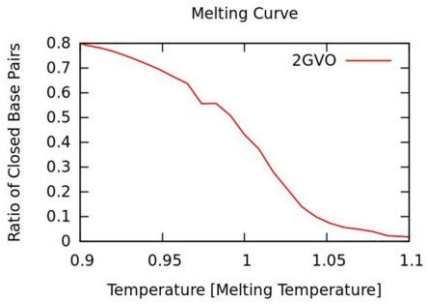


Figure 7: Melting curve of RNA hairpin 2GVO

6. CONCLUSION

We present $_{E}$SBMTools as a customizable, extendable and flexibly scriptable python package that aims at setting up and evaluating SBM molecular dynamics simulations with GROMACS. The toolbox provides a diverse collection of functions, that can be integrated in existing projects or build the foundation for new projects. It can be downloaded and it is open source, which gives the user complete control over all of its functionality. It is compatible with standard installations of clusters that provide GROMACS, biopython and scientific python.

ACKNOWLEDGMENTS

We thank the Helmholtz Association for funding our work within the "Impuls- und Vernetzungsfond" and the bwGRiD as part of the National Grid Initiative (NGI) for providing cluster resources for our simulations.

REFERENCES

Adcock, S. A. and J. A. McCammon (2006). "Molecular dynamics: Survey of methods for simulating the activity of proteins." Chem Rev **106**(5): 1589-1615.

Baxter, E. L., P. A. Jennings, et al. (2012). "Strand swapping regulates the iron-sulfur cluster in the diabetes drug target mitoNEET." Proceedings of the National Academy of Sciences of the United States of America **109**(6): 1955-1960.

Bryngelson, J. D., J. N. Onuchic, et al. (1995). "Funnels, Pathways, and the Energy Landscape of Protein-Folding - a Synthesis." Proteins-Structure Function and Genetics **21**(3): 167-195.

Cho, S. S., Y. Levy, et al. (2006). "P versus Q: Structural reaction coordinates capture protein folding on smooth landscapes." Proceedings of the National Academy of Sciences of the United States of America **103**(3): 586-591.

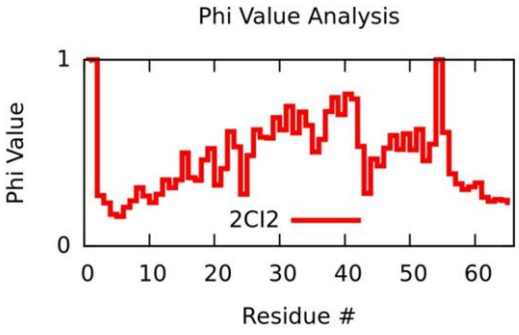


Figure 6: Phi value analysis for CI-2

Clementi, C., P. A. Jennings, et al. (2001). "Prediction of folding mechanism for circular-permuted proteins." Journal of Molecular Biology **311**(4): 879-890.

Clementi, C., H. Nymeyer, et al. (2000). "Topological and energetic factors: What determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins." Journal of Molecular Biology **298**(5): 937-953.

Cock, P. J., T. Antao, et al. (2009). "Biopython: freely available Python tools for computational molecular biology and bioinformatics." Bioinformatics **25**(11): 1422-1423.

Dago, A. E., A. Schug, et al. (2012). "Structural basis of histidine kinase autophosphorylation deduced by integrating genomics, molecular dynamics, and mutagenesis." Proceedings of the National Academy of Sciences of the United States of America **109**(26): E1733-1742.

Dill, K. A., S. B. Ozkan, et al. (2008). "The protein folding problem." Annu Rev Biophys **37**: 289-316.

Fersht, A. R. (1995). "Characterizing Transition-States in Protein-Folding - an Essential Step in the Puzzle." Current Opinion in Structural Biology **5**(1): 79-84.

Gambin, Y., A. Schug, et al. (2009). "Direct single-molecule observation of a protein living in two opposed native structures." Proc Natl Acad Sci U S A **106**(25): 10153-10158.

Gambin, Y., V. VanDelinder, et al. (2011). "Visualizing a one-way protein encounter complex by ultrafast single-molecule mixing." Nature methods **8**(3): 239-241.

Hess, B., D. Kutzner, et al. (2008). "GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Dynamics." J. Chem. Theory Comp. **4**: 435-447.

Hopf, T. A., L. J. Colwell, et al. (2012). "Three-dimensional structures of membrane proteins from genomic sequencing." Cell **149**(7): 1607-1621.

Jamros, M. A., L. C. Oliveira, et al. (2010). "Proteins at work: a combined small angle X-RAY scattering and theoretical determination of the multiple structures involved on the protein kinase functional landscape." The Journal of biological chemistry **285**(46): 36121-36128.

Klimov, D. K. and D. Thirumalai (1998). "Cooperativity in protein folding: from lattice models with sidechains to real proteins." Folding & Design **3**(2): 127-139.

Kussell, E., J. Shimada, et al. (2002). "A structure-based method for derivation of all-atom potentials for protein folding." Proceedings of the National Academy of Sciences of the United States of America **99**(8): 5343-5348.

Lammert, H., A. Schug, et al. (2009). "Robustness and generalization of structure-based models for protein folding and function." Prot: Struct, Funct and Bioinf **77**(4): 881-891.

Lange, O. F., N. A. Lakomek, et al. (2008). "Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution." Science **320**(5882): 1471-1475.

Lee, E. H., J. Hsin, et al. (2009). "Discovery through the computational microscope." Structure **17**(10): 1295-1306.

Lindorff-Larsen, K., S. Piana, et al. (2010). "Improved side-chain torsion potentials for the Amber ff99SB protein force field." Proteins **78**(8): 1950-1958.

Mickler, M., R. I. Dima, et al. (2007). "Revealing the bifurcation in the unfolding pathways of GFP by using single-molecule experiments and simulations." Proc Natl Acad Sci U S A **104**(51): 20268-20273.

Noel, J. K., P. C. Whitford, et al. (2012). "The Shadow Map: A General Contact Definition for Capturing the Dynamics of Biomolecular Folding and Function." The journal of physical chemistry. B.

Okazaki, K., N. Koga, et al. (2006). "Multiple-basin energy landscapes for large-amplitude conformational motions of proteins: Structure-based molecular dynamics simulations." Proc Natl Acad Sci U S A **103**(32): 11844-11849.

Oliveira, L. C., A. Schug, et al. (2008). "Geometrical features of the protein folding mechanism are a robust property of the energy landscape: a detailed investigation of several reduced models." J Phys Chem B **112**(19): 6131-6136.

Olivier, S., J. Prins, et al. (2007). Porting the GROMACS Molecular Dynamics Code to the Cell Processor. IEEE International Parallel and Distributed Processing Symposium, IPDPS.

Onuchic, J. N. and P. G. Wolynes (2004). "Theory of protein folding." Curr Opin Struct Bio **14**(1): 70-75.

Perez-Sanchez, H. and W. Wenzel (2011). "Optimization methods for virtual screening on novel computational architectures." Current computer-aided drug design **7**(1): 44-52.

Sadqi, M., D. Fushman, et al. (2006). "Atom-by-atom analysis of global downhill protein folding." Nature **442**(7100): 317-321.

Schug, A., T. Herges, et al. (2005). "Comparison of Stochastic optimization methods for all-atom folding of the Trp-cage protein." Chemphyschem **6**(12): 2640-2646.

Schug, A. and J. N. Onuchic (2010). "From protein folding to protein function and biomolecular binding by energy landscape theory." Curr Opin Pharmacol **10**(6): 709-714.

Schug, A., M. Weigt, et al. (2010). "Computational modeling of phosphotransfer complexes in two-component signaling." Methods in enzymology **471**: 43-58.

Schug, A., M. Weigt, et al. (2009). "High resolution complexes from integrating genomic information with molecular simulation." Proc Natl Acad Sci U S A **106**(52): 22124-22129.

Schug, A. and W. Wenzel (2004). "All-atom folding of the trp-cage protein with an adpative parallel tempering method." Europhysics Letters **67**(2): 307-313.

Schug, A., P. C. Whitford, et al. (2007). "Mutations as trapdoors to two competing native conformations of the Rop-dimer." Proc Natl Acad Sci USA **104**: 17674-17679.

Shaw, D. E., P. Maragakis, et al. (2010). "Atomic-level characterization of the structural dynamics of proteins." Science **330**(6002): 341-346.

Sulkowska, J. I., F. Morcos, et al. (2012). "Genomics-aided structure prediction." Proceedings of the National Academy of Sciences of the United States of America **109**(26): 10340-10345.

Thirumalai, D. and C. Hyeon (2005). "RNA and protein folding: common themes and variations." Biochemistry **44**(13): 4957-4970.

Thirumalai, D., D. K. Klimov, et al. (2002). Insights into specific problems in protein folding using simple concepts. Computational Methods for Protein Folding. **120:** 35-76.

Verma, A., A. Schug, et al. (2006). "Basin hopping simulations for all-atom protein folding." Journal of Chemical Physics **124**(4).

Verma, A. and W. Wenzel (2009). "A free-energy approach for all-atom protein simulation." Biophys J **96**(9): 3483-3494.

Weigt, M., R. A. White, et al. (2009). "Identification of direct residue contacts in protein-protein interaction by message passing." Proc Natl Acad Sci U S A **106**(1): 67-72.

Whitford, P. C., A. Ahmed, et al. (2011). "Excited states of ribosome translocation revealed through integrative molecular modeling." Proceedings of the National Academy of Sciences of the United States of America **108**(47): 18943-18948.

Whitford, P. C., J. K. Noel, et al. (2009). "An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields." Proteins-Structure Function and Bioinformatics **75**(2): 430-441.

Whitford, P. C., A. Schug, et al. (2009). "Nonlocal helix formation is key to understanding S-adenosylmethionine-1 riboswitch function." Biophys J **96**(2): L7-9.

MODELING OF RHEOLOGICAL BEHAVIOUR OF TOMATO SPREADS

Rosa, E., Peinado, I., Heredia, A. and Andrés*, A.

Institute of Food Engineering for Development
Universitat Politècnica de València
Camino de Vera s/n
46022 Valencia (Spain)
*aandres@tal.upv.es

ABSTRACT
The knowledge of the rheological properties of a material is very useful in the quality control of foods, equipment design and analysis of the structural changes produced during the processing/storage. In addition, it plays an important role in the consumer acceptance and can be used in the development of new products since allows to evaluate the effect of new ingredients/elaboration methods on its structure. The aim of this work was to model the rheological behaviour of tomato spreads elaborated under different processing variables (type of sugar, elaboration method and pectin percentage). All the products presented a non-Newtonian behaviour and the results could be fitted to Herschel-Bulkley model with a yield stress between 4.3 and 71.2 Pa. The statistical analysis showed higher consistency for the parameters obtained in the sucrose spreads compared to the products elaborated with isomaltulose. Finally, different equations are presented for predicting the rheological parameters as a function of the type of sugar and pectin percentage.

Keywords: rheology, isomaltulose, fruit spreads, osmotic dehydration.

1. INTRODUCTION

In recent years, changes in the lifestyle have increased the consumption of prepared food or ready-to-eat. However, consumers are aware that can demand products with a high nutritional and functional value and in some cases as close as possible to the fresh product as it is the case of fruit and vegetables.
Additionally, several epidemiological studies have pointed out the relation between a regular consumption of tomatoes and tomato based products and the prevention of cardiovascular diseases and determinate types of cancers (Clinton 1998; Giovannuci 1999; Miller *et al.* 2002).
In this sense, the development of new tomato products with high lycopene content and rich in other original nutrients such as ready-to-eat spreadable products could be interesting.
Fruit spread products present similar characteristics to fresh fruit but they are more stable than the fresh ones as the water activity and moisture of the products are reduced (Peinado *et al.* 2009, Rosa *et al.* 2009).
Osmotic dehydration can be used for water removing and sugar addition in concentration processes, avoiding the heating process since water removal takes place by osmosis. Besides, osmotic process can also be applied by using dry sugar instead of sugar solutions, similar to in dry salting process commonly applied to meat and fish products. In these processes, a concentrated solution, rich in volatile compounds, vitamins and water soluble minerals from the fruit itself is generated due to water out-flow and its volume is considerably lower than the volume used in wet osmotic dehydration processes.
Furthermore, sucrose could be replaced by other sugars with low glycemic index, such as isomaltulose, in order to obtain healthier products.
On the other hand, the syneresis process (separation into pulp and serum) is a defect very usual in conventional tomato-based products which could be reduced by increasing the consistency of the product; a high-consistency tomato juice has almost no syneresis (Porretta *et al*. 1995). In this way, spreadable tomato products would have a lower risk of suffering syneresis during the storage because they present a higher consistency than natural tomato puree by addition of pectin and increasing of soluble solids in their formulation.
Rheological measurements of these innovative products might be very useful to evaluate the effect of the different processing variables on its structural changes.
The aim of this study was to model the rheological flow behaviour of 20 Brix tomato spreadable products formulated with different processing variables: type of sugar (sucrose or isomaltulose), elaboration method (dry or wet osmotic dehydration, and proportion of fruit), and pectin percentage (0.5, 1, 1.5 %).

2. MATERIAL AND METHODS

2.1. Raw Material

Pear tomatoes (*Lycopersicon esculentum Mill*.) acquired in a local supermarket were manually selected to homogenise the samples in terms of shape, colour and ripeness, avoiding those with physical damage. Samples

were cleaned with chlorinated water, peeled and cut into cubes of approximately 1 cm^3.

2.2. Methodology

Equilibrium stage: two different osmotic dehydration methods at 25 ºC were used, Wet Osmotic Dehydration (WOD), in which the tomato cubes were immersed in a hypertonic osmotic solution (40 Brix) and Dry Osmotic Dehydration (DOD), in which the tomato cubes were directly covered with the solid osmotic agent, similar to the dry salting process commonly applied to meat and fish products. The ratio tomato:osmotic agent was calculated by mass balance to reach an equilibrium concentration of 20 Brix. In this stage either sucrose or isomaltulose were used like osmotic agents.

Product formulation: the ingredients in the spread formulations were dehydrated tomato, final osmotic solution, olive oil 2 %, salt 0.4 %, apple pectin (0.5, 1, 1.5 %) as a gelling agent and potassium sorbate (500 ppm) as a preservative (% refers to the final product). According to the different proportions of dehydrated tomato:final osmotic solution and depending on the type of osmotic dehydration (W or DOD), three elaboration methods were selected:

- W: WOD and formulation with a dehydrated tomato:final osmotic solution ratio of 70:30.
- D1: DOD and formulation with a dehydrated tomato:generated osmotic solution ratio comparable to the proportion achieved during the equilibrium stage (≈50:50).
- D2: DOD and formulation with a dehydrated tomato:generated osmotic solution ratio of 70:30.

Finally, 18 tomato spreads were formulated depending on the different processing variables: kind of sugar, elaboration method and pectin percentage (table 1).
Products were homogenized with a mixer for 3 minutes. Then they were stored for 24 hours to allow correct gel stabilization before to perform the analysis.

2.3. Rheological Analysis

Rheological properties were obtained with a controlled stress rheometer (RheoStress 1, Haake), at 25 ºC. All measurements were carried out in duplicate with plate-plate geometry and a 2.0 mm gap (Sato and Cunha 2009).
Steady state measurements were carried out with shear rate ranging from 0 to 300 s-1, in 3 sweeps (up, down and up-cycles), in order to eliminate the thixotropy. The data obtained in the third sweep were fitted to rheological models using the software Rheowin Data Manager (Haake) to find the best suitable model.

Table 1: Selected processing variables to obtain the different 20 Brix spreadable tomato products.

	Osmotic Agent	Elaboration	% Pectin
1	Sucrose	W	0.5
2			1
3			1.5
4		D1	0.5
5			1
6			1.5
7		D2	0.5
8			1
9			1.5
10	Isomaltulose	W	0.5
11			1
12			1.5
13		D1	0.5
14			1
15			1.5
16		D2	0.5
17			1
18			1.5

2.4. Statistical Analysis

Statgraphics Centurion was used to perform the statistical analyses. Analyses of variance (multifactor ANOVA) were carried out to estimate the significant effect of the processing variables (type of sugar, elaboration method and pectin percentage) on the final spreadable product.

3. RESULTS AND DISCUSSION

3.1. Rheological Modeling

After removing the thixotropy, the flow curves of the tomato spreads showed a shear thinning behaviour with yield stress. Therefore, Herschel–Bulkley model (equation 1) was the model that best fitted the experimental data.

$$\tau=\tau_0+k\cdot\gamma^n \quad (1)$$

Where, τ is the shear stress (Pa), τ_0 the yield stress (Pa), γ the shear rate (s^{-1}), k the consistency index ($Pa\cdot s^n$) and n the flow behaviour index.
This rheological behaviour is one of the most common in fruit and vegetables products like purees, jams or spreadable products (Sharma *et al.* 1996; Maceiras, Álvarez and Cancela 2007; Tonon *et al*. 2009; Sato and Cunha 2009; Peinado 2011).
Table 2 shows the rheological parameters obtained for the 18 tomato spread products with values of R^2 between 0.839 and 0.999. The differences observed between samples could be attributed to the different ratios fruit-osmotic solution as well as to the pectin percentages.

Table 2: Rheological parameters of Herschel–Bulkley model fitted to data of 20 Brix tomato spreadable products: τ_0 is the yield stress (Pa), k the consistency index (Pa·s^n) and n the flow behaviour index.

	τ_0	k	n
1	8.3 (0.6)	4.38 (0.05)	0.3757 (0.0007)
2	29.6 (1.2)	13.2 (0.3)	0.354 (0.005)
3	54 (15)	26 (5)	0.306 (0.009)
4	4.5 (1.9)	2.308 (1.015)	0.38 (0.03)
5	15.5 (1.5)	4.0 (0.8)	0.39 (0.04)
6	53.030 (-)	19.070 (-)	0.3500 (-)
7	7.2 (0.2)	2 (2)	0.3169 (-)
8	23 (4)	3.1 (0.3)	0.449 (0.004)
9	64.970 (-)	20.980 (-)	0.3575 (-)
10	8.0 (0.6)	1.784 (1.119)	0.4263 (-)
11	32.4 (0.5)	14 (3)	0.34 (0.05)
12	60 (3)	21.9 (0.3)	0.3356 (0.0109)
13	4.263 (-)	0.107 (-)	0.6900 (-)
14	20 (3)	7.1 (0.6)	0.395 (0.008)
15	32 (11)	9 (3)	0.373 (0.009)
16	13.8 (0.7)	1.6 (0.7)	0.53 (0.05)
17	25.830 (-)	4.208 (-)	0.3364 (-)
18	71.22 (1.03)	20.17 (0.08)	0.3277 (0.0007)

To evaluate the influence of the processing variables on the rheological parameters obtained, an ANOVA factorial analysis was performed. The table 3 shows the F-ratio values and the significance levels (p-value) of the effect of the different variables obtained in the statistical analysis.

Table 3: F-ratio values and significance levels (p-value) of Herschel-Bulkley model parameters obtained in the ANOVA factorial analysis.

Variables	F-ratio	P-value
τ_0		
A: Type of sugar	0.15	0.6983
B: Elaboration	**17.08**	**0.0000**
C: % Pectin	**220.04**	**0.0000**
AxB	2.78	0.0839
AxC	1.11	0.3462
BxC	**4.78**	**0.0063**
k		
A: Type of sugar	**6.09**	**0.0218**
B: Elaboration	**35.10**	**0.0000**
C: % Pectin	**230.47**	**0.0000**
AxB	1.67	0.2106
AxC	**8.64**	**0.0017**
BxC	**9.52**	**0.0001**
n		
A: Type of sugar	**15.27**	**0.0008**
B: Elaboration	**9.72**	**0.0009**
C: % Pectin	**23.01**	**0.0000**
AxB	**4.91**	**0.0173**
AxC	**27.54**	**0.0000**
BxC	2.53	0.0699

The statistical results revealed than the variable with a higher influence on the yield stress (τ_0) and the consistency index (k) was the pectin percentage (F-ratio value much higher than the rest of variables). However, in the case of flow behaviour index (n), both the kind of sugar and the pectin percentage presented an important effect.

These results were similar to those obtained for strawberry spread products elaborated with the same sugars (Peinado 2011).

3.2. Prediction Equations for the Rheological Parameters of Herschel-Bulkley Model

The equations for predicting the rheological parameters of Herschel-Bulkley model were obtained as a function of the type of sugar and pectin percentage, since the statistical results showed a higher influence on these parameters (table 4).

It can be noticed that in the case of sucrose spreads, the flow behaviour index (n) is constant, regardless of the percentage of pectin used.

Table 4: Prediction equations obtained for the parameters of Herschel-Bulkley model of spreadable tomato products depending on the kind of sugar and the pectin percentage (P).

	Prediction Equations	R^2
Sucrose		
τ_0	$\tau_0=24.7795P^{1.9457}$	0.9956
k	$k=8.9967P^{1.7627}$	0.9377
n	$n=0.36\pm0.04$	-
Isomaltulose		
τ_0	$\tau_0=27.0999P^{1.6593}$	0.9990
k	$k=7.0306P^{2.4840}$	0.9844
n	$n=0.3908P^{-0.4411}$	0.9119

It can be observed than in the case of yield stress parameter (τ_0), the proposed equations for both sugars (sucrose and isomaltulose) are very similar. This fact agrees with the previous statistical results, where no significant influence was found in the yield stress (τ_0) depending on the type of sugar. However, the equations of the consistency index (k) for both sugars show a higher increase of 'k' with the pectin addition in isomaltulose tomato spreads as compared to sucrose products. This fact could be related to the lower solubility of the isomaltulose, which implies more free water in the food matrix and the pectin addition would be more effective.

The suitability of the predicted models for describing the type of sugar and the pectin percentage-dependent behaviour of the tomato spreads was checked by comparing the experimental data with the theoretical values (figures 1 and 2).

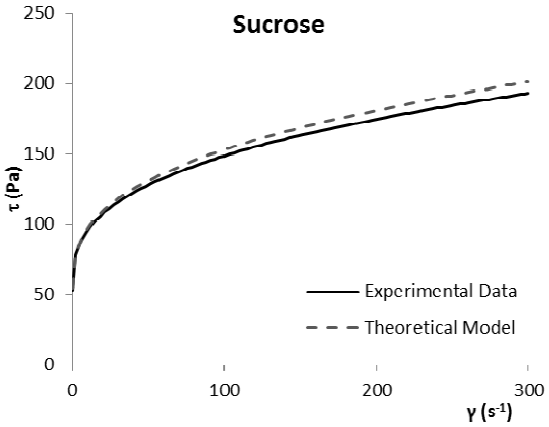


Figure 1: Example of the correlation between the experimental data and the theoretical curve for 20 Brix sucrose tomato spreads.

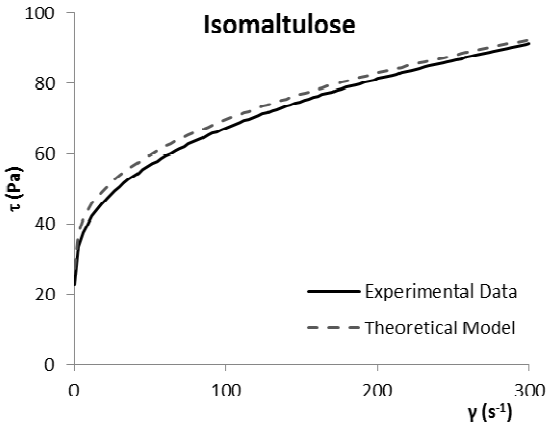


Figure 2: Example of the correlation between the experimental data and the theoretical curve for 20 Brix isomaltulose tomato spreads.

It can be observed than the theoretical values were in accordance with experimental data obtained for both sugars (sucrose and isomaltulose). Therefore, the proposed equations appear to be suitable to describe and predict the rheological parameters of the Herschel-Bulkley model for tomato spreads within the range of the studied processing variables.

4. CONCLUSIONS

Sugar and pectin percentage variables were seen to have a greater influence than the elaboration method on the rheological behaviour of spreadable tomato products. Since this type of product presents a shear thinning behaviour with yield stress, the Herschel-Bulkley model was found to be the one that best fitted the experimental data.

The pectin percentage had a very strong influence on the yield stress values (τ_0), but they did not exhibit significant differences for the different sugars used (sucrose and isomaltulose).

The values of the consistency index (k) and the flow behaviour index (n) were influenced by all the processing variables (type of sugar, elaboration method and pectin percentage). The highest values for consistency and therefore the lowest for fluidity were obtained in the tomato spreads elaborated with the highest proportion of fruit (W and D2) and percentage of pectin.

The proposed mathematical equations can be used to estimate the rheological parameters in the Herschel-Bulkley model of 20 Brix spreadable tomato products taking into account the type of sugar and the pectin percentage used in their elaboration.

These results could be very useful for designing and developing new tomato-based products, though subsequent studies would be necessary to increase the range of pectin percentage or to introduce new ingredients into the formulation.

ACKNOWLEDGMENTS

E. R. would like to thank the financial support from JAE-Predoc program of CSIC. Authors would like to thank Ministry of Science and Education's General Directorate of Research (AGL2008-01745/ALI) for the financial support given to this research.

REFERENCES

Clinton, S. K., 1998. Lycopene: Chemistry, biology, and implications for human health and disease. *Nutrition Reviews*, 56, 35–51.

Giovannucci, E., 1999. Tomatoes, tomato based products, lycopene, and cancer: review of epidemiologic literature. *Journal of the National Cancer Institute*, 91, 317-331.

Maceiras, R., Álvarez, E. and Cancela, M.A., 2007. Rheological properties of fruit purees: Effect of cooking. *Journal of Food Engineering*, 80 (3),763-769.

Miller E.C., Giovannucci E., Erdman J.W. Jr, Bahnson R., Schwartz S.J. and Clinton S.K., 2002. Tomato products, lycopene, and prostate cancer risk. *The Urologic Clinics of North America*, 29 (1), 83-93.

Peinado, I., Rosa, E., Heredia, A. and Andrés, A., 2009. Influence of Dry and Wet Osmotic Dehydration on Colour and Texture of a Spread Strawberry Product. 'New Challenges in Food Preservation'. European Federation of Food Science and Technology (EFFoST). (11-13/11/09) Budapest, Hungary.

Peinado, I., 2011. *Estudio de la utilización de isomaltulosa en el desarrollo de productos untables de tomate de bajo índice glicémico*. Thesis (PhD). Universidad Politécnica de Valencia.

Porretta, S., Birzi, A., Ghizzoni, C. and Vicini, E., 1995. Effects of ultra-high hydrostatic pressure treatments on the quality of tomato juice. Food Chemistry, 52, 35 –41.

Rosa, E., Peinado, I., Heredia, A. and Andrés, A., 2009. New Processing Methods and Healthier Sugars in

Strawberry Spread Fruit Manufacturing. 'New Challenges in Food Preservation'. European Federation of Food Science and Technology (EFFoST). (11-13/11/09) Budapest, Hungary.

Sato, A.C. and Cunha, R.L., 2009. Effect of particle size on rheological properties of jaboticaba pulp. *Journal of Food Engineering*, 91, 566–570.

Sharma, S. K., Le Maguer, M., Liptay A. and Poysa, V., 1996. Effect of composition on the rheological properties of tomato thin pulp. *Food Research International*, 29 (2), 115-117.

Tonon, R.V., Alexandre, D., Hubinger, M.D. and Cunha, R.L., 2009. Steady and dynamic shear rheological properties of açai pulp (Euterpe oleraceae Mart.). *Journal of Food Engineering*, 92, 425–431.

AUTHORS' BIOGRAPHIES

A. Andrés is a Full Professor in the *Universitat Politècnica de València* and the Director of the Institute of Food Engineering for Development in Valencia (Spain). **A. Heredia** is a Lecturer in the *Universitat Politècnica de València* and a Researcher member of the Institute of Food Engineering for Development in Valencia (Spain). **I. Peinado** acquired her PhD. in Food Technology in the *Universitat Politècnica de València* in 2011 and at present is a carrying out a post-doctoral fellowship in Northumbria University (United Kingdom). **E. Rosa** has a Bachelor's Degree in Food Science and Technology and currently is completing her PhD. in the *Universitat Politècnica de València* (Spain).

SIMULATION OF DILUTE-SOLUTION PROPERTIES OF BIOLOGICAL MACROMOLECULES WITH THE AID OF HIGH-PERFORMANCE COMPUTING

R. Rodríguez Schmidt[(a)], D. Amorós Cerdán[(b)], J.G. Hernández Cifre[(c)], F.G. Díaz Baños[(d)], J. García de la Torre[(e)]

[(a,b,c,d,e)] Dep. Química Física, Universidad de Murcia, 30100 Murcia, Spain

[(a)]ricrogz@um.es, [(b)]damoros@um.es, [(c)]jghc@um.es, [(d)]fgb@um.es, [(e)]jgt@um.es

ABSTRACT

The determination of dilute solution properties of macromolecules (hydrodynamic coefficients, radius of gyration and scattering-related properties, NMR and viscoelastic relaxation, etc) is of interest to characterize their conformation (size and shape) and dynamics in usual working environment (e.g. physiological conditions in case of biomacromolecules). Over the years, the Polymer Group at the University of Murcia has made computational developments intended for the prediction of dilute solution properties of synthetic polymers and biological macromolecules. Most of these developments are of public domain (see our web site http://leonardo.inf.um.es/macromol).

In this work, we present some improvements in our methodology aimed to achieve high-performance computing. The strategy is based on using parallelized versions of the LAPACK and similar mathematical libraries, and implementing in-house written codes. We show some results obtained by applying that methodology to some macromolecular models.

Keywords: bead model, biomacromolecule, conformation, hydrodynamic, parallel computing

1. INTRODUCTION

The calculation of conformational and hydrodynamic properties (like radius of gyration, sedimentation and diffusion coefficients, or intrinsic viscosities) of rigid and flexible macromolecules (like globular and denatured proteins respectively) is frequently based on constructing bead models that represent the size and shape of the particle (see Fig. 1) and then apply some fundamental hydrodynamic equations for the sphere to get the global particle properties. Those calculations involve linear-algebra problems, essentially the solutions of sets of linear (hydrodynamic interaction) equations (Carrasco and García de la Torre 1999), in which the number of unknowns is proportional to the number of elements in the model. This number has to be sufficiently large as to describe the complex structural details typical of biological macromolecules, particularly when modelling is done at atomic level (García de la Torre, Huertas, and Carrasco 2000), which quite often entails a large computation time.

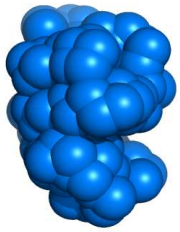


Figure 1: Example of bead model for lysozyme

The increasing number of cores in a processor and the availability of computer clusters, open the possibility of parallelizing some stages of the calculus and speeding up the global simulation process by running several simulations simultaneously. That is achieved by dividing the global problem into independent pieces that can be processed at the same time and then collect the results. Thus, an update in the software developed by our group is in order to embody some of those features and improve its efficiency.

We present in this work some strategies leading to easily implement parallelization with minor changes in our codes, mainly focused to run our programs under multi-core processor computers. On the one hand, we have replaced some of our homemade functions and subroutines by some others included in public domain mathematical libraries like LAPACK that are optimized to carry out parallel calculation if required. On the other hand we have built a set of ancillary tools intended to run a number of generations (cases) of some model in several cores simultaneously.

2. METHODOLOGY FOR RIGID PARTICLES

Conformational and hydrodynamic properties of rigid structures like many globular proteins can be calculated using our HYDRO programs suite (Carrasco and García de la Torre 1999). Nowadays, there is a big amount of accessible experimental information about the atomic structure of proteins, nucleic acids and many other biopolymers, which can be used to model those molecules and subsequently calculate solution properties. Our program HYDROPRO (García de la

Torre, Huertas, and Carrasco 2000; Ortega, Amorós, and García de la Torre 2011) is able to construct bead models based upon the atomic-level information contained in the Protein Data Bank that is a public domain data base for proteins and other related complex structures (www.rcsb.org/pdb/home/home.do).

The program places a sphere per atom (except hydrogen) with a radius that includes the hydration layer. Then, and after some model simplifications, it calculates basic solution properties of the structure: radius of gyration, diffusion and sedimentation coefficients, intrinsic viscosity, and relaxation times. Those calculations involve time consuming matrix and vector operations that can be parallelized by using public domain mathematical libraries. In particular, the most time consuming mathematical operation is the inversion of the hydrodynamic interaction tensor, the dimensions of which, for a model made of N beads, is 3Nx3N. Because the CPU time of such an operation scales as N^3, the calculation becomes slow quite soon.

Other aspect to take into account is the memory required to store such a big tensor. Because of the symmetry of the hydrodynamic tensor, it is possible to store just the elements in the upper or lower triangle of the matrix (what is called a packed matrix). Sometimes that strategy slows down the calculation and it is preferred to work with the whole square matrix. The program HYDROPRO has been rewritten to improve both memory management and calculation speed. For that purpose, subroutines to perform LU and Cholesky matrix factorization included in the LAPACK mathematical library (www.netlib.org) have been employed. On the other hand, bead models of complex structures like that in Fig. 1 for the protein lysozyme (so-called shell models) were generated starting from atomic-level structures taken from the Protein Data Bank.

Fig. 2 is a histogram showing the CPU time consuming by different versions of HYDROPRO in calculating a set of solution properties when running on a processor Intel Xeon x5660 6 cores (x2 CPU).

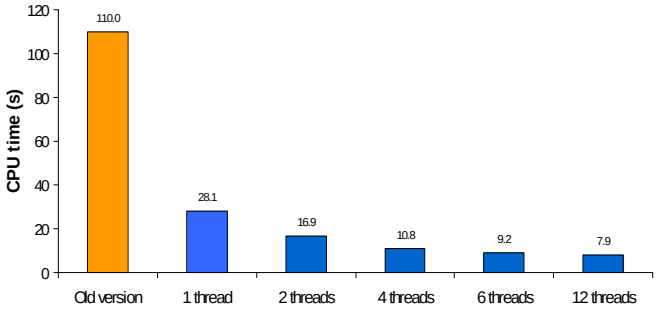


Figure 2: Gain in CPU time due to parallelization

The first bar on the left, corresponding to our old version of HYDROPRO, is exceedingly large in comparison with the other bars that measure the CPU time consumed by the parallel version of HYDROPRO that includes the possibility of using several calculation threads at a time. As appreciated, we can diminish the duration of the process more than ten times.

Experimental and calculated values for the radius of gyration, Rg, and the sedimentation coefficient, s, of several complex structures are compared in Table 1. As observed, the agreement is quite good (as it already occurred with our old version).

Table 1: Comparison of solution properties of complex structures calculated by HYDROPRO with their experimental values: [a]Tomonao et al, Biophys J 2008, 94, 1392-1402; [b]Behlke et al, Biochemistry 1997, 36, 5149-5156; [c]Armstrong et al, Biophys J 2004, 87, 4259-4270; [d]Hill et al, J Mol Biol 1969, 44, 263-277.

Structure	**Rg / Å**		**sx10^{13}/ s**	
	Calc.	Exp.	Calc.	Exp.
Chaperone GroEL	66.1	67.0[a]	21.5	22.13[b]
IgM antibody	127	121[c]	18.4	17.5[c]
Ribosme 30S	68.3	69[d]	36.9	31.8[d]
Ribosme 50S	74.6	77[d]	53.6	50.2[d]
Ribosme 70S	86.0	91.5[d]	69.3	70.5[d]

3. METHODOLOGY FOR FLEXIBLE CHAINS

The skeleton of the long and (usually) linear chains that constitutes biological macromolecules consists of chemical bonds arranged in such a manner that bond lengths and bond angles are nearly constant, but there is an important degree of freedom in the internal rotation around all or some of the bonds in the chain. The consequence of such conformational variability about each of so many bonds along the chain skeleton is that it could adopt multiple conformations like it is the case of denatured proteins. With a dynamic point of view, a single macromolecule is continuously changing its own conformation. Thus macromolecular chains are, in principle, essentially flexible entities that can be represented by bead and connector models like that in Fig. 3.

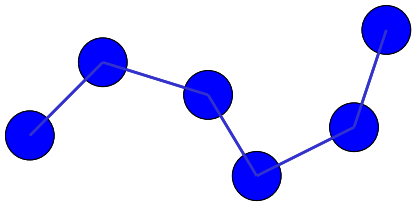


Figure 3: Example of bead and connector model

The theoretical foundation for the description of flexible chain macromolecules has its origin in the pioneering works of Kirkwood and Zimm (Kirwood and Riseman 1948; Zimm 1956).

As in the bead models used for rigid-body modeling, in the mechanic model of flexible particles, the parts composing the particle are represented by spherical beads, which is a convenient representation for hydrodynamic calculations. Instead of having a unique shape, the beads in the array are linked by a series of internal interactions that determine the conformational variability of the ensemble.

Firstly, there must be some connectors between (neighbor) beads, like the bonds in a chemical molecule. They behave as springs of some degree of flexibility, and are the primary interactions in the mechanic model, which -- regardless of the instantaneous conformation -- determine its topology (linear, branched, etc). Additionally, there can be other short-term interactions, mainly bending interactions between two neighbor connectors, and perhaps torsional or internal-rotation restrictions. Also, it may be present long-term interactions between non-connected beads (so-called excluded volume interactions) which are properly represented by several potential expressions depending on the features of the system (hard sphere potential, Lennard-Jones potential…)

The conformational statistics of flexible macromolecules can be simulated by a standard Monte Carlo (MC) procedure. However, the full, rigorous simulation of the macromolecular dynamic behavior has to be done by means of Brownian dynamics (BD).

3.1. Monte Carlo

Generically speaking, the MC simulation methods are intended to generate possible states of a system, which in our case are conformations of a flexible macromolecule, in order to obtain observable properties that correspond to averages over all the accessible conformations or properties. The various states have different probabilities, which are proportional to the Boltzmann factor, exp(-V/kT), associated to the potential energy of the system in such conformation, V (being k the Boltzmann constant and T the absolute temperature). If the possible conformations are generated in an absolutely random manner, this factor should be used as a statistical weight in the evaluation of the averages.

A usual choice to carry out a MC simulation is the importance-sampling procedure which is quite simple. From a previous conformation of the chain, a new one is generated making small random displacements of the beads. The potential energy of the new conformation, V, is evaluated as a sum of the several contributions from the various kinds of intramolecular or external interactions that the mechanic model may include. The new potential is compared to that of the previous conformation, Vprev. The new conformation is accepted if V < Vprev. Otherwise, if exp[(Vprev-V)/kT] > u, where u is a random number with uniform distribution in (0,1) generated each time that this decision is to be made, then the new conformation is accepted. If exp[(Vprev-V)/kT] < u, the conformation is rejected and the resulting conformation after the MC step is a copy of the previous one.

For MC simulations of bead-and-connector models, we have developed the public domain MONTEHYDRO program (García de la Torre, Ortega, Pérez Sánchez, and Hernández Cifre 2005) that implements an importance-sampling Monte Carlo simulation coupled to rigid-body hydrodynamics which is based on the procedures of the above mentioned HYDRO programs suite. Thus, the hydrodynamic coefficients are calculated using the "Monte-Carlo rigid-body" approach (Zimm 1980): the macromolecular conformations generated by MC are considered as instantaneously rigid and the rigid body hydrodynamic equations are applied to each one. Then a conformational average is performed in order to obtain global equilibrium properties.

Fig. 4 shows the excellent agreement obtained between experimental and simulation data when comparing the radius of gyration of a set of flexible proteins with different number of residues, N, (Amorós 2012). Simulations were carried out using the MONTEHYDRO program including parallelization. The model included excluded volume interactions represented by a hard sphere potential where the hard sphere radius was r_{HS}=2.25 Å.

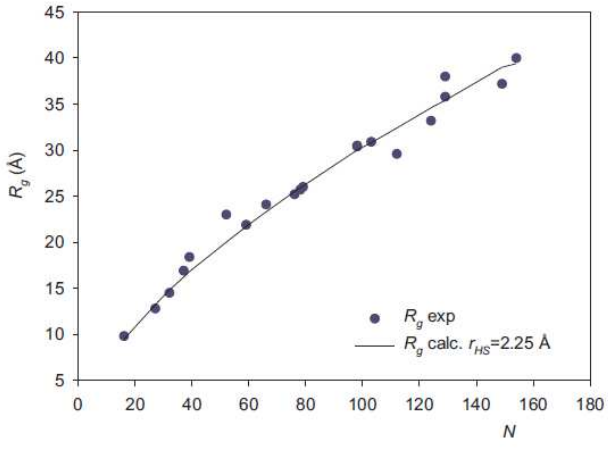


Figure 4: Rg of proteins: experimental (dots) and calculated by MONTEHYDRO (line)

3.2. Brownian dynamics

To study dynamic aspects of flexible macromolecules in solution, such as relaxation processes and non-equilibrium behavior, it is necessary to solve the

equation of motion that governs the macromolecular dynamics. This can be done by using molecular dynamics (MD) or Brownian dynamics (BD). MD is not adequate for long time and size scales. BD is a numerical technique to solve the stochastic equation of motion that arises from considering the solvent as a continuum, thus eliminating the solvent degrees of freedom. In other words, BD simulations describe the Brownian motion of a collective of frictional elements, beads in our model, which can interact with each other through different potentials.

An essential aspect in the BD simulation is the inclusion of the so-called hydrodynamic interaction (HI) effect, which determines the solvent-mediated influence of the motion of every element of the model on the others. A first order algorithm to solve the stochastic equation of motion and performed Brownian dynamics simulations including HI is that of Ermak and MacCammon (Ermak and McCammon 1977).

We have developed a BD simulation scheme that enables for the calculation of solution properties of flexible macromolecules with arbitrary complexity. Our procedures take into account fluctuating (non-preaveraged) hydrodynamic interaction as well as the possibility of including different types of intra-molecular potentials to represent excluded volume conditions and electrostatic interactions. That computational scheme is implemented in a suite of public domain programs, named SIMUFLEX. The suite consists mainly of two programs: BROWFLEX and ANAFLEX. The program BROWFLEX generates a Brownian trajectory of a flexible bead-and-connector model with arbitrary connectivity, and the program ANAFLEX analyses that trajectory to obtain several steady and time-dependent macromolecular quantities (García de la Torre, Hernández Cifre, Ortega, Rodríguez Schmidt, Fernandes, Pérez Sánchez, and Pamies 2009).

However, BD algorithms with fluctuating HI demands long simulation (CPU) time, which increases dramatically with the number of elements forming the chain model, N. The main reason is that the BD-HI methodology requires the calculation of the square root of the 3Nx3N HI tensor, a very time-consuming operation carried out every time that tensor is updated during the simulation. An exact mathematical procedure to get the square root of the symmetric matrix is the already mentioned Cholesky factorization. Some approximations have been proposed to speed up BD-HI simulations. Those mathematical approaches are able to increase enormously the efficiency of BD-HI simulations (Rodríguez Schmidt, Hernández Cifre, and García de la Torre 2011), moreover when they are used in combination with parallelization strategies for the matrix and vector operations.

It is worth to see how we can gain in CPU time due to the mathematical approaches themselves without using parallel computing. Fig. 5 shows the CPU time needed for the BD-HI simulations of linear Gaussian chains with excluded volume with increasing number of beads. For that purpose, we have measured the duration of 10^6 Brownian steps on a 3.0 GHz Intel Xeon Quad-Core L5450. As appreciated, the use of Geyer (Geyer and Winter 2009) and Kröger (Kröger, Alba Pérez, Laso, and Öttinger 2000) algorithms, when applicable, entails a fabulous gaining in CPU time respect to those of Jendrejack (Jendrejack, Graham, and de Pablo 2000) and Cholesky (Ermak and McCammon 1977). (Note: algorithms are named here after the first author of the paper where they were presented, except that of Cholesky that refers to algorithms that use the rigorous Cholesky factorization).

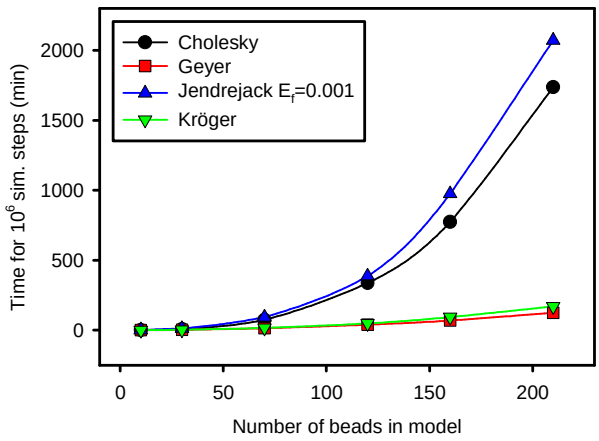


Figure 5: Runtime behavior of different BD-HI algorithms

4. DISTRIBUTION OF A GLOBAL TASK

Nowadays, even processors of modest PCs are multi-core (e.g. Dual, or Quad) and it is common to work with multiprocessor units, or even clusters of many units. Then, if the kind of problem under simulation study meets some requirements, one can make use of these resources in order to split a large task, like the generation of a macromolecular trajectory, into many shorter independent trajectories that will run in parallel using all of the available processors. The final results will be produced after a proper combination of the partial results obtained from those short independent trajectories. Obviously, that simulation strategy reduces the time needed to solve the global task approximately so many times as the number of available processors.

With the aim of dividing a global task into smaller independent ones, we are designing the programs suite MULTISIMUFLEX, a set of ancillary tools to be used along with MONTEHYDRO and SIMUFLEX. Thus, one can take advantage of multi-core computers and distribute through the available computing cores the task of generating conformations of bead-and-connector models either via Monte Carlo or via Brownian dynamics. In order to generate independent trajectories we resort to two procedures: a) the use of different initial conformations for each sub-simulation, b) the use of different sequence of random numbers for each sub-simulation. The MULTISIMUFLEX suite is not available in our web site yet.

We have carried out several tests to verify the right work of MULTISIMUFLEX. Fig. 6 is a sketch of a 'homemade' tetra-block A-B-C-B copolymer that represents a chimerical protein. It consists of a globular domain plus a hanging chain having an internal helix. We ran a large Brownian dynamics simulation by dividing it into independent simulations. Thus, the CPU time needed for the whole dynamics and then getting accurate values of the steady-state properties is reduced by an order of magnitude. Table 2 shows the values of the analyzed properties.

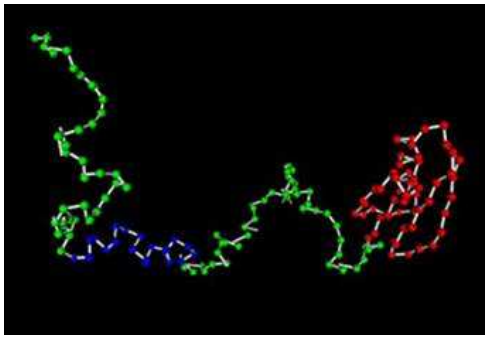


Figure 6: Bead-connector model of chimerical protein

Table 2: properties obtained for the chimerical protein after running a simulation using MULTISIMUFLEX.

Radius gyration (cm)	Diffusion coefficient (cm^2/s)	Intrinsic viscosity (cm^3/g)
2.73×10^{-7}	9.622×10^{-7}	4.726

5. CONCLUSIONS

It is clear that the use of multiprocessor units speed up the calculation of solution macromolecular properties without loss of precision. We have presented here some simple parallelization strategies used to improve the efficiency of our public-domain programs available at our web site http://leonardo.inf.um.es/macromol.

One strategy is the substitution of parts of our code by equivalent parallelized functions obtained from public-domain mathematical libraries like LAPACK. Thus, a faster version of our program HYDROPRO to calculate solution properties of rigid structures is now available. Other possibility is distributing efficiently the calculation among every available processor. In order to help in the task of splitting a large simulation into shorter ones that will run in parallel, we have designed the suite MULTISIMUFLEX. That program must be used in combination with our MONTEHYDRO and SIMUFLEX programs which are used to run Monte Carlo and Brownian dynamics simulations of flexible structures and then get their solution properties.

The above commented parallelization strategies along with the improvement of mathematical aspects of the simulation algorithms were tested by applying them to several bead models of proteins. It was verified that those strategies are able to yield excellent results for the solution properties and diminish the computational time more than one order of magnitude, thus increasing enormously the performance of our algorithms.

ACKNOWLEDGMENTS

This work was performed within a Grupo de Excelencia de la Región de Murcia (grant 04531/GERM/06). Support also provided by grant CTQ-2009-06831 from Ministerio de Educación y Ciencia, including FEDER funds.

REFERENCES

Amorós, D., 2012, *Metodologías para la predicción de propiedades conformacionales y dinámicas de macromoléculas biológicas en disolución*. PhD. Thesis, University of Murcia.

Carrasco, B., García de la Torre, J., 1999, Hydrodynamic Properties of Rigid Particles. Comparison of Different Modelling and Computational Procedures. *Biophysical Journal* 76, 3044–3057.

Ermak, D..L., McCammon, J.A., 1978, Brownian dynamics with hydrodynamic interactions. *Journal of Chemical Physics* 69, 1352–1360.

García de la Torre, J., Huertas, M.L., Carrasco, B., 2000, Calculation of Hydrodynamic Properties of Globular Proteins from their Atomic-Level Structure. *Biophysical Journal* 78, 719–730.

García de la Torre, J., Ortega, A., Pérez Sánchez, H.E., Hernández Cifre, J.G., 2005, MULTIHYDRO and MONTEHYDRO: Conformational search and Monte Carlo calculation of solution properties of rigid and flexible macromolecular models. *Biophysical Chemistry* 116, 12–128.

García de la Torre, J., Hernández Cifre, J.G., Ortega, A., Rodríguez Schmidt, R., Fernandes, M.X., Pérez Sánchez, H.E., Pamies, R., 2009, SIMUFLEX: Algorithms and tools for simulation of the conformation and dynamics of flexible molecules and nanoparticles in dilute solution. *Journal of Chemical Theory and Computation* 5, 2606-2618.

Geyer, T., Winter, U., 2009, A O(N^2) approximation for hydrodynamic interactions in Brownian dynamics simulations. *Journal of Chemical Physics* 130, 1149051 (8 pages).

Jendrejack, R.M., Graham, M.D., de Pablo, J.J.., 2000, Hydrodynamic interactions in long chain polymers: Application of the Chebyshev polynomial approximation in stochastic simulations. *Journal of Chemical Physics* 113, 2894-2900.

Kirkwood, J.G.., Riseman, J., 1948, The intrinsic viscosities and diffusion constants of flexible macromolecules in solution. *Journal of Chemical Physics* 16, 565–573.

Kröger, M., Alba Pérez, A., Laso, M., Öttikger, H.C., 2000, Variance reduced Brownian simulation of a bead-spring chain under steady shear flow considering hydrodynamic interaction effects. *Journal of Chemical Physics* 113, 4767-4773.

Ortega, A., Amorós, D., García de la Torre, J., 2011, Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophysical Journal* 101, 892–898.

Rodríguez Schmidt, R., Hernández Cifre, J.G., García de la Torre, J., 2011, Comparison of Brownian dynamics algorithms with hydrodynamic interaction. *Journal of Chemical Physics* 135, 084116 (10 pages).

Zimm, B.H., 1956, Dynamics of polymer molecules in dilute solution: viscoelasticity, flow birefringence and dielectric loss. *Journal of Chemical Physics* 24, 269–277.

Zimm, B.H., 1980, Chain molecule hydrodynamics by the Monte-Carlo method and the validity of the Kirkwood-Riseman approximati. M*acromolecules* 13, 592–602.

AUGMENTED GALLERY GUIDE

Zuzana Haladova[(a)], Csaba Boylos[(b)]

[(a)] Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava, Slovakia
[(b)]Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava, Slovakia

[(a)]zhaladova@gmail.com, [(b)]bladeszasza@gmail.com,

ABSTRACT
The possibility of creating museum guides utilizing not just audio or textual information leads to emerging of new augmented and multimedia solutions. In this work in progress paper we want to propose Augmented Gallery Guide developed for the common smartphone device with the Android operating system. The guide combines the audio with the augmented reality and creates an emerging user experience. This paper discusses several technical issues of the creation of the guide, especially the recognition of the exponates.

Keywords: museum guide, augmented reality, cultural heritage, smartphones, local features

1. INTRODUCTION

The progress in the field of Augmented reality and the context aware systems since the 90's induce the strong interest in cooperation of computer scientist and the museums and galleries. The possibility of presenting the cultural heritage through new and popular devices, such as the smartphones became a way of attracting young people. The research in this area is nowadays mostly focused on extending the information about museum exhibits with virtual textual or visual information.

In this paper we will propose a novel concept of augmented reality guide adjusted for the use in galleries. Our guide will consist of augmented reality with synchronized audio.

In the area of the museum/gallery guides we can recognize several types of guides: a person, a book, an audio guide, a visual (interactive) guide and an augmented reality guide. We will now focus not only on the augmented reality guide but also on the visual (interactive) guide as it is closely connected with the first one. Both of these types can possibly be multimedia guides when engaging audio or another media.

The difference between these two methods is the fact that the visual (interactive) guide does not fulfil all 3 conditions on AR system proposed by Azuma [Azuma 1997]:

1. Combines real and virtual
2. Interactive in real time
3. Registered in 3D.

The visual guide solution usually interactively recognizes different exponates and display the virtual content, but they don't register the virtual objects within the real environment.

Many different implementations of the systems of one of these two types have been published since the first content-aware system presented in [Abowd et al. 1997]. We will focus on these in the following section. For completeness, we have to mention that there is also some research on Augmented audio guides, with no visual information, for example [Zimmermann and Lorenz 2008].

This paper is organized as follows. In the second section we will define the area of museum/gallery guides, specify different aspect of such systems and present different previous approaches. In the third section we will focus on different methods for the recognition of the exponates. In the following section we will propose the new Augmented gallery guide concept based on conclusions from the previous section. Next section will evaluate the system and formulate a conclusion. In the last sections we will focus on the future work and the acknowledgements.

2. PREVIOUS WORKS

In the previous section we have defined two museum guide types we will be focusing on: visual (interactive) guide and augmented reality guide.

We can divide the hardware solutions of these two approaches together on 3 types:

1. Head mounted
2. Spatial
3. Handheld.

The Head mounted solution such as [Flavia 2002] uses the Head mounted display to provide the user with the immersive experience. This device is usually owned by the museum and can be borrowed by the user. Problem with HMD is the ratio between the ergonomic parameters of the device and the resolution of the displayed augmented reality. Although several different types of HMD are known (Optical see through, Video see through, HMD Projectors, Retinal displays - for the

details see [Bimber and Raskar 2005]), we think that the future applications will be developed for the devices such as Google glasses. The main advantage of the HMD concept is the hands-free setup.

The spatial augmented museum guide was proposed by [Kusunoki et al. 2002]. The guide consisted of interactive sensing board and was capable of recognizing different objects using RFID technology, which will be discussed in the next section. Other types of spatial augmented museum solutions were investigated by the O. Bimber [Bimber et al. 2006; Bimber and Raskar 2005]. This category encloses all the different solutions utilizing the transparent displays, mirror beam combiners or holographic set-ups. The main advantage of the spatial solutions is that they are usually suitable for more users cooperation, provide hands-free set-up and have theoretically unlimited field of view. They seem to be the best choice for static applications.

The last category encloses all the handheld solutions including smartphones, PDAs, pocket PCs, Palmtops, Tablets, netbooks and small notebooks. Different systems implemented for such device can be found in [Fockler and et al. 2005; Bay et al. 2006; Abowd et al. 1997; Miyashita et al. 2008; Bruns et al. 2007]. Handheld devices became more and more common and popular and their performance is increasing rapidly. Nowadays smartphones are equipped with quad cores processors and with their popularity and availability became the best platform for the guiding systems. Main disadvantage (according to [Bimber and Raskar 2005]) is the non-hands-free setup and the relatively small field of view of the device.

3. RECOGNITION OF EXPONATES

In this section we will deal with recognition of the exponates and different methods proposed in the previous works. The previous augmented and visual museum guide solutions can be divided by the method of recognition of the exponates on:

1. Visually based
2. Outside-in, Inside out systems
3. Dead-reckoning systems
4. Combination of systems
5. User input based

3.1. Visually based

The visually based approaches utilize the image from the camera to recognize the exponate and estimate its proper 3D position (when creating Augmented reality). There are three different approaches to the visual recognition of the exponates. In the first case, system utilize binary markers (e.g. black and white ARToolkit tags) which has to be printed and placed (registered) near the exponate ([Wagner and Schmalstieg 2003]).

The second approach is based on the matching of the local features in the camera frame with the preliminary acquired database of photographs of exponates ([Bay et al. 2006]). The method consists of the detection of the interesting points in the image (frame), their description by the feature vectors and the matching of these feature vectors with feature vectors from the objects (exponates) in the database.

The local feature methods usually used are based on the SURF [Bay et al. 2006] or other methods such as SIFT [Lowe 1999], ORB [Rublee et al. 2011] or combinations of different detectors (FAST [Rosten and Drummond 2006], Harris corners [Harris and Stephens 1988]) and descriptors (BRIEF [Calonder et al. 2010]).

The third approach consists of recognition of the exponates using global features (for example colour histograms, histograms of gradients...). As the representative of this approach we can mention *PhoneGuide: museum guidance supported by on-device object recognition on mobile phones* [Fockler and et al. 2005] which uses the global features and the neural networks for the recognition of exponates.

The visual methods usually require more computations, and the recognition is slower and not hundred per cent precise. On the other hand they are very cheap and more portable to different museums as only the photographs of all exponates are required (except the first one with binary markers).

3.2. Outside-in, Inside out systems

The emitter-receiver (or sensor) based approaches are used for the visual museum guides purposes as they do not provide us with the exact 3D position of the exponate in the space. These solutions are usually very precise however there is a need to place additional components (such as the emitters, receivers or sensors) in the museum area. The typical implementation of this approach uses the Infrared emitters and readers ([Flavia 2002]) or the RFID tags and readers ([Kusunoki et al. 2002]).

3.3. Dead-reckoning systems

In the paper Personal positioning based on walking locomotion analysis with self-contained sensors and a wearable camera [Kourogi and Kurata 2003] Kourogi and Kurata proposed the indoor user dead reckoning (calculating current position by using a previously determined position) tracking system composed of an accelerometer, gyroscope, magnetometer, camera and a head tracker. The System is not dependent on any external markers, chips or sensors and determines the user's relative position from the variation of the vertical and horizontal acceleration (caused by human walking locomotion). To estimate the absolute position they used additional method of matching the camera stream with the database of images (prepared beforehand) utilizing the Kalman filter framework [Kalman 1960].

A system like this can be used in the museum for standalone recognition of exponates based on user position (not very robust). Another possibility is to

provide the user position as an additional information to the visual recognition system.

3.4. Combination of systems

The fourth category encloses all the solutions which combine some of previous approaches. For example in [Bruns et al. 2007] the authors utilize the rough user position from the Bluetooth emitters placed in every room with the combination of exponate recognition using global features and neural networks. This method averages all of the positives and negatives of previous solutions. It is more accurate and faster than the visual methods and also fewer components need to be placed in the museum. On the other hand there is still the necessity of acquiring photographs of all the exponates.

3.5. User input based

The last category encloses the oldest solutions which utilize the user input instead of an automatic recognition. This method is not appropriate for the augmented reality guides, because it cannot register the exponate in 3D. However it will allow for the easy and straight forward implementation of an interactive guide without special exponate recognition algorithms.

When comparing different systems it is necessary to take another things into account: the price of the solution for the museum, necessity of additional components, portability and necessity of the pre-acquired photographs.

Apart from the type of the guide and the discussed recognition method we can mention three additional important criteria:

- The type of the information displayed to the user (text, image, video, 3D object),
- The type of the device (common handheld device/special components required),
- The necessity of sending the computations to the external server.

4. AUGMENTED GALLERY GUIDE

As we mentioned before our goal is to create an Augmented Gallery Guide. The main reason, why we decided to focus on the galleries instead of the museums is the fact that in galleries majority of exponates are paintings. The main advantage of paintings when compared to other exponates is that they are planar and it will be easier to properly register an augmented layer on them.

Based on the previous conclusions we have created several criteria on the Augmented Gallery Guide:

1. It has to run on common smartphones on Android platform
2. It does not require the installation of any additional components in the museum area,
3. It combines the augmented reality (as defined in [Azuma 1997]) with the synchronized audio comments and all the computations will be carried on the device.

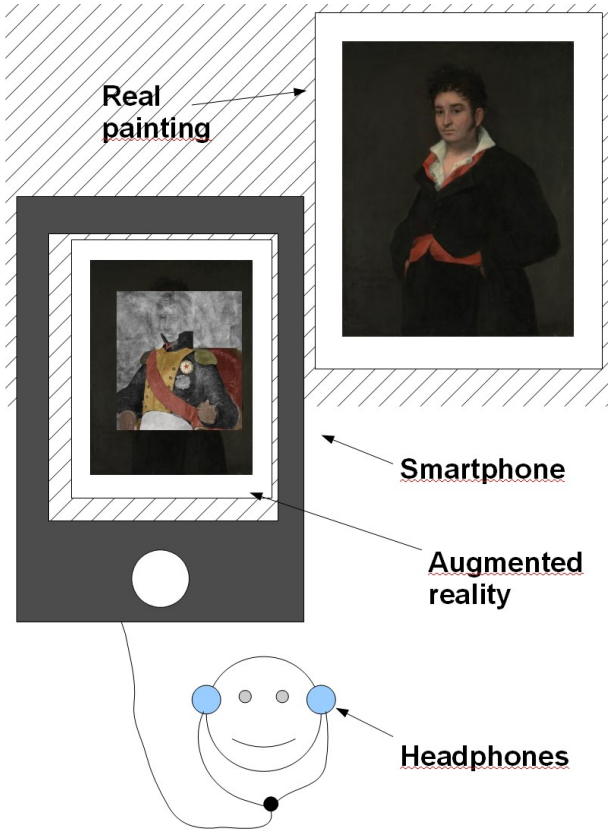


Figure 1: The scheme of the proposed Augmented Gallery Guide. The user holds a smartphone and wears the headphones connected to the smartphone. Camera on the smartphone streams recorded reality. The virtual information is augmented with the recognized and registered painting.

To fulfil all of the criteria we have to exclude all the spatial and head-mounted solutions. Also the system which recognizes the exponates with the sensor based or the combine solution and the visual solution using markers will be excluded. We have decided to use the local features for the recognition and registration of the exponates. We choose the FAST [Rosten and Drummond 2006] feature point detector and the BRIEF [Calonder et al. 2010] binary descriptor.

The main advantage of the BRIEF descriptor is the fact that it produces vectors of binary feature which can be easily matched using the Hamming distance metric (instead of the L2 norm commonly used for the matching). This causes that BRIEF descriptor can be matched very efficiently in comparison to SURF [Bay et al. 2006] or SIFT [Lowe 1999].

The main disadvantage however is, that BRIEF does not produce scale or rotation invariant descriptors. The rotational invariance is avoidable, and can be solved by utilizing the gyroscope of the smartphone. On

the other hand the scale invariance is more important and can be solved by the storage of the paintings database in several scales.

The input of our system is a pre-processed database of the paintings we want to augment. For each painting we have a virtual image, object or video and the audio track stored. In the pre-processing phase we process the database of the paintings images and compute the keypoints using FAST detector and then feature vectors for all keypoints using BRIEF descriptor. For every image we have the file containing these feature vectors stored.

In the run time, our system works as follows: First the frame is grabbed from the video camera of the smartphone. Then the keypoints in the image are found using the FAST detector and the descriptors are computed using BRIEF. Afterwards these descriptors are matched with the database of the paintings' descriptors. A good match of two keypoints is estimated using the second nearest neighbour strategy. The painting in the frame is recognized as the database painting with most matches.

Then we use RANSAC algorithm [Rublee et al. 2011] to exclude the outliers and estimate the homography between the painting from the database and the one from the frame. We can then estimate the position of painting's corners and draw a quadrilateral on the frame. In this moment the user is allowed to trigger the (augmented) visual and audio content.

In the next frame after the recognition, the following loop starts. The frame is grabbed and the keypoints and the descriptors (feature vectors) of these keypoints are computed. These are then matched only with the descriptors of the database image recognize in the previous frame. RANSAC is used and homography is computed. This loop proceeds until the painting is still present in the image, i.e. the number of good matches exceeds a threshold.

When the painting is no longer presented in the frame, the visual content is no longer available. However the audio track is still proceeding until stopped by the user. The pipeline of the run-time can be found on Figure 2.

If we look closely on the application, in the first step the application's main thread starts two threads: one video thread which streams the input video from the camera, and one computing thread. The Video thread is needed to satisfy the user, because it shows always high frame per second (FPS) video.

Some attention has to be paid to the concept of the mixed augmented and audio solution. Our solution is mainly created for the augmentation of the paintings. If we want to provide the user with some additional interesting information on painting and also create an augmented experience on the small screen, we have to somehow cover the display with the text frames.

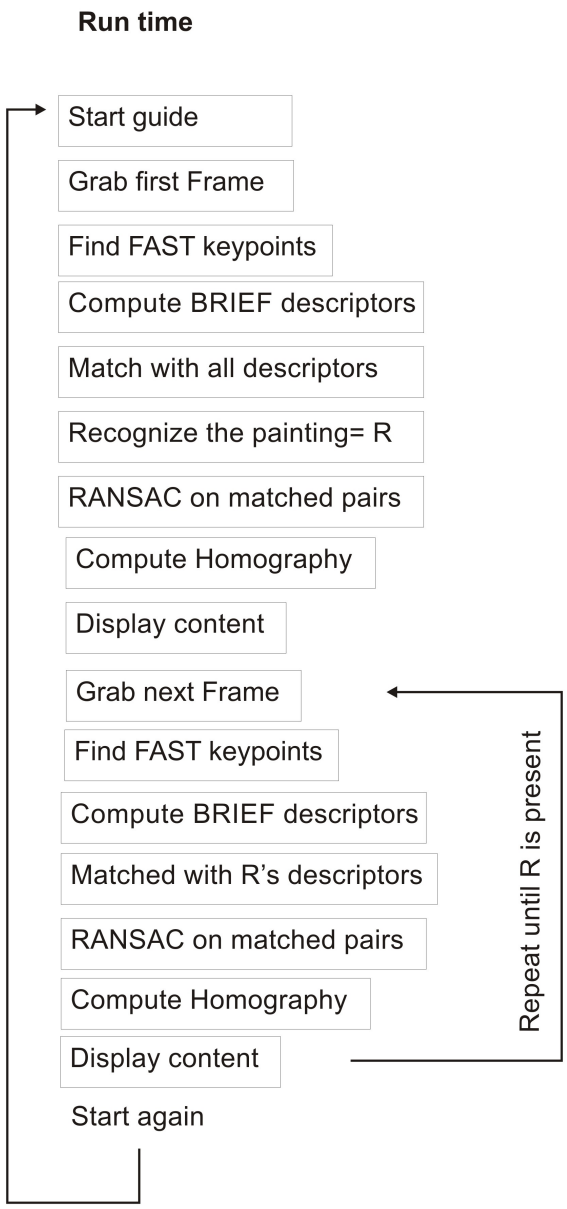


Figure 2: The pipeline of the run-time of the Augmented Museum guide system.

However this approach has no advantage compared to visual guide solution as it does not take advantage of the augmentation. On the other hand if we eliminate all the text information we can lose the role of the actual guide. To preserve both the augmentation and the guide at the same time and to utilize all the possibilities of the augmented reality guide we have decided to add the audio information. The scheme of our proposed system can be seen on the Figure 1.

Based on our tests of the application, we have also decided to create two different modes in our

application. The first one is the proposed Augmented reality solution and the second one is the Virtual reality mode which can be displayed after the painting is recognized in the previous mode. This mode was created because it is not very comfortable to point your smartphone's camera to the painting for several minutes (to see displayed visual content while listening to the audio track). The virtual reality mode enable user to watch the paintings "augmentation" in the lastly processed frame.

5. CONCLUSION

In this paper we have proposed the Augmented Gallery guide system based on comparison of different systems and conclusions we have made. Proposed system utilize the recognition of the exponates (paintings) using the matching of the local feature vectors of paintings in the database and in the camera frame. We have implemented our prototype for the Android platform using OpenCV library. Our current prototype is running on 5 fps on the common smartphone.

We have decided to design the first solution for the small gallery as this eliminate two shortcomings. The first one is the fact that the preparation of the material (audio tracks, images...) for each painting is manual and though time consuming. The second shortcoming is connected with the matching of the feature vectors. To search the database of hundreds paintings can apparently slow down the application.

6. FUTURE WORK

In the next phase we want to complete the proposed Augmented Gallery Guide and provide the complex user study in which we want to focus on the several aspect of user gallery visit. Firstly we want to measure the time spend in the gallery by the user with Augmented Gallery Guide, book guide or no guide at all. In the second phase we will provide the visitors with the questionnaire containing questions about user experience with the guide, but also test questions. The goal is to investigate if the user can acquire more interesting information using a guide.

ACKNOWLEDGEMENTS

This work is partially funded from project KEGA 068UK-4/2011.

REFERENCES

Abowd, G., Atkeson, C., Hong, J., Long, S., Kooper, R., and Pinkerton, M. 1997. Cyberguide: A mobile context aware tour guide. *Wireless Networks 3*, 421–433. 10.1023/A:1019194325861.

Azuma, R. 1997. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments* 6, 4, 355–385.

Bay, H., Fasel, B., and Van Gool, L. 2006. Gool. Interactive museum guide: Fast and robust recognition of museum objects. *In Proc. Int. Workshop on Mobile Vision*.

Bimber, O., and Raskar, R. 2005. Spatial Augmented Reality: Merging Real and Virtual Worlds. *A K Peters/CRC Press*, July.

Bimber, O., Frohlich, B., Schmalstieg, D., and Encarnac¸˜A O, L. M. 2006. The virtual showcase. *In ACM SIGGRAPH 2006 Courses*, ACM, New York, NY, USA, SIGGRAPH'06.

Bruns, E., Brombach, B., Zeidler, T., and Bimber, O. 2007. Enabling mobile phones to support large-scale museum guidance. *Multimedia, IEEE* 14, 2 (april-june), 16 –25.

Calonder, M., Lepetit, V., Strecha, C., and Fua, P. 2010. Brief: Binary robust independent elementary features. *In Computer Vision ? ECCV* 2010, K. Daniilidis, P. Maragos, and N. Paragios, Eds., vol. 6314 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 778–792.

Flavia, S. 2002. The museum wearable: real-time sensor-driven understanding of visitors' interests for personalized visually augmented museum experiences. *In In: Proceedings of Museums and the Web* (MW2002, 17–20.

Fockler, P., and et al. 2005. PhoneGuide: museum guidance supported by on-device object recogn. on mob. phones. *In MUM '05: Proc. of the 4th intern. conf.*, ACM, USA, 3–10.

Harris, C. and Stephens, M. A combined corner and edge detection. *In Proceedings of The Fourth Alvey Vision Conference*, pages 147-151, 1988.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. *Transactions of the ASME- Journal of Basic Engineering* 82, Series D, 35–45.

Kusunoki, F., Sugimoto, M., and Hashizume, H. 2002. Toward an interactive museum guide system with sensing and wireless network technologies. *In Wireless and Mobile Technologies in Education*, 2002. Proceedings. IEEE International Workshop on, 99 – 102.

Kourogi, M., and Kurata, T. 2003. Personal positioning based on walking locomotion analysis with self-contained sensors and a wearable camera. *In Proc. of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality*, 103–112.

Lowe, D. 1999. Object recognition from local scale-invariant features. In Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, vol. 2, 1150 –1157 vol.2.

Miyashita, T., Meier, P., Tachikawa, T., Orlic, S., Eble, T., Scholz, V., Gapel, A., Gerl, O., Arnaudov, S., and Lieberknecht, S. 2008. An augmented reality museum guide. *In Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, Washington, DC, USA, ISMAR '08, 103–106.

Rosten, E., and Drummond, T. 2006. Machine learning for high-speed corner detection. *In Computer Vision ? ECCV 2006*, A. Leonardis, H. Bischof, and A. Pinz, Eds., vol. 3951 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 430–443.

Rublee, E., Rabaud, V., Konolige, K., and Bradski, G. 2011. Orb: An efficient alternative to sift or surf. *In Computer Vision (ICCV)*, 2011 IEEE International Conference on, 2564–2571.

Wagner, D., and Schmalstieg, D. 2003. First steps towards handheld augmented reality. *In Wearable Computers, 2003. Proceedings. Seventh IEEE International Symposium on*, 127 – 135.

Zimmermann, A., and Lorenz, A. 2008. Listen: a useradaptive audio-augmented museum guide. *User Modeling and User-Adapted Interaction* 18, 389–416. 10.1007/s11257-008-9049-x.

AUTHORS BIOGRAPHY

Zuzana Haladova was born in 1987 in Trnava, Slovakia. She received Msc. Degree in Applied Informatics- Computer graphics from Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava in 2010. She is a PhD. student on Informatics at the department of Applied Informatics at the same university. She is the member of IVG (Image and Vision group). Her research interests include Augmented reality and Object recognition. Currently she is working on her PhD. thesis called: Augmented reality in the cultural heritage applications.

Csaba Boylos was born in 1989 in Lučenec, Slovakia. He received Bc. Degree in Applied Informatics from Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava in 2010. He is a Msc. student of Computer graphics and Geometry at the same university. This work is connected with his diploma thesis. His research interests include Augmented reality on smartphones and Local feature based recognition.

A BPMN GENERAL FRAMEWORK FOR MANAGING TRACEABILITY IN A FOOD SUPPLY CHAIN

Giovanni Mirabelli [(a)], Teresa Pizzuti [(b)], Fernando Gómez-González [(c)] , Miguel A. Sanz-Bobi [(d)]

[(a)] [(b)]Department of Mechanical Engineering, University of Calabria, Rende 87060, Italy
[(c)] [(d)] Department of Information Systems Engineering, Comillas Pontifical University, Madrid 28015, Spain

[(a)] g.mirabelli@unical.it, [(b)] teresa.pizzuti@unical.it, [(c)] fgomez@upcomillas.es, [(d)] masanz@upcomillas.es

ABSTRACT

In the research area of the supply chain, traceability is the result of many developments aimed at improving food quality and safety management. This paper presents the results of the first phase of elaboration of a Global Track&Trace (T&T) System for Food. A general framework is obtained through the definition of a T&T Information System. The development of an information system requires modeling of business processes and associated data results. In this research work, the whole supply chain has been modeled according to a Business Process Modeling Notation (BPMN). A general data model is proposed enough flexible and variable for developing the strategy of traceability and open the door to incorporate new future features to be taken into account. Processes and data management are achieved through the creation of a web-based system. The final model permits the supply chain optimization and the food quality management.

Keywords: Food Supply Chain, Tracking and Tracing, Information Systems, BPMN, ER model

1. INTRODUCTION

Traceability is a newer policy that is used to improve supply management, increase safety and quality, and differentiate finished goods on the basis of credence attributes (E. Golan, B. Krissoff, and F. Kuchler, 2004a). The increasing interest in food traceability directly interfaces with the introduction of new regulations and customer demands on food quality and safety.

In many developing countries, traceability initiatives have been started in the last decade and, within the European Union, it has been enshrined in a number of regulatory initiatives such as the Regulation of the European Community n.178/2002 (European Commission, 2002) . As a consequence, different types of traceability systems are emerging as a result of regulatory interventions, at an industry-wide level or as a competitive strategy at the level of individual supply chain. Moreover, the key issue frustrating the job of food safety agents are (i) the inability to link food chains records, (ii) the inaccuracy and the errors in the records and (iii) the delays in obtaining essential data. The recent cases of E.Choli in Germany are one example of the strong reaction of the market to a food outbreak disease. The E.Choli diffusion highlighted that when a food crisis occurs, rather than leaving potentially contaminated food in the market, authorities recommend the recall and removal from the market of all suspected food or recommend that consumers stop consuming the food products. This creates a huge amount of financial problems for companies which were not involved with the production or manipulation of the contaminated food.

In such a context, a traceability system may serve many purposes. Essentially, it functions as a tool for communication, making information available along the food supply chain. In order to maintain food safety, the information can be used to trace back and to find what the source and the cause of a problem is, to stop the problem or prevent it from happening again.

From a regulatory viewpoint, traceability is a requirement limited to ensure the ability for businesses to identify at least the direct supplier of a product as well as the immediate client, with the exemption for retailers (European Commission, 2002; European Commission, 2004). Notwithstanding, other requirements should be satisfied to ensure food security and to improve food quality (Food Standard Agency, 2002). Additional information should be collected in each stage of the Supply Chain to ensure the availability of data for the production analysis and optimization (Thompson et al., 2005).

The paper is structured as follows. Section 2 describes the food chain traceability and presents a brief review of the state of the art about traceability systems developed. Section 3 describes the main issues which deal with the implementation of a traceability system. Section 4 describes the track and trace information system proposed. Finally in Section the conclusions are discussed.

2. TRACEABILITY IN THE FOOD SECTOR

This section provides an overview on food supply chain traceability and food supply chain and presents the problem statement in order to provide some background and highlights the main goal of this research work.

2.1. Food Supply Chain Traceability: Current State and Future Work

The European Commission defined food chain traceability as "the ability to follow a food component intended to be, or expected to be into a food product through all stages of Food Supply Chain" (European Commission, 2002).

The Food Supply Chain (FSC) is a complex structure formed by several actors that contribute to the production, distribution, marketing and supply of food products. On the basis of the definition provided by the Food Traceability Handbook (Revision Committee on the Hanbook for Introduction of Food Traceability Systems, 2007) in a typical FSC are involved five basic entities: the primary producer, the processing company, the distributor, the retailer and the transporter or third part carrier. Each actor performs a specific task. The primary producer, such as the fisherman, the grower or the farmer, is devoted to the production of raw material and ingredient that are successively transformed by the processing companies; the transporter moves the products from one actor to another; the distributor handles the food commodities; the retailer sells food directly to the consumer. The presence of these actors highlights that the concept of food chain is extended both to the individuals upstream and downstream in the supply chain. In order to maintain the traceability, each actor must collaborate and share information in a coherent and shared form. In such a way it is possible to trace the path followed by a food product from "farm to fork".

Food Supply Chain (FSC) differs from the other supply chain because of the perishability which characterizes food product. As specified by Nishantha, Wanniarachchige, and Jehan 2010 the time windows in which food products moves from the from the raw material producer until the consumer remains relatively shorter in FSC. Food products, in fact, are extremely time critical and, by their nature, they are characterized by a short shelf. Food products are perishable and their shelf life is conditioned by the harvesting means, transformation processes, transporting ways, and storage conditions. This aspects, along with the wide variety of food products, contribute to making more difficult the design, implementation, and management of an efficient system of traceability (De Cindio et al., 2012)

Traceability is obtained through the combination of two different processes: tracking and tracing. These terms are often used in an interchangeably way even though they have different meanings. Tracking is the process by which the product is followed by upstream to downstream in the supply chain by recording date in each production stage. Tracing is the reverse process. Through tracking systems it is possible to trace the global history of the product and the responsibilities at different processing stages.

The operations required by a traceability management system can be divided into two main activities which refer to internal traceability and supply chain traceability. The internal traceability is realized by internal procedures, different for each business, that allow tracing the origin of materials used, the process operations and the food destination. The food supply chain traceability or external traceability is guaranteed by the integration and coordination of the tracking procedure adopted by each operator of the chain, and represents the ability to follow the path of a specific unit of product along the production chain.

The definition and implementation of a FSC traceability system depends on both the supply chain and the relationships between the various partners which collaborate in the production process. Manufacturers, distributors, authorities and consumers should be able to track and identify food and raw materials used for food production to comply with legislation and to meet the requirements of food safety and food quality (Ruiz-Garcia et al., 2010). This result can be conveniently achieved if each company along the supply chain is able to adopt a system of internal control and recording (internal traceability) information and if transitions between actors are regulated and managed in a coherent and shared form (De Cindio et al., 2011).

2.2 Problem Statement

Traceability systems are emerging proposing different approaches, as result of both regulatory and industry initiatives. Three key functions of a traceability systems have been identified (Hobbs et al., 2002; Sanderson and Hobbs, 2006;E. Golan, B. Krissoff, and F. Kuchler, 2004b). The firs key function is to allow efficient trace-back of products and inputs when a food safety or herd health problem occurs. In such a case, efficient and timely trace-back could limit the size of product recalls and the number of people exposed to tainted food, thereby limiting human-health impacts, minimizing productivity losses from illness, etc. The second key function is that traceability can be used to reduce information costs for consumers by identifying credence attributes through the labeling of environmentally-friendly production practices, or assurances about feed, other ingredients or production practices. In this case the traceability system is directly connected with the quality system of the company. The third function of traceability may be as a means of strengthening liability incentives to produce safe food.
Potential of traceability systems and the numerous advantages that can be obtained through its implementation have been well documented in literature (Moe, 1998; Lo Bello et al., 2004).

Despite multifaceted potential benefits, nowadays the operational conditions of current traceability systems are kept at bare minimum merely to fulfill legal requirements. Particularly Small and Medium Enterprises (SMEs) are either do not use traceability systems or use non-digital systems due to the limited scale of their operations, necessity of heavy investment and nature of their manufacturing process (Nishantha et al., 2010). Many barriers, in fact, hinder the successful

implementation of traceability, the most important are: necessity of costly investments, reluctance to change, lack of skilled staff to handle advanced systems and limitation of existing traceability systems. Meeting the traceability standard set by an industry organization or by government regulation, affects the cost of production per unit of food. Nevertheless, tracing and tracking capabilities are crucial to confine the reaction to possible hazards and reduce the recovery cost (Bechini et al., 2005).

The issue of food recall, and the consequently issue of money loss, can be easily solved through the introduction of a global traceability systems capable of enabling more targeted recalls, of identifying more strategically the product origin and consequently constraining the product recall only to the products actually affected by contamination (Pouliot and Sumner, 2009). In this way, in case of emergency, the outbreak of a disease can be immediately identified.

In recent years the traceability of food products has attracted the attention of many researchers for several reasons (Jansen-Vullers et al., 2003): first traceability, according to the Regulation of the European Community N. 178/2002, has become a legal requirement within the European Union from January 1, 2005 (European Commission, 2002); secondly, food companies tend to view traceability as a strategic tool needed to increase consumer confidence and improve the both image of the company and of a specific product.

Moreover, currently consumers have no access to the information on the real origin of products, the activities in which the products were involved and the operators who manipulated it. Many initiatives have been started in the area of food traceability in the last decade and several authors have been interested in the development of food traceability systems (Jansen-Vullers et al., 2003) ;Regattieri, Gamberi, and Manzini (2007); Bechini et al. (2008); Thakur and Hurburgh (2009); Thakur and Donnelly (2010); Thakur, Martens, and Hurburgh (2011), (Thakur, Martens, et al., 2011b; Thakur, Sørensen, et al., 2011); Bevilacqua, Ciarapica, and Giacchetta (2009); Ruiz-Garcia, Steinberger, and Rothmund (2010); Verdouw et al. (2010)). Despite the numerous efforts for developing effective traceability systems, current results obtained reveal some critical limitation of existing traceability systems (Bechini et al., 2005). Successful implementation of traceability systems requires elevate investment costs, staff training and global legal requirements.

2.3 Research Objective

The research objective of this work is the development of a new general framework for the traceability of food products able to support quality and safety control. The solution adopted for the information management to support traceability is generally applicable, which means that it meets requirements from various kinds of industries. This model can be applied in real-life situations that might benefits from traceability solutions.

This paper aims to contribute to the development of a reference model in food traceability and it presents the result of the first part of a complex research work. At this step, a new Traceability Information System is developed. The system is obtained through the modeling of business processes and the development of a data model which contains all the information required for traceability. In order to ensure traceability and to implement an effective traceability system, the supply chain has been modelled using the Business Process Modelling Notation (BPMN) (Object Management Group, 2010), a common language understandable by analysts and developers. On the basis of the FSC model, data has been modeled following the Entity-Relationship (E-R) model (Hoffer et al., 2010). The integration of the process model with the data model has led to the generation of a web application that can be used for traceability purposes and supply chain management and optimization.

3. MODEL DEFINITION

In this section the food Track&Trace general framework is described and a short introduction to Business Process Modeling Notation is provided.

3.1 Business Process Modelling Notation

The Business Process Modelling Notation (BPMN) is a new standard to model business process flows and web services. It is a graphical notation that depicts the steps in a business process. BPMN depicts the end to end flow of a business process. The notation has been specifically designed to coordinate the sequence of processes and the messages that flow between different process participants in a related set of activities (Object Management Group, 2011). A business process currently spans multiple participants and coordination can be complex. Until BPMN, there has not been a standard modelling technique developed that addresses these issues. BPMN has been approved to provide users with a royalty free notation.

The primary goal of BPMN is to provide a notation that is readily understandable by all business users, from the business analysts that create the initial drafts of the processes, to the technical developers responsible for implementing the technology that will perform those processes, and finally, to the business people who will manage and monitor those processes.

BPMN aims at bridging the gap between business process design and process implementation. It allows the automatic translation from a graphical process diagram to a BPEL process representation that may be then executed using a Web services technology. Another goal, but no less important, is to ensure that the XML language designed for the execution of business processes, such as WSBPEL (Web Services Business Process Execution Language), can be visualized with a business-oriented notation.

The Business Process Modelling Notation is especially used in Service Oriented Architecture (Object Management Group, 2011). In a Service Oriented Architecture (SOA) approach, business processes models are leading in routing event data among multiple software components that are packaged as interoperable services (Erl, 2005; Papazoglou et al., 2007). The main elements of the BPMN are showed in Figure 1.

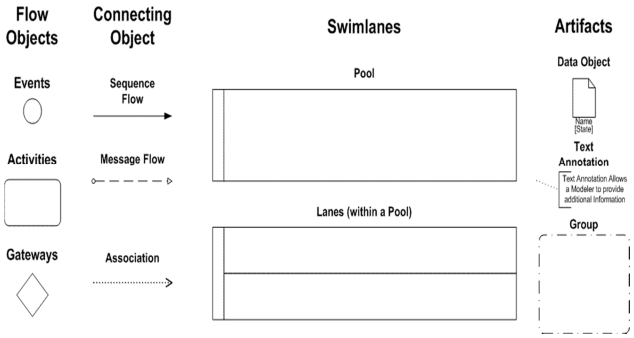


Figure 1-Core elements of the BPMN

BPMN allows reconstructing patterns of process or the Business Process Diagram (BPD) by means of graphs or networks of objects. These objects represent the activities of the process and are connected by control flows that define logical relationships, dependencies, and the execution order. The advantage of using BPMN concerns with the model dynamicity: in fact, the transition from one version to another one permits to add or cancel some elements of the model without the necessity of reprogramming the application using a specific language. According with BPMN, the actors involved in the supply chain have been classified into pools and external traceability is obtained through the flow of messages. The choice of the BPMN as standard to model the process flow is directly connected with the advantage of integrating actors, tasks and data in a single model. The flow of products lots along the supply chain is associated with information exchanges among responsible actors and possibly third-party organizations.

3.2 Food Track&Trace General Architecture

In this paragraph, the general framework of the proposed Food Track&Trace System is described. The development of efficient traceability information systems in food chains has assumed considerable importance in recent years. The ability to trace and track every single unit of product depends on the supply chain traceability system which in turn depends on the internal data management system and the information exchanged between the actors.

Generally, an information system is formed by the business process models and associated data resources.

In the proposed research work traceability is obtained through the development of a SOA application. Process and actors involved in the supply chain and the relationships between them have been modeled using BPMN. As mentioned in the previous paragraph, the BPMN is the new standard for modelling business processes and web service processes, as put forth by the Business Process Management Initiative (BPMI – www.BPMI.org) that makes web services work in a four-stage process, as follows (Owen et al., 2003):

1. Design the processes using BPMN.
2. Simulate the processes and modify them for efficiency.
3. Make the services available by publishing them using a Business Process Execution language.
4. Orchestrate the web services into end-to-end business flows by assembling them and coordinating their behavior. Business Process Management Systems (BPMS) are employed for this stage.

Process models represent specific ordering of work activities across time and place, including clearly identified inputs and outputs (Davenport, 1993). They represent the sequence of activities, events and control decisions.

In the general framework proposed, the whole food supply chain where initially modeled in order to visualize the basic processes, the actors involved and the related product and information flows. Taking into account the Supply Chain Operations Reference (SCOR) model of the Supply Chain Council (Supply Chain Council, 2010), for each actor present in the model, have been defined the most important operations.

A simple schema of the BPMN model is described in Figure 2. The general model shows the different processes operated by each actor along with the information flow. Actors involved in the supply chain have been classified into pools and information has been organized in the form of business objects. Each actor records data of products and processes and collaborated with other operators in the industry by making available all the information necessary for traceability. The main agents modeled in the General Framework are the primary producer, the processor, the transporter, the wholesaler and the retailer. Because of the different features characterizing each primary producer, this agent has been successively modeled and dived into three other actors: primary raw material producer, secondary raw material producer and third raw material producer. For example, when referring to the vegetable field, these actors represent respectively the seeder, the nursery and the grower.

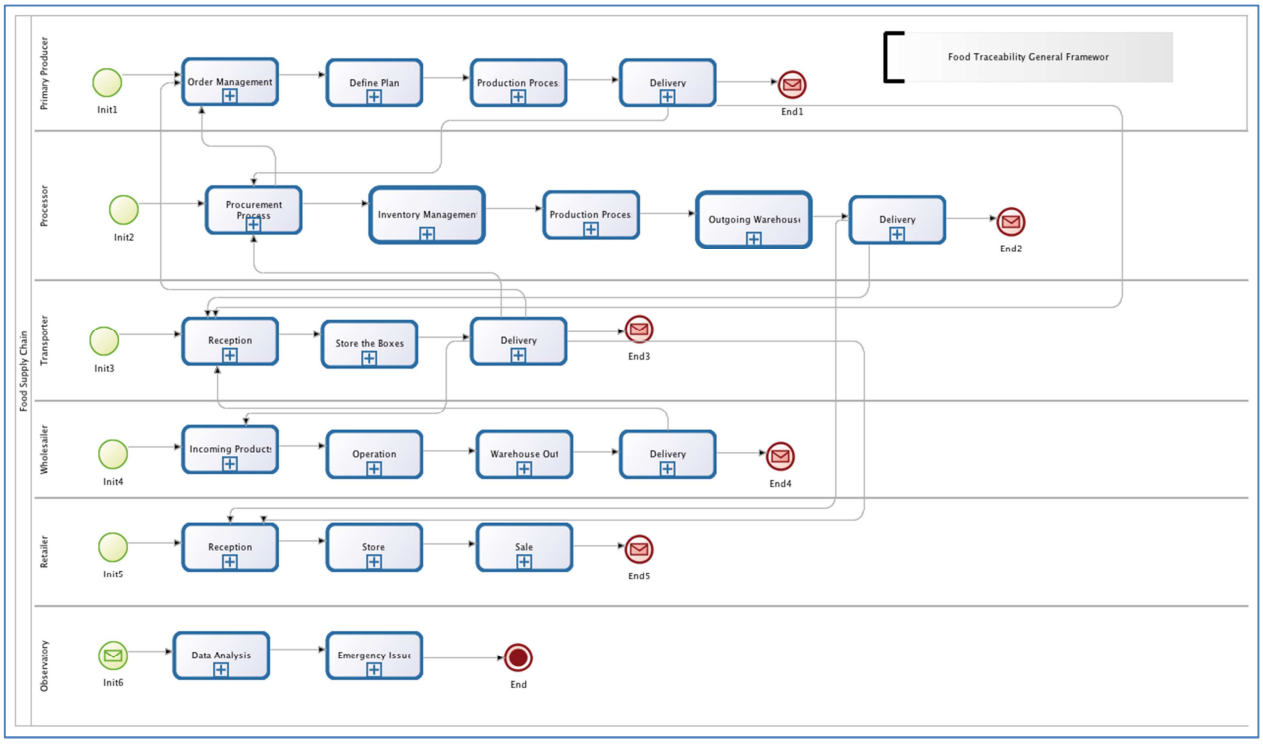


Figure 2- General Model

Transformations and logistics operation such as sourcing and delivering have been considered for each actor. An operation of transformation is operated each time that a lot is manipulated and each time that it leads to the definition of new lots. In these cases, according to Bechini et al. 2008 the lot behavior has been modeled by the following activity pattern: lot integration, lot division, lot alteration, lot movement.

The General BPMN Model of Figure 2 shows common processes for some actor, such as "production process" for primary producer and processor, "delivery" or "sale" for all the actors, "Reception for transporter and retailer and similar to this "incoming products", "procurement process" and "order management" respectively in wholesaler, processor and primary producer.

Each activity showed in Figure 2 represents a call activity which connects with another BPMN process.
A QR code is generated each time that a Traceability Resource Unit is formed. Different units of aggregation are defined according to the GS1 Global Traceability Standard (GS1 Standards Document, 2010). In particular:

- A Consumer Unit (CU) represents a single products, bags, and packages with a certain amount, volume or weight of goods.
- A Trade Unit (TU) is represented by cartons, boxes, pallets or bulk lots (in weight or volume).
- A Logistic Unit (LU) is generally represented by pallets and containers.

Each time that products are moved from one actor to another, the transporter generates a QR for each Shipping Unit (SU) and for each Logistic Unit (LU) manipulated. A Shipping Unit represent a truck or vessels loads.

Transportation can be done in different ways and using several means of transportation and different carries can be involved in the process of movement. The carries can refer to different companies. The general idea adopted in the Food Track&Trace System is that each carrier read the QR code of the product to move and generate a new QR code with the information on the carrier, the mean of transportation, the route, the date time of manipulation and delivery. The unit of transportation is manipulated each time that it is operated and activity of division or merge.

A central role in case of food accidents is played also by the Observatory. The Observatory is responsible for the management of the traceability system and recall activities in case of food outbreak disease. Each actor communicates with other actors of the supply chain and it constantly provides the Observatory with the information required for the traceability management. The Observatory contains a repository with all the information of the products produced in a particular location. The Observatory BPMN model is formed by two different lanes: the Data Analysis Department and the Emergency Task Department. In case of food accident or outbreak disease, with infected or dead people, the Emergency Task introduces in the System all the information about person, location and products eaten. For each food eaten the system generates a new table with all the ingredients and raw material used for transforming and producing the food in all the supply

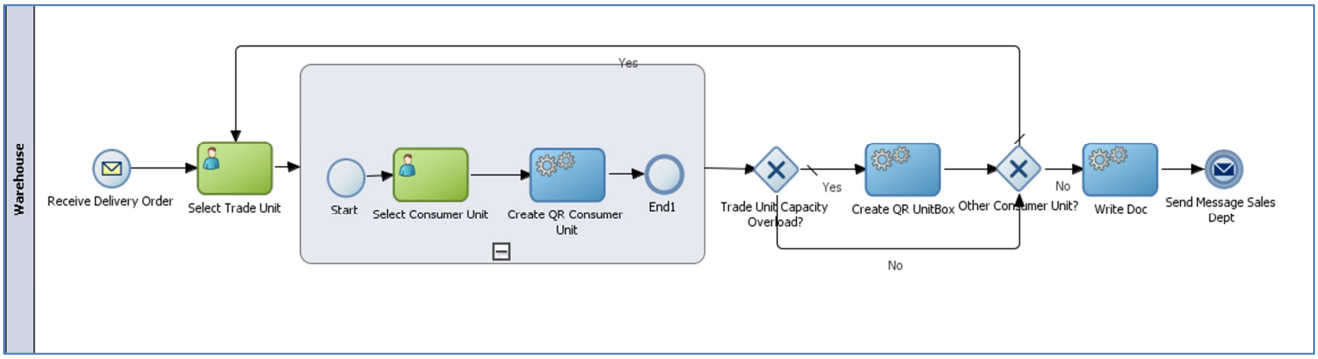


Figure 3- Process of Packaging and Labeling

chain stages. The system will show the maps relating to the introduced information and the probable location in which the infection has been generated.

One of the most important operations in the traceability maintenance is represented by the process of packaging and labelling. Figure 3 shows an example of the packaging process in which a QR Code is generated for each Consumer Unit and Trade Unit.

The core of the proposed process modeling is to identify common data and parameters and to construct a data model enough flexible to be adapted to different supply chains. The main objective of our work was to make available all the information at the final consumer. The tracking system is based on a service-oriented architecture (SOA) and the communication is based on messages in XML.

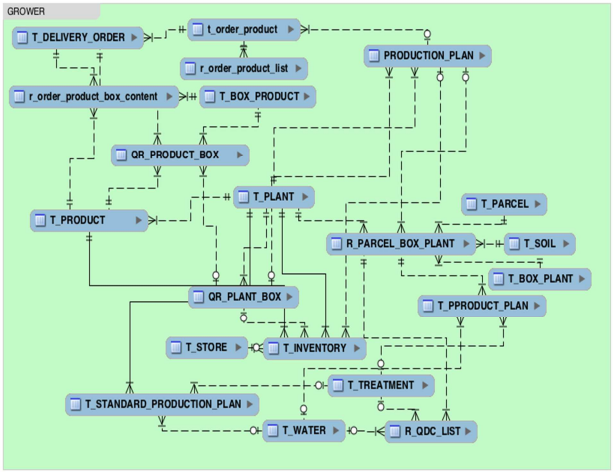


Figure 3- E-R Diagram of the Grower

It is required data to be recorded for the modeling and the analysis of product processes in the chain. An extended data model has been created following the Entity-Relationship (E-R) diagram (Hoffert et al., 2006) to support the whole architecture. The main elements of the data model are entities, or containers of data elements, and relationships, defined as semantic connections between entities. Entities are characterized by attributes. A general data model is proposed enough flexible for developing the strategy of traceability and open to incorporate new future features to be taken into account. The data modelling and management approach is achieved through the creation of a web based system. A MySQL database has been generated for each actor involved in the supply chain.

Figure 4 shows the database diagram that has been generated for a particular actor, a grower, who is responsible for seedling seeds and for the growing of a plant until a certain period or dimension. In the model all the information about treatments and location are recorded, including the geographical information of the land in which the parcel is located.

By using information from the production environment it is possible to provide relevant details on local environmental conditions which contribute to the particular uniqueness of the products, as soil, landscapes and climatic conditions, and to certify the origin of a particular product. In addition, the indication of origin (soil, region, country) becomes objective data with special regards to the new requirements of food safety and environmental protection.

A general problem of the previous traceability systems is that actors positioned downstream in the supply chain generally have no information about operations and treatments operated by the previous actor in the chain. In order to solve this problem, it was important to consider that all the information on a product is directly included in a QRcode. The QR-code, which was developed by Denso Wave (http://www.qrcode.com), is known as a kind of 2D barcode. The features of this code symbol are large capacity, small printout size and high speed scanning etc. A QR code is generated each time that a traceability unit is generate, moved, or manipulated.

4. CONCLUSIONS

The outbreak food diseases of the past years show as more information is necessary, and that a global traceability system is fundamental in a global market. In addition to systematically storing information that must be made available to inspection authorities on demand, a traceability system should take also food safety and quality improvement into account. To take into account the current requirements on food quality for health care into consideration, additional data that is not strictly necessary for traceability must be stored. For instance, for a cooking activity, oven temperature and humidity

can be considered important parameters in case of hazard. For a cultivation activity, operations on the parcel are fundamental to trace the proximity of the land for cultivation to a source of pollution.

The traceability system prototype presented in this paper is designed under a flexible and open perspective in order to facilitate integration of information across the entire supply chain, ensuring consumer trust and compliance with legal and quality standard. Communication exchange, which implies information transmission in a secure and reliable way, is regarded as an e-business transaction. The software tool utilized for the process modeling directly generates a web application model that assures the connection to each operator of the supply chain.

The framework includes a set of process models that are understandable by business manager in a notation that can be interpreted by SOA-based Information System (BPMN).

The main features of the general proposed framework are: (i) high flexibility, (ii) reduced development time, (iii) reduced implementation costs, (iv) high usability, (v) management and control of actors, processes and data, (vi) easy information exchange between the different actors of the supply chain, (vii) appropriate level of integration with the data system. However, there are critical issues to be addressed, such as the tightness of a mandatory standard for all companies in the food sector for the encoding of information.

Through the final system, each FSC operator can: (i) guarantee the origin and the quality of a food product; (ii) assure the compliance with regulation; (iii) improve logistics; (iv) improve the inventory management; (vi) management of the whole products lifecycle.

In addition, recorded data can be used for several analyses such as the definition of: (i) type and quantity of cultivation (plant, animals or fresh) per locality or region; (ii) type and quantity of cultivation per period or year; (iii) land surface availability to be allocated to a particular product; (iv) level of activities of a particular locality/region/country; (v) previsions; (vi) recommendations.

Finally, data and time information recorded at each production step will help identifying non-compliance in the case of storage. The system, in fact, can be used in order to avoid food fraud such as off-season sales and certify the total quality of the product.

5. FURURE RESEARCH WORK

In this paper are presented the results of the first phase of elaboration of a global Track and Trace System. A new traceability framework can be defined combining the advantages of an information system for food traceability with the advantage of a geographical information system. This system can help governmental authorities in case of food outbreak disease and offer more information to the customers on the products that eat.

The integration of the Global T&T Food System with the Geographic Information system is a strategic approach that can help in the Geotraceability maintenance. Nowadays, the use of data processing applications that can process geographical information is becoming more widespread. In parallel to this development, geographical data are becoming more widespread and increasingly accessible, in particular due to the generalization of spatial data infrastructure and the diffusion of tool like Google Earth (http://earth.google.com).

ACKNOWLEDGMENTS

This paper is co-funded with the support of the European Commission, European Social Fund and Regione Calabria (Italy).

REFERENCES

Bechini, A., Cimino, M.G.C.A., Lazzerini, B., Marcelloni, F., Tomasi, A., 2005. A General Framework for Food Traceability, in: Proceedings of the 2005 Symposium on Applications and the Internet Workshops, SAINT-W '05. IEEE Computer Society, Washington, DC, USA, pp. 366–369.

Bechini, A., Cimino, M.G.C.A., Marcelloni, F., Tomasi, A., 2008. Patterns and technologies for enabling supply chain traceability through collaborative e-business. Information and Software Technology 50, 342–359.

Bevilacqua, M., Ciarapica, F.E., Giacchetta, G., 2009. Business process reengineering of a supply chain and a traceability system: A case study. Journal of Food Engineering 93, 13–22.

Davenport, T.H., 1993. Process Innovation: Reengineering Work through Information Technology. Harvard Business School Press, Boston, Massachusset.

De Cindio, B., Longo, F., Mirabelli, G., Pizzuti, T., 2011. Modeling a traceability system for a food supply chain: Standards, Technologies and software tools. Presented at the MAS: The International Conference on Modeling & Applied Simulation, Rome, pp. 488–494.

De Cindio, B., F. Longo, G. Mirabelli, and T. Pizzuti. 2012. "Vegetable and Fruit Chain: Tracing Critical Aspects." Industrie Alimentari 51 (522): 9-20+28.

Erl, T., 2005. Service-Oriented Architecture: Concepts, Technology and Design. Prentice-Hall PTR, Englewood Cliffs.

European Commission, 2002. Regulation (EC) N. 178/2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety.

European Commission, 2004. GUIDANCE ON THE IMPLEMENTATION OF ARTICLES 11, 12, 16, 17, 18, 19 AND 20 OF REGULATION (EC) N° 178/2002 ON GENERAL FOOD

LAW: CONCLUSIONS OF THE STANDING COMMITTEE ON THE FOOD CHAIN AND ANIMAL HEALTH.

Food Standard Agency, 2002. Traceability in the Food Chain: A preliminary Study.

Golan, E., Krissoff, B., Kuchler, F., 2004a. Food traceability: One ingredient in a safe and efficient food supply.

Golan, E., Krissoff, B., Kuchler, F., 2004b. Food traceability: One ingredient in a safe and efficient food supply.

GS1 Standards Document, 2010. GS1 Global Traceability Standard: Business Process and System Requirements for Full Supply Chain Traceability.

Hobbs, J.E., Fearne, A., Spriggs, J., 2002. Incentive structures for food safety and quality assurance: an international comparison. Food Control 13, 77–81.

Jansen-Vullers, M.., van Dorp, C.., Beulens, A.J.., 2003. Managing traceability information in manufacture. International Journal of Information Management 23, 395–413.

Lo Bello, L., Mirabella, O., Torrisi, N., 2004. Modelling and Evaluating traceability systems in food manufacturing chains, in: Enabling Technologies: Infrastructure for Collaborative Enterprises, 2004. WET ICE 2004. 13th IEEE International Workshops On. pp. 173–179.

Moe, T., 1998. Perspectives on traceability in food manufacture. Trends in Food Science & Technology 9, 211–214.

Nishantha, G.G.D., Wanniarachchige, M.K., Jehan, S.N., 2010. A pragmatic approach to traceability in food supply chains. pp. 1445–1450.

Object Management Group, 2010. Business Process Model and Notation (BPMN) Version 2.0.

Object Management Group, A.S., others, 2011. Business Process Model and Notation (BPMN). Business 50, 504–507.

Owen, B.M., Raj, J., others, 2003. BPMN and Business Process Management Introduction to the New Business Process Modeling Standard. Popkin Sofware 2678.

Papazoglou, M.P., Traverso, P., Dustdar, S., Leymann, F., 2007. Service-oriented computing: State of the art and research challenges. Computer 40, 38–45.

Pouliot, S., Sumner, D.A., 2009. Traceability, food safety and industry reputation.

Regattieri, A., Gamberi, M., Manzini, R., 2007. Traceability of food products: General framework and experimental evidence. Journal of food engineering 81, 347–356.

Revision Committee on the Hanbook for Introduction of Food Traceability Systems, 2007. Handbook for Introduction of Food Traceability Systems (Guidelines for Food Traceability).

Ruiz-Garcia, L., Steinberger, G., Rothmund, M., 2010. A model and prototype implementation for tracking and tracing agricultural batch products along the food chain. Food control 21, 112–121.

Sanderson, K., Hobbs, J.E., 2006. Traceability and process verification in the Canadian Beef industry. Department of Agricultural Economics, University of Saskatchewan.

Supply Chain Council, 2010. Supply Chain Operations Reference (SCOR®) modelOverview - Version 10.0.

Thakur, M., Donnelly, K.A.M., 2010. Modeling traceability information in soybean value chains. Journal of Food Engineering 99, 98–105.

Thakur, M., Hurburgh, C.R., 2009. Framework for implementing traceability system in the bulk grain supply chain. Journal of Food Engineering 95, 617–626.

Thakur, M., Martens, B.J., Hurburgh, C.R., 2011a. Data modeling to facilitate internal traceability at a grain elevator. Computers and Electronics in Agriculture.

Thakur, M., Martens, B.J., Hurburgh, C.R., 2011b. Data modeling to facilitate internal traceability at a grain elevator. Computers and Electronics in Agriculture.

Thakur, M., Sørensen, C.-F., Bjørnson, F.O., Forås, E., Hurburgh, C.R., 2011. Managing food traceability information using EPCIS framework. Journal of Food Engineering 103, 417–433.

Thompson, M., Sylvia, G., Morrissey, M.T., 2005. Seafood traceability in the United States: Current trends, system design, and potential applications. Comprehensive Reviews in Food Science and Food Safety 4, 1–7.

Verdouw, C.N., Beulens, A.J.M., Trienekens, J.H., Wolfert, J., 2010. Process modelling in demand-driven supply chains: A reference model for the fruit industry. Computers and electronics in agriculture 73, 174–187.

FOOD TRACEABILITY MODELS: AN OVERVIEW OF THE STATE OF THE ART

Giovanni Mirabelli [(a)], Teresa Pizzuti [(b)], Fernando Gómez-González [(c)] , Miguel A. Sanz-Bobi [(d)]

[(a)] [(b)]Department of Mechanical Engineering, University of Calabria, Rende 87060, Italy
[(c)] [(d)] Department of Information Systems Engineering, Comillas Pontifical University, Madrid 28015, Spain

[(a)] g.mirabelli@unical.it, [(b)] teresa.pizzuti@unical.it, [(c)] fgomez@upcomillas.es, [(d)] masanz@upcomillas.es

ABSTRACT
The evolution of the information technologies and their impact in the human life promotes an increasing demand of reliable information when security and safety plays a primary role. Nowadays, food traceability represents one of the main concerns in public authorities and industry. Traceability has become a critical part of the agro-food industry. The aim of the agro-food traceability is to allow the full monitoring of a product in the supply chain and trace the history of a good from the producer to the consumer. It is therefore a preventive instrument of quality and safety management. This paper presents an overview about the state of the art of models and systems developed for food traceability. In the first part, a short overview of the regulatory state of the art is presented. Then, the main research works in the area of food traceability are discussed. Finally some considerations are provided.

Keywords: tracking and tracing, food supply chain, information systems,

1. INTRODUCTION

In recent years the traceability of food products has attracted the attention of many researchers for several reasons (Jansen-Vullers, van Dorp, e Beulens 2003): first traceability, according with the Regulation of the European Community N. 178/2002, has become a legal requirement within the European Union from January 1, 2005 (European Commission 2002); secondly, food companies tend to view traceability as a strategic tool needed to increase consumer confidence and improve both the image of the company and of a specific product.

The term traceability refers to the "ability to trace the history, application or location of an entity by means of recorded identifications" (United Nations Food and Agriculture Organization (FAO) 1999). The European Community Regulation N. 178/2002 of the European Commission defines traceability as the "ability to trace a food, feed or producing animal or substance intended to be part of a food or feed, through all stages of production, processing and distribution" (European Commission 2002).

The final scope of a traceability system is generally to ensure the complete monitoring of the Food Supply Chain and to assure the observation of two primary functions, tracking and tracing. Tracking refers to the ability to trace a product through the Supply Chain, from upstream to downstream, recording data in each production stage. Tracing is the reverse process of tracking: through the tracing systems it is possible to identify the source of a food or group of ingredients through the information recorded upstream in the Supply Chain (Schwägele 2005). Kim, Fox, and Gruninger 1995 state that a traceability system must be able to track both products and activities. Therefore, the maintenance of traceability is a complicated and expensive process especially with regard to processed foods. In case of processed foods, in fact, different lots of various raw materials are combined into several production batches typically distributed in various points of sale (Hu et al. 2009). Hence, data to record must include information on products and on processes that operate on products (such as transport, transformation or combination). This goal can be reached through the implementation of an efficient traceability system supported by appropriate architectural solutions (Bechini et al. 2008).

The development of a global and efficient food traceability system faces with the knowledge of the previous research work conducted in this field. The need to analyse and classify the previous works present in the field, led the authors to the definition of a review of the state of the art of the food traceability models.

A general classification of the scientific literature on food traceability has been proposed in the contest of the Trace Project (http://www.trace.eu.org/index.php). Scientific contributions have been classified into two main categories: General Traceability and Chain Traceability. The General Traceability category includes general articles on traceability system, such as papers about general frameworks, analysis of benefits, advantages and disadvantages. Under the category of Chain Traceability have been classified the scientific works which refer to a particular supply chain, such as meat/agro or seafood, those devoted to the analysis of particular technologies, the articles that define which benefits can be obtained through the implementation of a traceability systems in the supply chain, the papers that analyze the impacts in logistics, the articles which defines the regulatory framework in the European

Union. Moreover, there is no a clear definition of how these sub-categories have been obtained.
In this paper the authors propose a new classification of the previous works on traceability system for food products. Two different parts have been analysed:

- Regulation, recommendation, and guides of Governments;
- Scientific literature.

Scientific literature has been classified into two main categories: Mathematical models and Information models.

The category of Mathematical models includes a review of the works oriented to the definition of some mathematical models containing mixing and risk transmission problems. During the development of a traceability system the definition of the rules for the identification of product units and their complete monitoring plays an important role for reducing batch dispersion and optimizing products recall. To this end, particular attention has been devoted to the analysis of those papers focused on the analysis and modelling of the lot behaviour.

The category of Information models includes works on the definition and development of innovative traceability information systems. In this area, important considerations have been done on the evaluation of the different technologies that can be used for recording, managing and transferring information. This technologies, known as auto-identification technologies are: bard code, Radio Frequency Identification (RFID) technologies, Near Field Communication (NFC) systems, Real Time Locating Systems (RTLS). The awareness of benefits and costs related to the introduction of these technologies is fundamental in the development of a traceability information system. In addition, new tendencies show that ontologies can be used to set up innovative traceability semantic models.

This paper is structured as follows. Section 2 presents a brief review of the regulatory state of the art on traceability. Section 3 presents an overview of the scientific literature.
The analysis of the state of the art is a requirement of any step previous at the development of a traceability system.

2. REGULATORY OVERVIEW

The increasing concern on food safety matter has promoted that many governments have begun thinking the adequacy of the private traceability system and the possibility of adopting mandatory traceability systems to improve the social food safety level. Some of the nations and regions have required mandatory food traceability systems or encouraged voluntary traceability programmes (F. Wang et al. 2011). In many developing countries, traceability initiatives have been started in the last decade. They mainly refer to perishable and high-risk food export products like beef and fish, fruits and vegetables, but also coffee or wine.

In Europe, Regulation (EC) 178/2002 of the European Parliament and of the Council lays down the general principles and requirements of a food law. The principal aim of this Regulation is to protect human health and consumer interests in relation to food. It applies to all stages of production, processing and distribution of food and feed, but there is an exemption for primary production for private domestic use, and the domestic preparation, handling, or storage of food for private domestic consumption. The traceability requested is known as "one step back-one step forward", which means to identify the immediate supplier of the product in question and the immediate subsequent recipient. In fact traceability is a requirement limited to ensure the ability for businesses to identify at least the direct supplier of a product as well as the immediate client, with the exemption for retailers (European Commission, 2002, European Commission 2004). Each food business operator must record and preserve information such as (1) name, address of supplier, and type of products supplied, (2) name, address of customer, nature of the products delivered to the customer, and (3) date of the transaction/delivery (European Commission, 2002).

In Japan, the Government has supported the development of traceability systems from 2003 with the establishment of the Food Safety and Consumer Affairs Bureau within the Japanese Ministry of Agriculture, Forestry and Fisheries (MAFF). The MAFF policy is to encourage food business operators to voluntary establish traceability systems (MAFF 2004, 2007). The government has taken decisions to support the development of traceability systems. In 2003, the Food Safety and Consumer Affairs Bureau was established within the (MAFF). Although traceability systems are not legally required except for domestic beef, MAFF policy is to encourage food business operators to voluntarily establish traceability systems (MAFF 2004, 2007). Supporting this policy, MAFF has provided funds for projects such as developing traceability systems utilizing advanced ICT and formulating a handbook to guide the establishment of traceability systems. The handbook for the introduction of food traceability systems was created for food business operators and aims to facilitate cooperation between the various operators throughout the food chain (Revision Committee on the Handbook for Introduction of Food Traceability Systems 2007). The handbook covers definitions, basic objectives of traceability, the role that each operator should play to establish traceability, and how to proceed with the introduction of a traceability system. It outlines examples of general traceability systems as well as guidelines for specific food items. An English translation has been produced for overseas suppliers. In June 2003 the Japanese Govern introduced the Beef Traceability Law and in October 2010 enacted the Act on Recording Source Data and Other Information Relate to the Trade of Rice and Other Gains. In this act is reported that "The government has drafted a plan calling for a new law to establish the

traceability of all food products" (Summary of FY2010, 2008).

In China the first Food Safety Law becomes effective in July 2009, and it requests the food company to keep the account book of procurement and sale for at least two years to be reviewed by food safety authorities once needed. No mandatory regulation was in effect till that the General Administration of Quality Supervision, Inspection and Quarantine of the P. R. C. released a new regulatory to ask diary industry to adopt IT system to record critical information (AQSIQ of PRC, 2010). The Announcement No.119, 2010 might be treat as a trend on the food traceability system adoption.

In Canada, traceability initiatives were mostly oriented to animal identification and tracking, through the creation of the Canadian Cattle Identification Agency (CCIA).

In 2001, Québec was the first province implementing a traceability procedure for cattle, sheep, and pigs under Agri-Traçabilité Québec, which provides a framework for identification of animals and premises, as well as animal transportation tracking.

- In the Health of Animals Act, cattle, bison and sheep identification became federally regulated by 2004.
- In 2003, Agriculture and Agri-Foods Canada consulted with federal, provincial and territorial governments, where a consensus “that traceability is necessary in a safe food supply” was established; this was incorporated into the Agricultural Policy Framework (APF).
- Can-Trace, created in 2003, released the 2nd version of the Canadian Food Traceability Data Standard in 2006, based on the EAN.UCC system. Can-Trace is a collaborative, multi-commodity effort to establish traceability standards for all food products in Canada.

Participation in the Can-Trace is currently on a voluntary basis.

In the United States, after the Bioterrorism Act program regulation of 2002, local and foreign food businesses that produce food products for sale in the United States must be registered with the U.S. Food and Drug Administration (FDA). Importers and processors are required to keep records of their immediate suppliers and buyers for 2 years after transaction, and must be able to reproduce these records upon request for inspection by the FDA. In 2007, the FDA issued the Food Protection Plan (FPP), which objective is improving the food safety and defense for all domestic and imported products in the United States. A component of the FPP is the emergency response development, under which traceability practices are in the process of being defined, in collaboration with the food industry and other stakeholders. Recently, the FDA Food Safety Modernization Act (FSMA) signed as law on January 4th, 2011establishes within the FDA a tracing system able to receive information that improves the capacity to effectively and rapidly track and trace food in United States, or offered for importation into the United States.

In Australia, under the Legislative Instrument Act, the Australian New Zealand Food Standards Code was stipulated in 2003. Through this standard code the food businesses must be able to identify where their products come from (Diogo et al. 2004).

South Korea and Taiwan have included a definition of traceability in their food legislation and they have also implemented traceability programmes for some categories of domestic products, where participation of food operators in most of these programmes is voluntary.

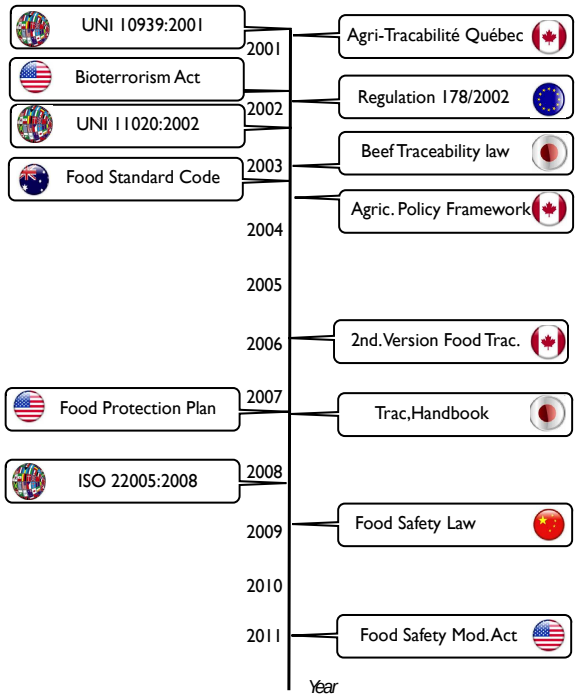


Figure 1- Food Traceability Regulatory Framework

An interesting finding by reviewing the history of food traceability regulations is that not only the approaches to establish the systems are different, but also the breadth, depth and precisions of these systems are different. While most chains allow only one step forward and one-back trace, a deeper traceability system back to the producer of raw materials is required e.g. for ensuring that products have not been genetically modified (Golan, Krissoff, and Kuchler 2004). Additional requirements should be satisfied to ensure food security and to improve food quality (Food Standards Agency 2002). Additional information should be collected in each stage of the supply chain in order to ensure the availability of data for the production analysis and optimization (Thompson, Sylvia, and Morrissey 2005). All trading partners in the supply chain must guarantee both the internal and external, or well-know supply chain, traceability.

Therefore, the no existence of a global standard for food traceability hinters the communication between the

different actors of the supply chain and hence the traceability chain. An important bottleneck for traceability is, in fact, the lack of a global standardization. Companies positioned in different geographical contexts (America, EU, Asia, Africa) have to deals with different implementation for the products responsibility and liability, and different standards for labeling. A global traceability system can be conveniently achieved if each company in the supply chain follows a common system for information encoding, registration and control, and the transactions between actors involved are regulated in a coherent and shared form (De Cindio et al. 2011).

3. SCIENTIFIC LITERATURE

In this section is presented a review of the scientific literature on food traceability, with a particular focus on the fruit and vegetable field.

The increased interest of the scientific world in the research area of the supply chain traceability is the result of a long series of developments aimed at improving food quality and safety management (Opara 2003).

3.1 Mathematical Models

The implementation of an effective shop floor traceability system does not only consist of recording, manually or using computers, the various supply chain batches. Indeed, it implies a deep modification of the organization and, sometimes, of the company fabrication processes. A traceability system must ensure the linkage between products and information, and must guarantee this connection through the supply chain. Moreover, as mentioned before, the maintenance of traceability is a complicated and expensive process especially with regard to processed foods. In case of processed foods different lots of various raw materials are combined into several production batches typically distributed through various points of sale (Hu et al. 2009). Some mathematical models have been proposed in order to solve this problem and to model the lot behavior.

Dupuy, Botta-Genoulaz, e Guinet (2005) propose a mathematical model to reduce batch dispersion, for controlling the mixing of production bathes in order to limit the size, and consequently the cost and the media impact of batches recalled in case of problem. The problem studied aims to minimize the quantity of products recalled in case of a problem that occurs in a particular situation with a 3-level ‘‘disassembling and assembling’’ bill of material. For the development of the mathematical model they implemented a model proposed in Dupuy, Botta-Genoulaz, e Guinet (2002), a method based on the concepts of Traceability Resource Unit (TRU) and batch dispersion.

Bollen, Riden, and Cox (2007) and Riden and Bollen (2007), studied and analyzed the traceability in fruit supply chains in order to improve the traceability control of the batches. They proposed a mixing model that was able to assign the probabilities of bin origin to individual fruit at the point that they are packed into their final packs. The model can significantly reduce fruit mixing and improve the traceability. They introduced concepts for quantify aspects of processing transformations, implementing a model based on enable simulations that examine the effect of splitting throughput into multiple output lines. They stated that there is a potential to implement high precision and fine granularity traceability in agricultural supply systems, which can also meet a number of other purposes such as improvement feedback to producers and benefits to supply system efficiency, as well as being acceptable for compliance purposes.

Hu et al. (2009) studied the traceable information flow and risk transmission throughout food supply which contains raw material, process and distribution. They propose a mathematical model based on dynamic programming in order to solve the risk transmission problem in a China dumpling factory, using Radio Frequency Identification (RFID) to identify and transfer traceable information. The purpose of the study in the factory is to minimize the cost due to a food safety crisis. If a food safety problem comes from a raw material batch, the factory will trace and identify all products. This model takes into consideration the previous research work of Dupuy, Botta-Genoulaz, e Guinet (2002). They propose a graphical model to describe the risk transference problem, according to the Gozinto graphs proposed by van Dorp 2003.

Tamayo et al. (2009) used the traceability information in order to reduce the size of products recalls. Three principal subjects are defined as follows: dispersion evaluation and optimization, criticality determination and final product delivery optimization. To achieve the final purpose of reducing the recall size and cost they stated that it is important to perform an intelligent delivery allocation. The developed expert system uses the information produced by a genetic algorithm and an artificial neural network to optimize product dispatches.

X. Wang et al. (2010) developed an integrated optimization model in which the product safety related traceability factor is incorporated with operations factors to develop an optimal production plan. The model aims to improve traceability and manufacturing performance by simultaneously optimizing the production batch size and batch dispersion with risk factors.

3.2 Information Models

In recent years, many works have been conducted on the development of traceability systems in food supply chain.

Jansen-Vullers, van Dorp, e Beulens (2003) proposed a reference-data model for tracking and tracing of food based on the Gozinto Graph, a tree-like graphical representation of raw materials parts, intermediates and subassemblies, in which a particular production process transforms an end-product through a sequence of operations. In the paper the tracking and

tracing requirements from three business cases situated in a production network (breeder, grower and egg producer) are discussed. A concise overview of the main requirements is identified for each business case and a data model is constructed. The development of the reference data model is described by explaining the model-part of the bill of lots and/or batches, the model-part of operations and variables and the integration of these two model-parts. The reference data model supports the registration of historic relations between lots and batches (where-from and where-used relations), the registration of operations on lots and batches in production, the registration of associated variables and values, on operation control, and the registration of capacity units on which operations are executed. The reference data model includes tracing of generating properties, which have been identified as an overlaying requirement.

Lo Bello et al. (2005) proposed a general approach based on distributed collaborative information systems where every company exchanges traceability data with the others over a network. XML was used as the format to represent data, for its ability to cope with data structures of different size. Web Services based technology has been adopted to interface different suppliers which communicate through HTTP protocol.

Regattieri, Gamberi, e Manzini (2007) develop a traceability system for Parmigiano Reggiano (the famous Italian cheese) introducing a general framework based on the integration of alphanumeric codes and RFID. The characteristics of a product are identified in its different aspects along the entire supply chain, from the bovine farm, the dairy, the seasoning warehouse, and lastly to the packaging factory. The complete supply chain of Parmigiano Reggiano is traced by an RFID system integrated with an alphanumerical code. Technically the system developed is based on a central database that collects data from bovine farms and from dairies. Manufacturers can check the progress made in production at any time and, if problems occur in the market place, they can re-trace the development of the portion of infected products and introduce effective re-call strategies.

Bechini et al. (2008) introduce a data model for identifying assets and actors and show a formal description of the lot behavior throughout the Supply Chain. The lot behaviour has been modelled by six activity patterns (integration, division, alteration, movement, acquisition and providing) using a UML activity diagram. The standard Unified Modelling Language (UML) notation is adopted to formally describe the different aspects of the modelled system. The model of a simply cheese supply chain with a UML communication diagram is presented. An independent, private data-sharing networks (PDSNs) is proposed as proper infrastructure for business process integration and Enterprise Service Bus (ESB) as architectural scheme for connecting third party applications. The ebXML Message Service (ebMS) is used for transporting business documents in a secure, reliable, and recoverable way in the inter-enterprise business collaboration scenario. In case that one of the business partners cannot manage ebMS messages (for instance, in the case of legacy systems), the communication is handled via ESB.

Thakur e Hurburgh (2009) developed a model for implementing internal traceability systems for a grain elevator that handles specialty grain and a model for information exchange among the supply chain actors. A UML sequence diagram shows the information exchange in the grain supply chain when a user requests additional information about a suspected product. The usage requirements of the traceability system are defined by the UML Use Case diagram technique. One of the most important goals of defining system requirements is to synchronize the requirements of all the actors. Integrated Definition Modeling (IDEF0) is used to develop the system for the internal traceability that they use and all the information is recorded in a RDBMS (Relational Database Management System) form by each actor. Finally some suitable technologies to enable this information exchange, such as the XML documents, are discussed. A relational database model to facilitate internal traceability at grain elevator is presented in Thakur el al. 2011. In this reference the entity-relationship modeling technique is used to develop the internal traceability grain handling a RDBMS for constructing and implementing the Entity Relationship model. The main purpose of the database is to connect the incoming grain lots with the outgoing grain lots. Once the data is stored in the database, the manipulation is accomplished through the use of queries written using the Structured Query Language (SQL).

Thakur and Donnelly (2010) presented a model for information capturing in the soybean supply chain. Actors involved in the supply chain are responsible for production, handling and processing. The soybean value chain and the main inputs and outputs of each stage are modeled using a simple flowchart. Conceptual process flow diagrams are created for farming, handling and processing sectors in the soybean value chain. Information capture points are identified for each sector and the corresponding products, processes and quality information to be captured are determined. A UML class diagram is developed for modeling products, processes, quality and transformed information. Finally some technologies available for transferring the information, such as the XML, are presented.

Maitri Thakur et al. (2011) presented a new methodology for modeling traceability information using the EPCIS framework and UML statecharts. EPCIS is an EPCglobal standard designed to enable EPC-related data sharing within and across enterprises. The model presented is used for mapping of food production processes in order to provide improved description and integration of traceability information. The method follows the approach of defining states and transitions in food production. A generic statecharts for food production is presented and applied to two supply chains: pelagic fish and grain. A state-transition model

with emphasis on identifying both traceability transitions and food safety and quality data are developed. The application of current EPCIS framework for managing food traceability information is presented by mapping the transitions identified in two product chains to the EPCIS events: Object Event and Aggregation Event. The corresponding states where the quality parameters are recorded are also identified and linked to these EPCIS events.

Bevilacqua, Ciarapica, e Giacchetta (2009) used the business process reengineering (BPR) approach to create a computer-based system for the management of the supply chain traceability information flows. They present a computer-based system for the traceability of fourth range of vegetables. They used an Event-Driven Process Chains (EPCs) technique to model the business processes. In order to ensure the traceability, each single unit or lot of the food products has been uniquely identified combining GTIN and the lot code. The business processes database follows the Entity Relationship Model (ERM). In the paper, moreover, the data model is not presented, and the front-and generated, based on the software ARIS, is only discussed.

Ruiz-Garcia, Steinberger, e Rothmund (2010) presented a web-based system to process, save and transfer data for tracking and tracing agricultural batch products along the SC. The development of the prototype involved the integration of several information technologies and protocols. The tracking system is based on a service-oriented architecture (SOA) and the communication is through messages in XML. Moreover, the work not deals with the problem of process and data modeling. In addition, there are only few authors using the BPMN standard for process modeling.

In the area of information modeling, several research works have been conducted on the analysis and evaluation of the different tools that can be used for recording, managing and transferring information such as barcode and Radio Frequency Identification (RFID) technologies. RFID systems have found applications in the agri-food sector especially in fresh-produce companies (Amador, Emond, and Nunes 2009; Gandino et al. 2009; Jedermann, Ruiz-Garcia, and Lang 2009; Martinez-Sala et al. 2009) and meat processing companies (Abad et al. 2009; LiWei, DongPing, and ChunHui 2009; Hsu, Chen, and Wang 2008; Reiners et al. 2009; Shanahan et al. 2009; Bo, Haiyan, and Caijiang 2008).

3.3 Future of Traceability Systems: Ontological Models

Currently there is a large variety of traceability mechanisms used in food supply chains. Some older traceability schemes are paper based while more recent are IT based. Notwithstanding, the information technologies can only be used to track information which were generated by product, management processes, manipulation operations. In fact, the information resources cannot be shared and reused in the process of tracing. A new research area is currently investigating how ontologies can be used to set up a traceability semantic model in order to reuse the information resources in the process of tracing and promote the accuracy and efficiency of information management. An ontology is an explicit specification of a conceptualization" (Gruber, T. R. 1993). The aim of ontology is to capture knowledge in related field, provide shared understanding to conceptual knowledge, definite common vocabulary in this field and give clear definition to the mutual relationship between these jargons and words from different levels of formal model (Heijst et al. 1995). Several ontologies have been proposed in the area of agriculture (Shoaib and Basharat 2010; Bansal and Malik 2011) and in particular in the vegetable supply chain domain (Yue et al. 2005) (Yue et al. 2005). Food ontologies are emerging related to nutritional concepts, such as the FOODS ontology (Snae and Bruckner 2008). The development process for a food ontology related to a specific health problem is shown in Cantais et al. (2005). This ontology is part of PIPS (Personalized Information Platform for Health and Life Sciences).

Kim, Fox, e Gruninger (1995) proposed the first quality ontology for products traceability and introduced two fundamental concepts, Traceable Resource Unit (TRU) and primitive activity. The data model presented by Kim et al. (1995) describes the main elements of the traceable ontology and underlines the relationships between terms. These concepts where revised by Moe (1998) which defined for each core entity a set of essential descriptors that must be included in order to secure ideal traceability of products and activities.

Yue et al. (2005) analysed the situation of China agricultural supply chain Informationization and discussed the use of Ontology in vegetable supply chain knowledge expressing. They put forward a process to build vegetable supply chain Ontology and gave a vegetable supply chain knowledge expressing frame. The frame can be used to express concepts and their relationships of vegetable supply chain as well as build vegetable supply chain knowledge base.

Salampasis et al. (2008) suggest a traceability solution which considers food traceability as a complex integration of business process problem which demands information sharing. To enable information sharing, data and the way they are organized should be standardized and their meaning and carrying semantics should be commonly agreed by the different operators along the food supply chain. To this end, they propose a generic framework for traceability applications which consists of three basic components: (i) an ontology management component based on OWL; (ii) an annotation component for "connecting" a traceable unit with traceability information using RDF; (iii) Traceability core services & applications.

Chifi, Salomie, and Chifu (2007) proposed an ontological model approach which allows semantic

annotation of Web Services aiming at automatic Web Services composition for food chain traceability. The model has been implemented in the framework of the Food-Trace project (Food Trace) for traceability in the domain of meat industry. The developed ontology is for the Romanian language. The model consists of core ontology and two categories of taxonomic trees: Business Service Description (BSD) tree and Business Product Description (BPD) trees. The core ontology defines six generic concepts: Business Actor, Service, Service Input, Service Output, Product and Feature.

Wang et al. (2012) proposed a quality and safety traceability system of fruit and vegetable products based on ontology. This work analyzed the whole process of fruit and vegetable products from farm to sale terminal, determined the collection scope of quality and safety traceability information of fruit and vegetables products, and established the traceability information ontology of the fruit and vegetables products. Through the definition of the traceability information ontology the authors introduced an example of traceability semantic net of control elements in the planting sub-system to explain the practical application process.

Bansal and Malik (2011) used the AGROVOC thesaurus as base vocabulary to develop the CROPont ontology. AGROVOC vocabulary developed by Food and Agriculture Organization is used for indexing and retrieving data in agricultural information systems. The idea is to develop ontology for crop production cycle that serves as a building block to an ontology driven Agriculture Information System Framework.

Yang et al. (2011) used the ontology theory and established an ontology knowledge base of food safety in emergencies domain which can provides semantic support for information retrieval in food safety domain. On the basis of this ontology knowledge base they developed an experimental semantic retrieval system FSSRS (Food Safety Semantic Retrieval System), trying to improve the retrieval performance and make up for the defect of traditional search methods.

4. DISCUSSION AND CONCLUSIONS

The main objective of this paper was to present a literature review on traceability models. Over the last decades, the development of traceability systems has received growing attention and several models have been developed. The authors have reviewed these models as they run through the literature by classifying them into mathematical and information models. The initial search identified a huge number of technical papers which were reduced according to content and quality.

The analysis of the state of the art highlights that several countries have developed traceability programs in the many sectors. However, more than often these programs do not provide information to the consumer about raw material management, processing**,** storage and distribution practices. Rather, traceability is viewed as a mechanism for improving food safety control by ensuring rapid product recall ability when a food safety incident occurs.

The scientific literature about the potential of traceability systems and the numerous advantages that can be obtained through its implementation is diverse (Moe 1998; Golan, Krissoff, and Kuchler 2004; Lo Bello, Mirabella, and Torrisi 2004; X. Wang and Li 2006; Pouliot and Sumner 2009; Nishantha, Wanniarachchige, and Jehan 2010).

Advantages of chain traceability and advantages of internal traceability in the production step are well discussed by Moe (1998). A traceability system can be used in order to guarantee food quality and safety and to improve the consumer trust (Lo Bello, Mirabella, and Torrisi 2004). Traceability, in fact, is becoming popular as a tool for winning consumer trust and managing complex supply chain while complying with ever increasing legal standards notwithstanding the nature of the industry (Nishantha, Wanniarachchige, and Jehan 2010). An efficient food traceability system can address information asymmetry in the supply chain, increase the speed of response to safety failures, and strengthen market and liability for precaution (Pouliot and Sumner 2009).

Tracing and tracking capabilities are crucial to confine the reaction to possible hazards and reduce the recovery cost (Bechini et al. 2005). Nevertheless, the recent food safety incidents (e.g. dioxins in animal feedstuffs in Belgium, E. Coli in Germany) have demonstrated that traceability is a “buzz word” with regard to food. Traceability systems have been shown to be weak or absent and hence slow or unable to assure consumers of food safety. In such case, food recalls or warnings have been applied to all suppliers, even to the supplier of products that do not contribute to the contamination

In addition, even if today a variety of lot code markings and systems exists for products identification and these have merit, they do not link across the life cycle of the world’s food supply.

The analysis of the discussed model highlight that the degree of coordination between the different actors of the supply chain is fundamental in the implementation of a traceability system. Also Álvarez et al. (2006) state that particular importance must be devoted to the degree of coordination between buyers and suppliers.

New traceability systems can be developed integrating the advantages of the previous works, in order to obtain a better solution at lower cost.

ACKNOWLEDGMENTS

This paper is co-funded with the support of the European Commission, European Social Fund and Regione Calabria (Italy).

REFERENCES

Abad, E., F. Palacio, M. Nuin, A. Zárate, A. Juarros, J. M. Gómez, and S. Marco. 2009. “RFID Smart Tag for Traceability and Cold Chain

Monitoring of Foods: Demonstration in an Intercontinental Fresh Fish Logistic Chain." *Journal of Food Engineering* 93 (4): 394–399.

Álvarez Gil, M. J., J. A. Alfaro, and L. A. Rábade. 2006. "Buyer-supplier Relationships Influence on Traceability Implementation in the Vegetable Industry." *Nº.: UC3M Working Papers. Bussiness Economics 2006-02*.

Amador, Cecilia, Jean-Pierre Emond, and Maria Nunes. 2009. "Application of RFID Technologies in the Temperature Mapping of the Pineapple Supply Chain." *Sensing and Instrumentation for Food Quality and Safety* 3 (1): 26–33. .

AQSIQ of P.R.C. Announcement No. 119, 2010 http://spscjgs.aqsiq.gov.cnlscxk/spscxk/201102/t20110209_177012.htm

Bansal, Nishu, and Sanjay Kumar Malik. 2011. "A Framework for Agriculture Ontology Development in Semantic Web." In , 283–286. IEEE. doi:10.1109/CSNT.2011.68.

Bechini, A., M. G. C. A. Cimino, F. Marcelloni, and A. Tomasi. 2008. "Patterns and Technologies for Enabling Supply Chain Traceability Through Collaborative E-business." *Information and Software Technology* 50 (4): 342–359.

Bechini, M. G. C. A. Cimino, B. Lazzerini, F. Marcelloni, and A. Tomasi. 2005. "A General Framework for Food Traceability." *2005 Symposium on Applications and the Internet Workshops SAINT 2005 Workshops*: 366–369.

Lo Bello, L., O. Mirabella, and N. Torrisi. 2004. "Modelling and Evaluating Traceability Systems in Food Manufacturing Chains." In *Enabling Technologies: Infrastructure for Collaborative Enterprises, 2004. WET ICE 2004. 13th IEEE International Workshops On*, 173–179.

———. 2005. "A General Approach to Model Traceability Systems in Food Manufacturing Chains." In *Emerging Technologies and Factory Automation, 2005. ETFA 2005. 10th IEEE Conference On*, 2:8–pp.

Bevilacqua, M., F. E. Ciarapica, and G. Giacchetta. 2009. "Business Process Reengineering of a Supply Chain and a Traceability System: A Case Study." *Journal of Food Engineering* 93 (1): 13–22.

Bo, Yan, Fu Haiyan, and Zhang Caijiang. 2008. "Application of RFID Technology in Meat Circulation Management." In *Control Conference, 2008. CCC 2008. 27th Chinese*, 808 –812.

Bollen, A.F., C.P. Riden, and N.R. Cox. 2007. "Agricultural Supply System Traceability, Part I: Role of Packing Procedures and Effects of Fruit Mixing." *Biosystems Engineering* 98 (4): 391–400.

Cantais, J., D. Dominguez, V. Gigante, L. Laera and V. Tamma, "An example of food ontology for diabetes control", Galway, Ireland, in *Proceedings of the ISWC Workshop on Ontology Patterns for the SemanticWeb, 2005.*

Chifi, V. R., I. Salomie, and E. S. Chifu. 2007. "Ontology-enhanced Description of Traceability Services." In *Intelligent Computer Communication and Processing, 2007 IEEE International Conference On*, 1–8.

Diogo M. S. , Julie A. C. The Economics of Implementing Traceability in Beef Supply Chains: Trends in Major Producing and Trading Countries. *Working Paper*, 2004

Dupuy, C., V. Botta-genoulaz, and A. Guinet. 2002. "Traceability Analysis and Optimization Method in Food Industry." In *Systems, Man and Cybernetics, 2002 IEEE International Conference On*, 1:494 –499.

Dupuy, C., V. Botta-Genoulaz, and A. Guinet. 2005. "Batch Dispersion Model to Optimise Traceability in Food Industry." *Journal of Food Engineering* 70 (3): 333–339.

European Commission. 2002. "REGULATION EC No 178_2002 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 28 January 2002 Laying down the General Principles and Requirements of Food Law, Establishing the European Food Safety Authority and Laying down Procedures in Matters of Food Safety". Official Journal of the European Communities.

———. 2004. "GUIDANCE ON THE IMPLEMENTATION OF ARTICLES 11, 12, 16, 17, 18, 19 AND 20 OF REGULATION (EC) N° 178/2002 ON GENERAL FOOD LAW: CONCLUSIONS OF THE STANDING COMMITTEE ON THE FOOD CHAIN AND ANIMAL HEALTH."

Food Standards Agency. 2002. "Traceability in the Food Chain: A Preliminary Study." Available at: http://www.food.gov.uk/multimedia/pdfs/traceabilityinthefoodchain.pdf.

FSMA, U.S. FDA FOOD SAFETY MODERNIZATION ACT Available at: http://www.gpo.gov/fdsys/pkg/BILLS-111hr2751enr/pdf/BILLS-111hr2751enr.pdf

FOOD TRACE, http://www.coned.utcluj.ro/FoodTrace/

Gandino, F., B. Montrucchio, M. Rebaudengo, and E. R. Sanchez. 2009. "On Improving Automation by Integrating RFID in the Traceability Management of the Agri-food Sector." *Industrial Electronics, IEEE Transactions On* 56 (7): 2357–2365.

Golan, E., B. Krissoff, and F. Kuchler. 2004. "Food Traceability: One Ingredient in a Safe and Efficient Food Supply."

Gruber, T. R. 1993. "A Translation Approach to Portable Ontology Specifications." *Knowledge Acquisition* 5 (2): 199–220.

Heijst, Gertjan van, Sabina Falasconi, Ameen Abu-hanna, Guus Schreiber, and Mario Stefanelli.

1995. *A Case Study in Ontology Library Construction*.

Hsu, Yu-Chia, An-Pin Chen, and Chun-Hung Wang. 2008. "A RFID-enabled Traceability System for the Supply Chain of Live Fish." In *IEEE International Conference on Automation and Logistics, 2008. ICAL 2008*, 81 –86.

Hu, Zhang, Zhang Jian, Shen Ping, Zhang Xiaoshuan, and Mu Weisong. 2009. "Modeling Method of Traceability System Based on Information Flow in Meat Food Supply Chain." *WSEAS TRANSACTIONS on INFORMATION SCIENCE and APPLICATIONS*, July.

Jansen-Vullers, M.H, C.A van Dorp, and A.J.M Beulens. 2003. "Managing Traceability Information in Manufacture." *International Journal of Information Management* 23 (5): 395–413.

Jedermann, Reiner, Luis Ruiz-Garcia, and Walter Lang. 2009. "Spatial Temperature Profiling by Semi-passive RFID Loggers for Perishable Food Transportation." *Comput. Electron. Agric.* 65 (2) (March): 145–154.

Kim, Henry M., Mark S. Fox, and Michael Gruninger. 1995. "An Ontology of Quality for Enterprise Modelling." In *Proceedings of the 4th Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET-ICE'95)*, 105–. WET-ICE '95. Washington, DC, USA: IEEE Computer Society.

LiWei, Geng, Qian DongPing, and Zhao ChunHui. 2009. "Cow identification technology system based on radio frequency." *Transactions of the Chinese Society of Agricultural Engineering* 25 (5): 137–141.

Ministry of Agriculture, Forestry and Fisheries (MAFF), What You Need to Know about Traceability Systems, Tokyo: MAFF, 2004

Ministry of Agriculture, Forestry and Fisheries (MAFF), Traceability as Common Sense!, Tokyo: MAFF, 2007

Martinez-Sala, Alejandro S., Esteban Egea-Lopez, Felipe Garcia-Sanchez, and Joan Garcia-Haro. 2009. "Tracking of Returnable Packaging and Transport Units with Active RFID in the Grocery Supply Chain." *Computers in Industry*.

Moe, T. 1998. "Perspectives on Traceability in Food Manufacture." *Trends in Food Science & Technology* 9 (5): 211–214.

Nishantha, G. G. D., M. K. Wanniarachchige, and S. N. Jehan. 2010. "A Pragmatic Approach to Traceability in Food Supply Chains." In *Advanced Communication Technology (ICACT), 2010 The 12th International Conference On*, 2:1445–1450.

Opara, L. U. 2003. "Traceability in Agriculture and Food Supply Chain: a Review of Basic Concepts, Technological Implications, and Future Prospects." *JOURNAL OF FOOD AGRICULTURE AND ENVIRONMENT* 1: 101–106.

Pouliot, Sebastien, and Daniel A. Sumner. 2009. "Traceability, Food Safety and Industry Reputation."

Regattieri, A., M. Gamberi, and R. Manzini. 2007. "Traceability of Food Products: General Framework and Experimental Evidence." *Journal of Food Engineering* 81 (2): 347–356.

Reiners, Kerstin, Alexander Hegger, Engel F. Hessel, Stephan Böck, Georg Wendl, and Herman F.A. Van den Weghe. 2009. "Application of RFID Technology Using Passive HF Transponders for the Individual Identification of Weaned Piglets at the Feed Trough." *Computers and Electronics in Agriculture* 68 (2) (October): 178–184.

Revision Committee on the Hanbook for Introduction of Food Traceability Systems. 2007. "Handbook for Introduction of Food Traceability Systems (Guidelines for Food Traceability)."

Riden, C.P., and A.F. Bollen. 2007. "Agricultural Supply System Traceability, Part II: Implications of Packhouse Processing Transformations." *Biosystems Engineering* 98 (4): 401–410.

Ruiz-Garcia, L., G. Steinberger, and M. Rothmund. 2010. "A Model and Prototype Implementation for Tracking and Tracing Agricultural Batch Products Along the Food Chain." *Food Control* 21 (2): 112–121.

Salampasis, Michail, Eleni P. Kalogianni, Sezen Ocak, Theodoros Lantzos, and Christos Kouroupetroglo. 2008. "Demands and Technology Pushes for Building ICT Infrastructure for Food Traceability." In *AWICTSAE 2008*. Alexandroupolis.

Schwägele, F. 2005. "Traceability from a European Perspective." *Meat Science* 71 (1): 164–173.

Shanahan, C., B. Kernan, G. Ayalew, K. McDonnell, F. Butler, and S. Ward. 2009. "A Framework for Beef Traceability from Farm to Slaughter Using Global Standards: An Irish Perspective." *Computers and Electronics in Agriculture* 66 (1) (April): 62–69.

Shoaib, Muhammad, and Amna Basharat. 2010. "Semantic Web Based Integrated Agriculture Information Framework." In , 285–289. IEEE.

Snae, C., and M. Bruckner. 2008. "FOODS: A Food-Oriented Ontology-Driven System." In *Digital Ecosystems and Technologies, 2008. DEST 2008. 2nd IEEE International Conference On*, 168–176.

Summary of FY2010: Measures for Food, Agriculture and Rural Areas, Ministry of Agriculture, Fishery and Food, Japan, 2008

Tamayo, Simon, Thibaud Monteiro, and Nathalie Sauer. 2009. "Deliveries Optimization by Exploiting Production Traceability Information."

Engineering Applications of Artificial Intelligence 22 (4-5) (June): 557–568.

Thakur, M., and K. A. M. Donnelly. 2010. "Modeling Traceability Information in Soybean Value Chains." *Journal of Food Engineering* 99 (1): 98–105.

Thakur, M., and C. R. Hurburgh. 2009. "Framework for Implementing Traceability System in the Bulk Grain Supply Chain." *Journal of Food Engineering* 95 (4): 617–626.

Thakur, M., B. J. Martens, and C. R. Hurburgh. 2011. "Data Modeling to Facilitate Internal Traceability at a Grain Elevator." *Computers and Electronics in Agriculture*.

Thakur, Maitri, Carl-Fredrik Sørensen, Finn Olav Bjørnson, Eskil Forås, and Charles R. Hurburgh. 2011. "Managing Food Traceability Information Using EPCIS Framework." *Journal of Food Engineering* 103 (4): 417–433.

Thompson, M., G. Sylvia, and M. T. Morrissey. 2005. "Seafood Traceability in the United States: Current Trends, System Design, and Potential Applications." *Comprehensive Reviews in Food Science and Food Safety* 4 (1): 1–7.

United Nations Food and Agriculture Organization (FAO). 1999. "Codex Alimentarius Commission Procedural Manual (11th Ed.)."

Wang, F., R. Cao, W. Ding, H. Qian, and Y. Gao. 2011. "Incentives to Enable Food Traceability and Its Implication on Food Traceability System Design." In *Service Operations, Logistics, and Informatics (SOLI), 2011 IEEE International Conference On*, 32–37.

Wang, X., and D. Li. 2006. "Value Added on Food Traceability: a Supply Chain Management Approach." In *Service Operations and Logistics, and Informatics, 2006. SOLI'06. IEEE International Conference On*, 493–498.

Wang, X., D. Li, C. O'brien, and Y. Li. 2010. "A Production Planning Model to Reduce Risk and Improve Operations Management." *International Journal of Production Economics* 124 (2) (April): 463–474.

Wang, Yongfeng, Yu Yang, and Yongming Gu. "Research on Quality and Safety Traceability System of Fruit and Vegetable Products Based on Ontology."

Yang, Y., J. Du, and M. Liang. 2011. "Study on Food Safety Semantic Retrieval System Based on Domain Ontology." In *Cloud Computing and Intelligence Systems (CCIS), 2011 IEEE International Conference On*, 40–44.

Yue, J., H. Wen, X. Zhang, and Z. Fu. 2005. "Ontology Based Vegetable Supply Chain Knowledge Expressing." In *Semantics, Knowledge and Grid, 2005. SKG'05. First International Conference On*, 130–130.

ANALYSIS OF AIRPORT CHECK-IN COUNTER ALLOCATION POLICIES USING SIMULATION

Özden Onur Dalgiç[a], Yusuf Seçerdin[b], Gizem Sultan Nemutlu[c], Nilgün Fescioglu Ünver[d]

[a] [b] [c] [d] TOBB University of Economics and Technology, Department of Industrial Engineering, Ankara, 06560, TURKEY

[a]oodalgic@etu.edu.tr, [b]yusufsecerdin@etu.edu.tr, [c]gnemutlu@etu.edu.tr, [d]nfunver@etu.edu.tr

ABSTRACT

This study tackles the check-in counter allocation problem of Ataturk International Airport. Check-in process is required for all passengers and has to be completed 30 minutes before the flight time. Resources of this operation are check-in counters which are allocated to airline firms. This study analyses two different check-in counter allocation policies. The first policy allocates a fixed number of counters for a fixed duration of time to each flight. Second policy is the dynamic allocation policy - which allows airline firms to rent new counters or release counters dynamically by considering the time left until flight and the number of passengers that have not checked in yet. We used simulation to analyze the effects of these policies on passenger waiting times and counter utilizations. Results showed that dynamic policy reduce the waiting times by using counters in a more efficient way.

Keywords: check-in counter allocation, airport resource allocation, resource allocation policy evaluation, airport simulation

1. INTRODUCTION

The check-in counter allocation problem is different from other resource allocation problems in terms of uncertainty about the amount of resources required to meet the demand. The check-in process is stochastic due to passenger arrival, and the number of required check-in counters varies with time since the total number of passengers per flight differs over time (Chun and Mak 1999). System requirements change dynamically with respect to the time and date of the flights. Considering this complexity, accurate prediction of the resource requirements is nearly impossible in a non-computerized way.

There are some variations of the problem in the practice. These differences arise from passengers' class, check-in policy and the flight type. The passengers can be divided into categories such as economy and first class, and they are serviced independently. Different check-in process policies include completely restricted system, common check-in system and composite check-in system (Lee and Longton 1959). Furthermore, there is some performance standards related to check-in process specified by the airport management. The airline companies have to satisfy these requirements along with providing more qualified service for customers, to survive in highly competitive market.

In this study, we considered Istanbul Atatürk International Airport and focused on a side of a specific counter block (island) which consists of 16 counters. There are 10 flights assigned to these counters in a given day. The check-in process policy is a completely restricted system for these counters.

The outline of this paper is as follows. First, related literature is expressed in Section 2. Section 3 describes the simulation model and alternative scenarios. Results and output analysis are presented in Section 4. Section 5 presents the conclusion and future work.

2. LITERATURE

Check-in counter allocation and check-in process problems are mostly approached in the literature with three methods: pure deterministic (operational research) method, deterministic method with the integration of stochastic nature of problem, pure stochastic approach (simulation).

Lee and Longton (1959) tried to determine the best type of check-in system among complete restricted, common check-in and composite check-in system. Indeed, the problem researchers handled is determining the optimal transfer time, and determining number of check-in clerks required for that transfer time in a composite check-in system. Wai Chun and Tak Mak (1999) present an integrated system which consists of an intelligent resource simulation system and check-in counter allocation system. Airlines enter their seasonal check-in counter requests through the database forms input facility. The authorities of airport use the intelligent resource simulation system to predict the minimum number of check-in counters needed to satisfy the predefined service levels of airline. The constraint-based scheduling system will then use both the simulated and requested values to perform check-in counter allocation. Dijk and Sluis (2006) consider check-in process as a two stage problem. Stages are the

stochastic stage which designates the number of check-in desks to meet a determined service level for each individual flight by simulation. Second stage is a scheduling and optimal capacity allocation stage in which the minimum total number of desks and desk hours required over all periods are optimized by integer programming under the realistic constraints. Parlar and Sharafali (2008) developed a stochastic dynamic programming model to determine the optimal number of counters.

Appelt et al. (2007) identified the delays in the check-in process and created scenarios to improve the efficiency of the check-in procedure. In the simulation model, three basic scenarios and three combinations of them were created to analyze whether the delay in system differ according to check-in modes serviced to passengers by the airline. Joustra and Dijk (2001) denoted that traditional queuing models are very limited for check-in process and simulation is a necessary tool to evaluate it in a more valid way. In their study they used a check-in simulation toolbox to evaluate the effects of various factors in check-in process. Takakuwa and Oyama (2003) built a simulation model to examine the passenger flow. The number of passengers who miss their flights is aimed to be reduced by considering the issues that affect the waiting times. A special-purpose data generator is designed to determine the experimental data to execute the simulation. Park and Ahn (2003) developed an effective model that determines the number of check-in counters to be opened and duration that these counters should remain operating. They conduct a survey to determine the time when passengers arrive at airport before their flight departure times. After that, they used passengers' arrival distributions to find the appropriate number of check-in counters and duration of operation. They also take the time of the flight into account - whether it is in peak hours or on a special day.

3. SIMULATION MODEL

3.1. System and Assumptions

Airport check-in process has stochastic and deterministic components. The stochastic nature of passenger arrivals and variable check-in process times requires analyzing the problem by simulation modeling. In the airport, the flight features and the initial allocation of check-in counters are deterministic. The airport management is informed about the flight schedules by the airline firms and allocates the current counters among the firms such that all counters allocated for a flight are adjacent. Considering this information in the simulation model, first we selected 10 flights randomly, and assigned a specific counter block to these flights in a given day. The counter opening time and the total duration of check-in process for each flight are assigned to the counters.

Arrival profiles of the passengers are modeled according to two different flights observed in the airport. One of these flights intensively consists of Turkish passengers (arrival profile-1); the other one intensively consists of passengers which are foreign tourists (arrival profile-2). The arrival profiles of these two flights are denoted in Figure 1 and Figure 2 respectively. The durations that the counters of these flights are open, are divided in 8 equal periods with a starting period (period 0). The length of period 0 is equal to 30 minutes and in this period the passengers who show up before counters are open, arrive. This period starts exactly 30 minutes before counters are open. Figure 1 shows the arriving passenger proportions which are calculated by: Number of passengers arrived period i/Total number of passengers. The proportions of all periods are determined according data obtained from the observations.

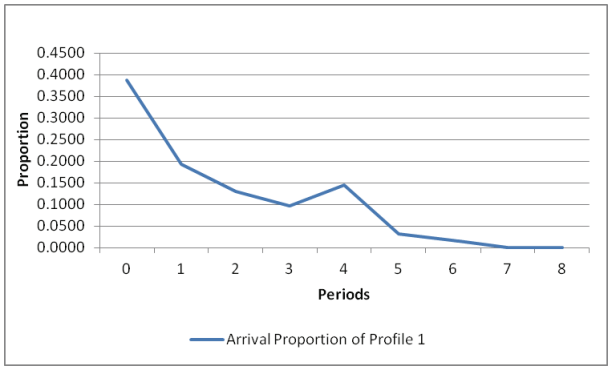


Figure 1: Arrival Profile 1

Since we are conducting the study for Istanbul Ataturk International airport, we assumed that arrival profile of the passengers for 60 percent of all flights fit to arrival profile-1 and the passenger profiles for the remaining flights fit to arrival profile-2. The arrival proportions used in the simulation model for each flight are denoted in Table 1.

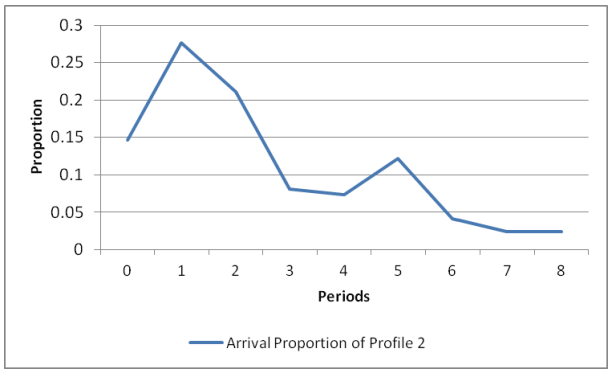


Figure 2: Arrival Profile 2

Table 1: The Passenger Proportions per Flight

Periods	Flight Numbers	
	1-3-5-7-9-10	2-4-6-8
0	0.39	0.15
1	0.19	0.28
2	0.13	0.21
3	0.10	0.08
4	0.15	0.07
5	0.03	0.12
6	0.02	0.04
7	0.00	0.02
8	0.00	0.02

Table 2 shows the total number of passengers for each flight. We used triangular distribution to express the total passenger count in each simulation run, as the number of passenger checking-in can change from time to time and planes are not always flying at full capacity.

Table 2: Number of Passengers for Each Flights

Flight	# of Passengers
1	TRIA(272,303,333)
2	TRIA(360,400,440)
3	TRIA(79,88,96)
4	TRIA(85,95,104)
5	TRIA(167,186,204)
6	TRIA(31,35,38)
7	TRIA(110,123,135)
8	TRIA(93,104,114)
9	TRIA(354,394,433)
10	TRIA(192,214,235)

Moreover, check-in process times differ according to the flight profiles. This difference can be caused by either the passenger profiles or the check-in staff or both. To determine the probability distributions of the check-in processing times of these two flights, the processing times were collected per passenger for each of the flights. After collection of the data, the distributions were determined for each of the flight profiles by using ARENA Input Analyzer (see Table 3).

Table 3: Check-in Processing Times

Profiles	Processing Time (in seconds)
Arrival Profile 1	24+WEIB(91, 1.05)
Arrival Profile 2	18+WEIB(134, 1.24)

3.2. Objectives and Performance Measures

The objective of this study is to analyze and compare different check-in counter allocation policies. The revenue of the airport management will increase when more counters are allocated to airline firms through an efficient allocation policy. The goal is to decrease the waiting times of passengers by increasing utilization of the counters without changing the total number of available counters.

We used the following performance measures:

- Average counter usage: This performance measure indicates average counter usage ratio for all counters. Counter usage ratio equals the total time a counter is open divided by simulation time.
- Average waiting time of passenger in queue.

3.3. Model Development and Alternative Scenarios

We developed a different simulation model for each policy, using ARENA 12.0 simulation software package. The first simulation model belongs to the base case scenario which is the fixed counter allocation policy in a 16 counters-10 flights airport day. The second model analyses the dynamic counter allocation policy in the same 16 counters-10 flights simulation day.

3.3.1. Base Case and Simulation Model

This study represents the currently used check-in system as the "base case". The check-in counters are pre-assigned to airline firms by the airport management based on a schedule. These counters are called as startup counters after this point in this paper. The counters are opened at the beginning time and serve for a specified time. Passengers arrive at the check-in service area and wait in the queue if all counters are busy. At the end of the service time the counters are closed. A basic representation of the check-in process system is given in Figure 1.

Simulation model of the base case is quite simple. Passengers are the only entities of this model. Passengers belonging to a specific flight start arriving at the check-in service area when the counters of that flight are opened. They leave the system when their check-in operation is completed. Resources of check-in process are counters in the system. A passenger's arrival at the counter, beginning and completion of check-in process constitute the events of base case model.

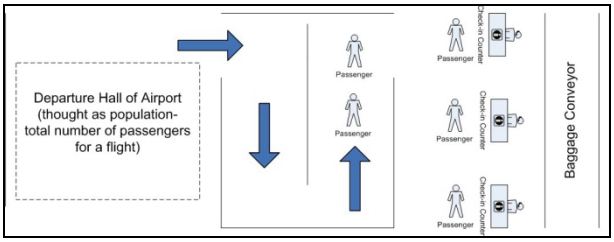


Figure 3: Basic Representation of the Check-in Process System

3.3.2. Dynamic Case and Simulation Model

In this model, the main assumption is that the number of counters which are used for any flight can be changed dynamically while the check-in process is continuing. The change can be made through the use of startup counters of other flights whose check-in operations have not started yet, or releasing some of the counters

which are currently used. By this way, the utilization of the counters which are defined as resources of the system will be improved. Dynamic change is depended on a value called k which is calculated periodically, as in

$$k = \left\lceil \frac{\text{Expected check-in process time*Remaining number of passengers to be served}}{\text{Remaining time}} \right\rceil$$

The dynamic allocation policy and the terms used in the policy are explained in detail below. Figure 2 illustrates the counter block and corresponding counter positions.

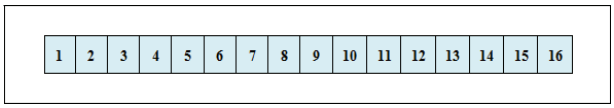


Figure 4: Illustration of Counter Block and Counter Positions

Definition 1 "Available counter" *is a counter that*;

i. In counter releasing operation, has the smallest ID or the largest ID of the dealing flight, edge of adjacent counters (resources) for any flight (counter 1 and 4 for flight 1, and counter 7 and 10 for flight 2 in Figure 3),

ii. In counter seizing operation, is adjacent to edge counters and will not be used for another flight not started yet for next 30 minutes (counter 5 for flight 1, and counter 6 and 11 for flight 2 in Figure 3).

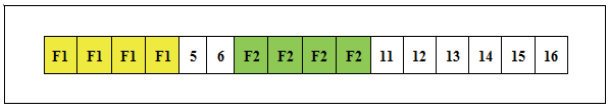


Figure 5: Illustration of Available Counters

Definition 2 "Additional counter" *refers to an available counter when a seizing operation is completed. This seized counter is an additional resource to startup counters for corresponding flight (counter 6 is seized by flight 2 in Figure 4).*

Definition 3 "Current flight" *is the flight that is analyzed through k value for seizing and releasing operations (flight 2 in Figure 4).*

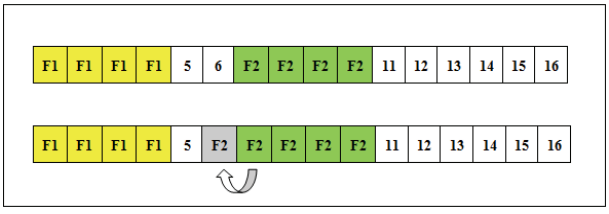


Figure 6: Illustration of a Seizing Operation

Definition 4 "neighbor flight" *refers to the flight which serves at the counters adjacent to the counters of current flight while check-in process of current flight continues (flight 1 is neighbor of flight 2, changing counters between them occurs as flight 2 releases counter 5 first, then flight 1 seize that counter in Figure 5)* .

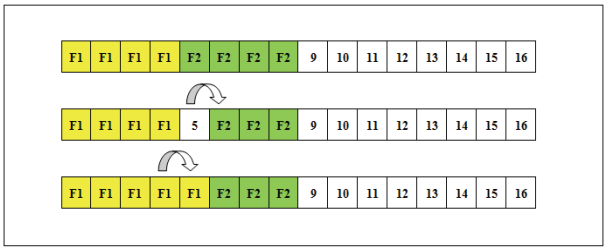


Figure 7: Illustration of Changing Counters between Flights

The seizing and releasing operations are made on available counter/s. The case in which the flight has only one counter, it will remain open until the end of check-in process regardless of the k value.

The mechanism of the dynamic counter allocation policy is as follows:

Phase 1: Each flight is initially allocated a number of startup counters (number and place of counters decided by airport management). The starting and ending time of the check-in operation for each flight refer to the opening and closing time of the startup counters respectively. The number of passengers of each flight, which expresses the limited population for the check-in operation, is assigned as an attribute to this flight. All the data is set into the model are summarized in Table 4.

Phase 2: Once the counters for a flight open, the k value is calculated every fifteen minutes. The decision for seizing or releasing a counter is made according to the k-value. In order to seize an additional counter, there must be at least one available counter.

Phase 2.1: If we decide to seize a counter, it is important to check at what time this additional counter will be opened as a startup counter of another flight. The starting time of next flight assigned to this additional counter must be taken into account and the operation of current flight at this counter must be terminated 30 minutes before flight starting time. If this time minus 30 minutes is greater than the ending time of the current flight check-in process, this counter can be used until the ending time (Simulation model is presented in Figure 8 in Appendix A).

Phase 2.2: In the case of releasing a counter, the counter selected should be suitable for the neighbor flights to seize. According to the k value, if there is need for one more counter, the released counter can be seized again if it is still an available counter (Simulation model is presented in Figure 8 in Appendix A).

Phase 3: When the check-in operation is completed for a flight, all the counters used for that flight are closed.

Table 4: Input Data for Simulation Model

Flight	Duration of Check-in	Starting and Ending Time (in minutes)	# of Startup Counters	Startup Counters
1	130	685-815	5	12,13,14,15,16
2	160	635-795	6	2,3,4,5,6,7
3	95	360-455	2	11,12
4	105	585-690	3	9,10,11
5	100	840-940	4	13,14,15,16
6	85	200-285	1	15
7	105	525-630	3	13,14,15
8	125	815-940	2	9,10
9	165	965-1130	9	1,2,3,4,5,6,7,8,9
10	120	800-920	4	2,3,4,5

3.4. Verification and Validation

Both of the simulation models are verified by using TRACE element in the ARENA 12.0 software. Trace is one of the most powerful techniques to verify a simulation model (Law 2007). TRACE element provides a list which keeps events happened, state variables, some other statistics about the system during the simulation. Thus, potential mistakes would exist in the simulation model can be detected by model developer. While development process, output of the TRACE element was checked and it can be seen that the both of simulation models operate as intended.

Although we used real input data about the flights and startup counters, we could not obtain any data about passenger waiting times in the base case therefore we could not use statistical validation tools. Instead we compared the system behavior with the simulation results and verified face validity of the model.

4. RESULTS AND OUTPUT ANALYSIS

In this section we compare the two scenarios in terms of average counter usage and average waiting time of passengers. A single run for each scenario starts when the counter opens for the first flight and ends when all the day's flights finish their check-in processes. The number of replications is determined using the relative error method (relative error of 5 percent). Thus, 20 replications are enough for both performance measures while simulation models are run for two hundreds replications to narrow the confidence intervals of the difference between the performances of the models. Table 5 presents the results for each simulation models with respect to performance measures.

Table 5: Results-Means and Half-widths

	Base Case	Dynamic Case
Waiting Times	28.67±2.9104	26.02±2.8130
Counter Utilization Rate	0.2713±0	0.3284±0.0114

Paired t-test is used to compare the policies and determine which policy is better in terms of the performance measures (Law 2007). Table 6 presents the 90% confidence intervals for the difference between the scenarios. Confidence intervals show that the waiting times in the check-in process of the dynamic system is shorter than that of the current system, in other words the proposed system is resulted statistically significant improvement of waiting times.

Results show that counter usage rates increased when dynamic allocation policy is used. In the current system, it is necessary to use additional resources for decreasing the queue length. To do this, the airline firms want to increase the number of counters they demand. However, the resources are limited and it can be impossible to allocate the counters to the airline firms based on their demands instead of their needs.

The illustrations below explain the differences between the base case and the dynamic case more clearly. In Figure 8, the counter labeled as C3 is not assigned to any flight. In the base case, this counter can be initially assigned to flight1 (F1), flight2 (F2), or none. If this counter is assigned to any of the flights, it can remain idle in the some portion of the time while it is open. If it not assigned to any of the flights, it cannot be utilized when flights need more counters.

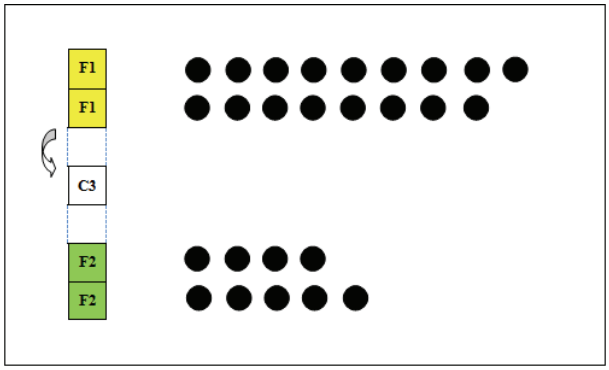


Figure 8: Illustration of Proposed System Procedure

In the proposed system, counter 3 (C3) can be utilized by both of the flights (one flight at a time) when one of them needs additional resources to minimize passenger waiting times. For instance, counter 3 is seized by flight 1 in the situation represented in Figure 8 as the queue length of flight 1 exceeded the k value. The seized counter will be released when the queue length of flight 1 fall below the k value as in the situation represented in Figure 9. This releasing operation also enables flight 2 to seize the same counter to minimize passenger waiting times since the queue length of flight 2 exceeds the limit now. Through the proposed method, flight 1 and 2 which are neighbor flights to counter 3 can reduce the passenger waiting times by seizing the additional available counter. In addition to this, utilization of counter 3 is increased by serving to neighbor flights when the number of passenger waiting in the queues increase.

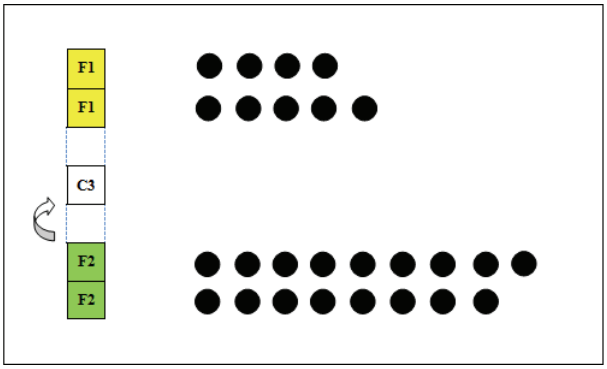


Figure 9: Illustration of Proposed System Procedure

In the proposed system, the number of counters is dynamically changed and as a result, counter usage rate increased. As it pointed out before, the proposed system also decreases the waiting times which means that the profit of airline firms may be increased through a positive impression on passengers. Additionally, since for a country international airports are the gates opening abroad, the quality of service is very important. The waiting times in the check-in process queue of passengers who travel overseas should be decreased even if an additional cost is bore by either airline firms or airport management. Moreover, the less time spend waiting in the check-in queue will result in more time spent in the duty free area which is a very important source of revenue for the airports.

Table 6: Confidence Intervals for Performance Measures

	Lower Limit	Mean	Upper Limit
Waiting Time (in minutes)	0.6008	2.6509	4.7010
Counter Utilization Rate	-0.0685	-0.0572	-0.0458

5. CONCLUSION AND FUTURE WORK

In this study, we considered the check-in counter allocation problem. Target of the airport management is to increase the counter usage of each flight which means that more counters are provided for each flight. Thus, airline firms are able to decrease the waiting times of the passengers by using more counters. By taking this into account, we compared two different counter allocation policies. In the first policy each flight uses a constant number of counters during the check-in process (base case). This policy is currently used in many airports. The other policy is one that the number of counters can change dynamically during the check-in process (dynamic case). Simulation models and scenarios are developed for these two policies. Results show that the dynamic case is better than the base case in terms of passenger waiting times.

As a future study we are planning to investigate the case where passengers are not of the same type. As it was mentioned before, all the passengers are assumed to be homogenous (apart from the arrival profile difference) in this study. However, the luggage sizes and the number of luggages per passenger can differ for each flight with respect to the distance of flight in reality and as a result, the check-in process time of a flight will be affected. Furthermore, it is known that there are differences in arrival behavior of passengers due to cultural differences. This study included two types of passenger profiles but in an international airport we may have more than two cultural profiles. The cultural features of passengers should also be considered in future studies.

REFERENCES

Appelt, S., Batta R., Lin, L., Drury, C. 2007. Simulation of passenger check-in at a medium-sized US airport. *Proceeding of the 2007 Winter Simulation Conference*, pp. 1252-1260. December 9-12, (Washington D.C., USA).

Chun H.W., Mak, W.T.R. 1999. Intelligent resource simulation for an airport check-in counters allocation system. *IEEE Transactions on System, Man, and Cybernetics* 29(3): 325-335.

Kesen, S.E., Baykoç, Ö.F. 2007. Simulation of automated guided vehicle (AGV) systems based on just-in-time (JIT) philosophy in a job-shop environment. *Simulation Modeling Practice and Theory* 15 272-284.

Joustra, P.E., Van Dijk, N.M.. 2001. Simulation of check-in at airports. *Proceeding of the 2001 Winter Simulation Conference*, pp. 1023-1028. December 9-12, Arlington (Virginia, USA).

Law, A.M. 2007. *Simulation Modeling and Analysis*. 4th ed. New York: Mc-Graw Hill Companies, Inc.

Lee, A.M., Longton, P.A. 1959. Queuing process associated with airline passenger check-in. *Operational Research Quarterly* 10(1): 56-71.

Parlar M., Sharafali, M. 2008. Dynamic allocation of airline check-in counters: a queuing optimization approach. *Management Science* 54(8):1410-1424.

Park, Y., Ahn, S. 2003. Optimal assignment for check-in counters based on passenger arrival behavior at an airport. *Transportation Planning and Technology* 26 (5): 397–416.

Takakuwa, S., Oyama T. 2003. Simulation analysis of international-departure passenger flows in an airport terminal. *Proceeding of the 2003 Winter Simulation Conference*, pp. 1627-1634. December 7-10, New Orleans (Los Angeles, USA).

Van Dijk N.M., Van der Sluis, E. 2006. Check-in computation and optimization by simulation and IP in combination. *European Journal of Operational Research* 171(1):1152-1168.

AUTHORS BIOGRAPHIES

ÖZDEN ONUR DALGIÇ is a master student in Industrial Engineering Department at TOBB University of Economics and Technology, in Ankara, Turkey. His research interests consist of machine scheduling and robotic cell scheduling. His current researches are

mainly on robotic cell scheduling with dual gripper robots. His email address is oodalgic@etu.edu.tr.

YUSUF SEÇERDIN is a master student in Industrial Engineering Department at TOBB University of Economics and Technology, in Ankara, Turkey. He has a bachelor's degree from business administration and his research interest includes transportation systems and simulation, and logistics. His email address is yusufsecerdin@etu.edu.tr.

GIZEM SULTAN NEMUTLU is a master student in Industrial Engineering Department at TOBB University of Economics and Technology, in Ankara, Turkey. She has a bachelor's degree from statistics and her fields of interests are supply chain management and, production system planning and control-with simulation. Her email address is gnemutlu@etu.edu.tr.

NILGÜN FESCIOGLU ÜNVER is an Assistant Prof. in Industrial Engineering Department at TOBB University of Economics and Technology, in Ankara, Turkey. Her research interests center on discrete event simulation, agent based systems and self adaptive systems. Her current research focus is on energy efficiency in manufacturing and healthcare systems. Her email address is nfunver@etu.edu.tr.

APPENDIX A

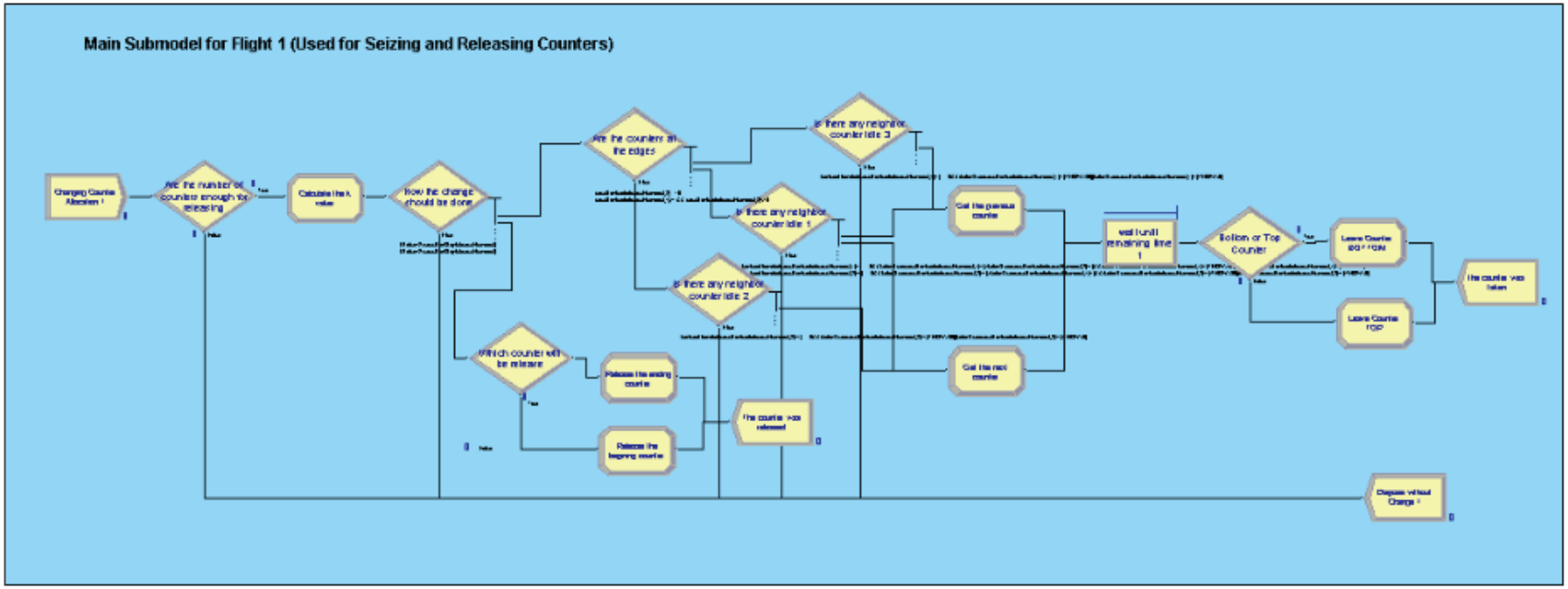


Figure 10: ARENA Simulation Submodel-Used for Seizing and Releasing Operations

INVESTIGATING SPATIAL POWER EFFECTS USING 3D REAL-TIME MODEL

Janos Sebestyen JANOSY

Senior Consultant, Centre for Energy Research, Hungarian Academy of Sciences,
P.O. Box 49, H-1525 Budapest, Hungary

Janos.S.Janosy@energia.mta.hu

ABSTRACT

Fuel assemblies were used in our nuclear power plant initially for 3 years, now for 4 years and soon they will stay in the core for 5 years. Each year only 1/3rd, 1/4th later 1/5th of them is replaced; therefore the change of the fuel type is a lengthy process, with mixed cores used in between. The full-scope simulator is upgraded to simulate the exact behavior of these mixed cores. The RETINA is a 3D thermo-hydraulic code (Hazi, 2001), the KIKO3D is a 3D neutron-kinetics code (Kereszturi, 2003), are operating parallel in real-time. Models were presented on WAMS2010 workshop in Buzios, Brasil (Janosy, 2010).

Verification and validation of these models are extremely difficult because the lack of experimental data. Since the Chernobyl accident no experiments with nuclear power are encouraged. The paper describes the experience gained during the V&V process – driving these coupled 3D models up to extreme conditions.

Keywords: NPP simulation, parallel processing, real-time simulation, Coupled 3D thermo-hydraulics and neutron kinetics.

INTRODUCTION

Fuel elements, integrated into fuel assemblies produce heat in the nuclear reactors in rather difficult, harsh conditions. The pressure and temperature is high - up to 160 bars and 320°C - and the power density in some reactors reaches 90 kW/liter and more. They are made from expensive metals using expensive technologies. They should not leak - the cladding represents the first barrier between the radio-active materials and the environment (usually there are at least three barriers). If there is a remarkable leak, the reactor should be stopped and the leaking fuel assembly replaced - a procedure causing significant economic loss.

Nevertheless, most of the fuel assemblies are well made and they practically never leak. During the 20-year-history of the four-unit Paks NPP there was detectable leak only once or twice. The fuel elements originally spent three years in the core, nowadays they stay for four years - with slightly higher uranium content, of course. If they should stay for five years, the increasing of the enrichment is not enough - the control system of the reactor is not designed to cover the excessive reactivity of the core, produced by the higher enrichment of the fresh fuel.

The solution is the Gadolinium (Gd) which is a burnable neutron poison. In the first year - or so - it helps to cover the excessive reactivity by absorption of neutrons, then it burns out and do not causes any problem in the upcoming years.

Now we replace every year 1/4th of the fuel elements with fresh ones. If we start to replace them with the new types, supposed to stay for five years, it means that we are going to use mixed cores at least for four years. These cores need special treatment and the operators should be trained to it. The core surveillance system must be fitted to these mixed cores, too.

THE 3D MODEL RQUIREMENTS

Earlier we could use simpler models with great success (Janosy, 2003, 2007 and 2008). Now we have 349 very different fuel assemblies in the core; each of them can be of different age and different composition. The core configuration is carefully optimized each year to ensure that the power distribution and burn-out corresponds to the maximal safety and to the best fuel economy.

The water flowing through the core of PWRs acts not only as coolant but as moderator, too - that means if fulfills the task of slowing the neutrons down in order to optimize the neutron balance and making the chain reaction stable and possible. Careful design of the reactors results in negative temperature and volumetric coefficients that means that the reactor is capable to self-regulate its power - because making the coolant hotter and thinner means worse neutron balance and therefore it decreases nuclear power.

These effects make the neutron kinetic model of the reactor and the thermo-hydraulic model of the primary cooling circuit tightly coupled, therefore they must be solved simultaneously. Describing very different physical phenomena they contain very different equations - that causes the problems of the simultaneous numerical solution. The required time step for the accurate numerical integration can be very different, too.

The crucial point is: how to nodalize the nuclear reactor and the primary circuit in order to achieve high fidelity of simulation with reasonable computer loads - in other words achieving accurate simulation still remaining in real-time. It looks easy to divide the equipment to very small parts, and solve the problem using them as coupled nodes.

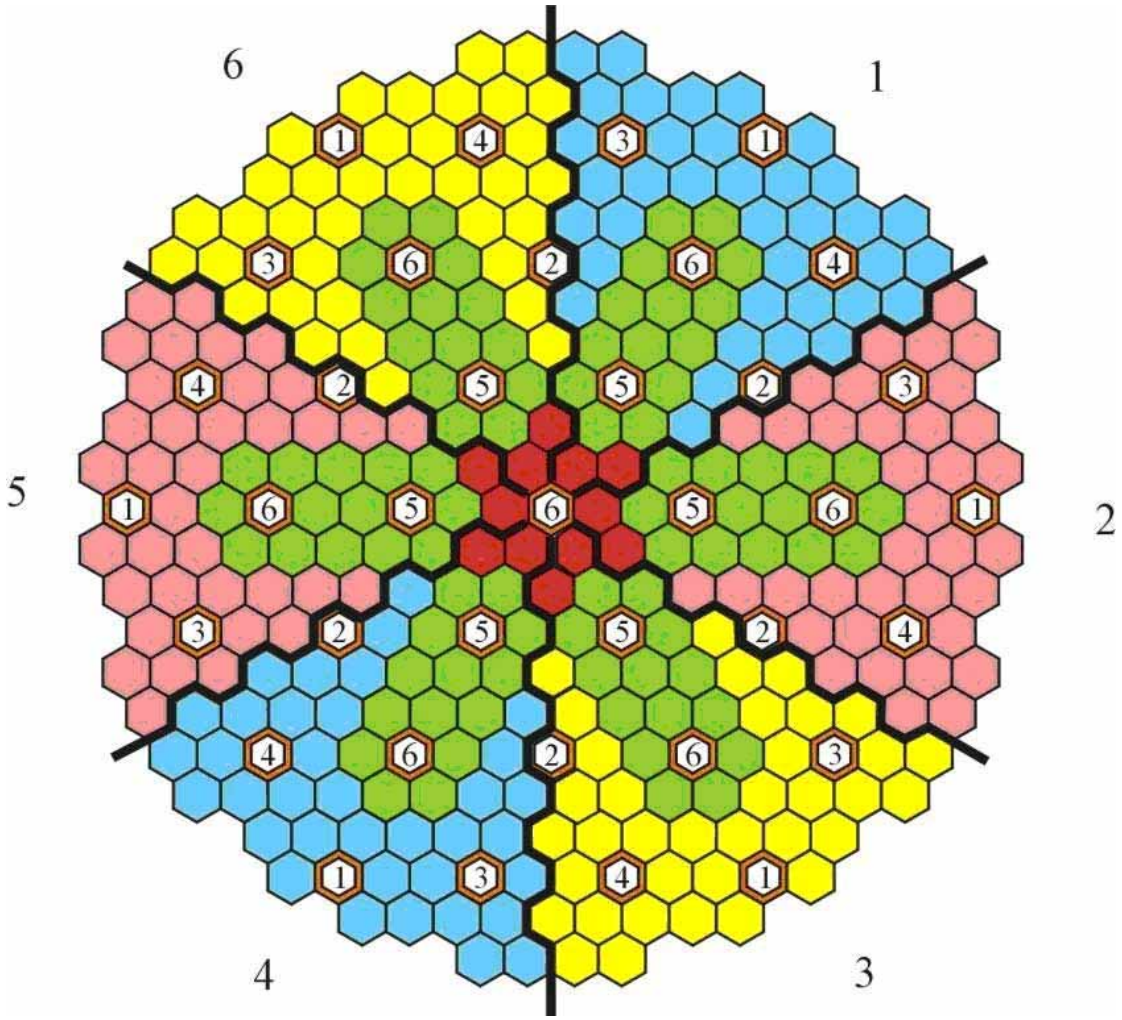


Fig 1. The map of the core with the 349 fuel assemblies, including the 37 control ones.

Decreasing the size of the individual nodes not only increases their number according to the third power, but in the same time it significantly decreases the necessary time step of the numerical integration.

NODALIZATION: NEUTRONICS

As it is shown on Pic. 1, we have in the core 349 hexagonal fuel assemblies (the numbers outside the core refer to the six cooling loops). The 37 numbered fuel assemblies are used to control the chain reaction. They are twice as long as a normal fuel assembly. The upper part is made from special steel designed do absorb the proper amount of neutrons in order to be able to control the chain reaction. The lower part is a usual fuel assembly containing usual amount of fuel. ***Pulling out*** this control assembly means that the lower part enters the core, ***lowering it*** causes this part to leave and to be replaced by the neutron absorber assembly.

The 37 control assemblies are organized into 6 groups, containing 6 assemblies except the 6th one, which contains 7 (this 7th is the central one). The first five groups with 30 assemblies are used as the "safety rods", fully pulled out during normal operation and fully lowered during reactor shut-down. The 6th group is normally used as "control rods", during normal operation they are always in different intermediate positions according to the prescribed power of the reactor. In some very rare situations the 5th group is helping to the 6th one, sometimes staying in intermediate position, too.

That evidently means that the first four groups do not influence the spatial distribution of the neutrons, they absorbents are pulled out and their fuel assemblies are inserted. Lowering them the reactor is shut down and the spatial distribution is not important any more. In the same time, the last two groups - the 5th and the 6th - can seriously influence the 3D distribution of the neutrons, being in different intermediate positions according to the different operating conditions of the reactor and the primary circuit. The nodalization of the core from the neutron kinetics point of view does not leave us too much freedom: each "neighbor" to each assembly can be of different "age" in the reactor (zero to four, later zero to five years), with or without Gadolinium content accordingly. Different "age" means different burn up, thus different stage of enrichment and different isotope content. That means that in horizontal plane each particular assembly should be a separate node – we are going to have as much as 349 nodes horizontally.

As to the vertical nodalization, we must have not less than 8 or 10 planes to get enough resolution (8 to 10 points) to describe the axial neutron (and heat) distribution. We have chosen 10 planes vertically - that means, we have finally 349 x 10 nodes for the KIKO3D model.

Real-time spatial (3D) simulation of 3490 nodes in several groups of neutrons according to their actual energy requires huge computer power. The only way to do it using finite number of processors means to separate the time and space problem. The result can be written as a product **of two functions:** the amplitude ***function of tim***e and the distribution ***function of space***.

NODALIZATION: THERMOHYDRAULICS

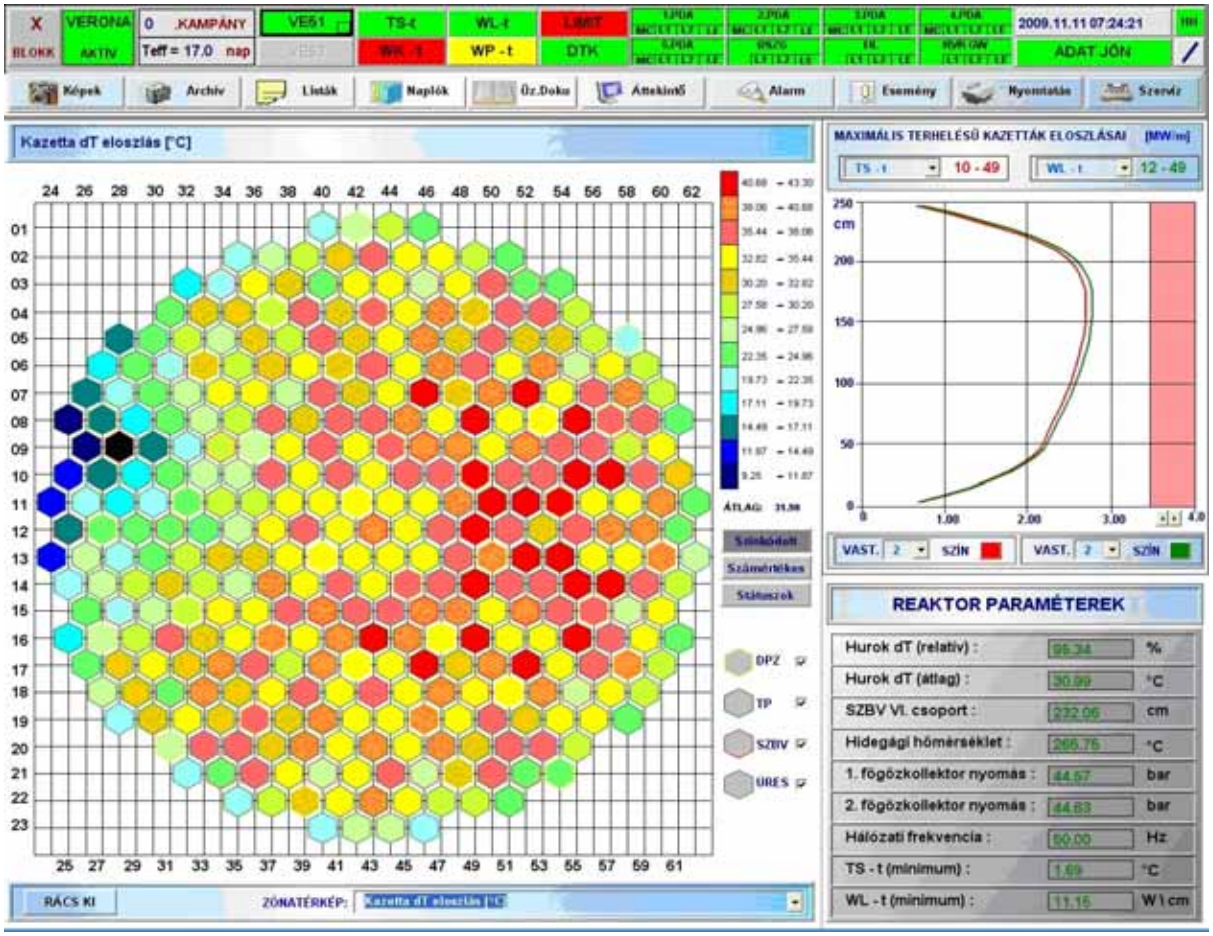
Thermo-hydraulic nodes should be much larger in space than the neutron-kinetic nodes. It is connected with the 0.2 sec. time step of the full scope replica simulator of the power plant. If we want to avoid large number of iterations, the amount of the steam/water leaving/entering the node each time step must be probably less than the full amount of the steam/water inside the node. It means that if we multiply the maximal feasible volumetric flow-rates with the 0.2 sec. integration time step, we get the minimal volumes for the nodes in question. Creating relatively large nodes we have to group fuel assemblies very carefully, in order to get as detailed simulation results as possible. The color coding of Pic. 1 shows our results.

The central (red) node on Pic. 1 contains 13 fuel assemblies, including the central "control rod" (from group No. 6). Six inner (green) nodes contain 16 fuel assemblies including 2 control assemblies each (one form Group 5 one form Group No. 6). The peripheral six large nodes, shown in different colors, contain 40 fuel assemblies each. Vertically we divide the core into 5 termo-hydraulical layers; it is easy to fit them with the 10 layers of the neutron-kinetical model.

This kind of thermo-hydraulic nodalization provides the following benefits:

- Only control rods of the 5th and 6th control rod group may have intermediate positions, influencing the spatial distribution of the neutrons. ***The inner 6 nodes and the central node*** are responsible for the calculation of these effects.
- One or more cooling loops may fail, usually because of the tripped main circulating pumps (MCPs). ***The six outer large nodes*** can respond spatially to these effects.

Thanks to the nodalization scheme described above, different spatial effects in the core can be studied. As an example, the "rod drop" malfunction is presented.

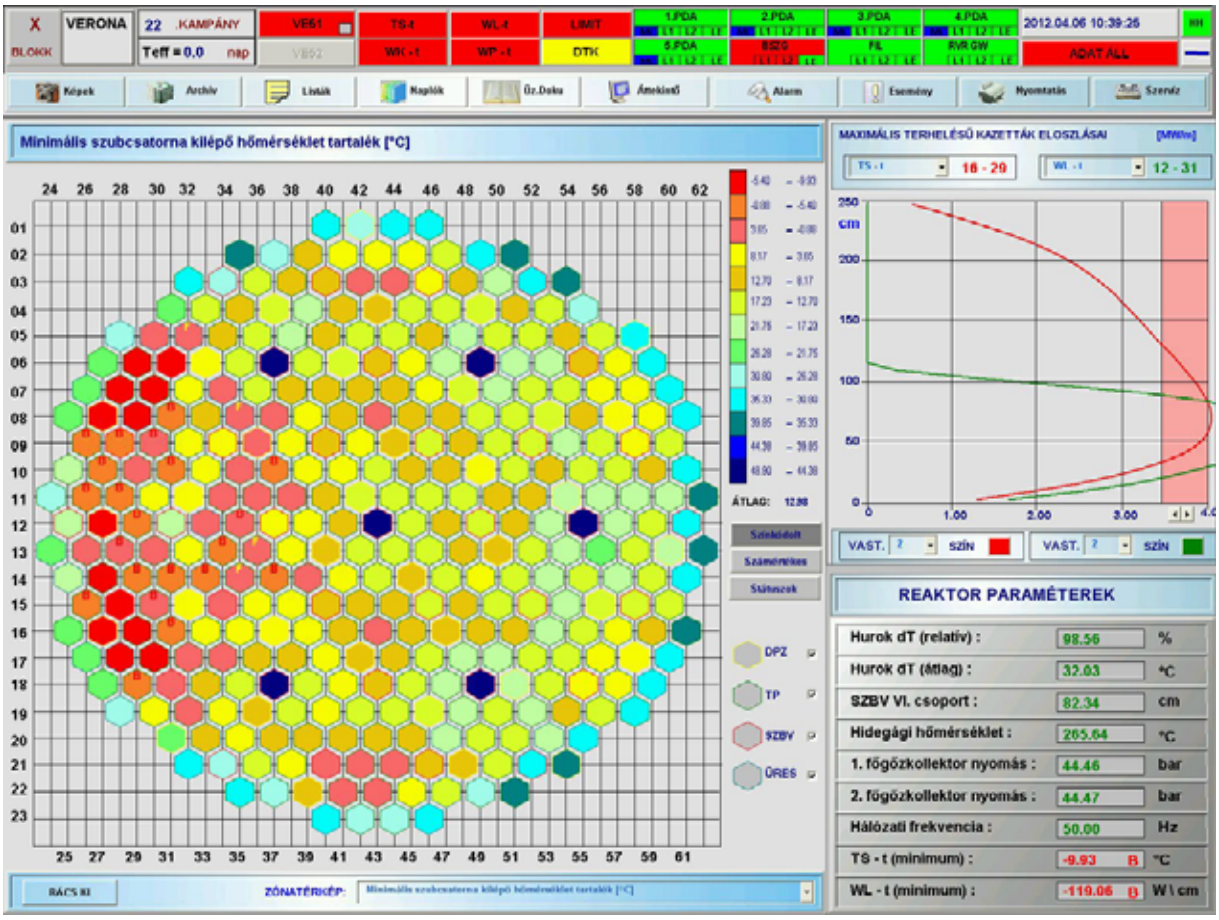


Pic. 2. Picture on the in-core surveillance system VERONA - driven by "rod drop" state data from the simulator

If a control rod erroneously drops into the core, the negative reactivity caused by it can be compensated by the power controller, pulling all the other rods a little out from the core. However, the power locally will be less around the fallen neutron absorber.

All well-designed reactors are self-regulating, that means overheating causes negative reactivity thus decreases the heat power, and overcooling does the opposite - it leads to positive reactivity and the power increases a little. This effect compensates the locally introduced (by the fallen rod) negative reactivity, and that's why the distortion of the power field - and the resulting temperature field – see Pic. 2. - is not so strongly distorted than it could have been expected. Details about out RETINA thermo-hydraulic are reported by Janosy, 2010, Hazi 2001, and Hazi, 2002.

Another experiment with the coupled spatial neutron kinetic and thermo-hydraulic models can be seen on Pic. 3. The controller keeps the power permanent moving the control rods. First we inserted some boron acid solvent to get all rods pulled out. (Boron acid is an absorber for the neutrons participating in the chain reaction). Next, we addressed the "control rod stuck" malfunction to the leftmost control rod. It will not move any more, remains in "pulled out" condition. Now extracting slowly the boron acid the controller has to move all the other control rods down, in order to keep the power constant. The restructuring neutron power distribution results in a serious overload at the left part of the core causing intensive boiling. The plant is not permitted to operate under such conditions: these situations can be studied on the simulator only.



Pic. 3. Six rods are "in", the leftmost 7th is "out". Note the overloaded and distorted axial distribution on the right side

VERIFICATION AND VALIDATION PROBLEMS

The cornerstone of all simulation is the verification and validation method used for the constructed models. In our case the ***verification*** means that the computer program representing our model is error-free and calculates the numerical solution of that differential equation system exactly that has been described in the Technical Design document During ***validation,*** the exact value of the rather numerous free parameters of the model should be determined optimally in order to achieve the highest confidence of the model system - ensuring the best similarity to the modeled real object.

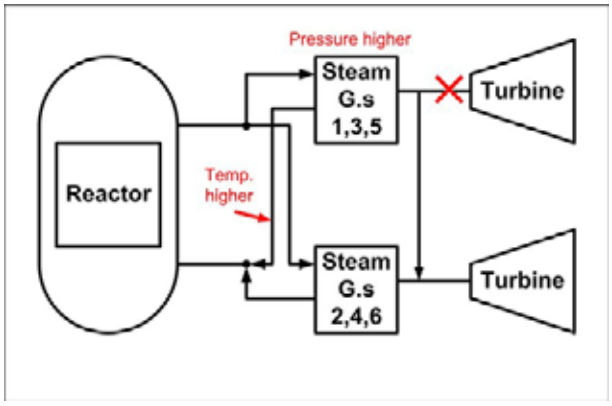
The validation is much more difficult than the verification process. The problem lies in the lack of proper experimental data. No experiments are allowed to perform on the real plant, and only the transient recordings of the **a**nticipated **o**perational **o**ccurrences (AOO) are available. This term stands for all events which are common: start-up, shutdown, operation on different power levels, control of the frequency of the electrical network, turbine start-up and trip, pump trips,

synchronization of the electrical generators, physical measurements verifying the parameters of a new core after re-fueling, etc. Careful investigation of the archives provided by the plant surveillance computer systems after a given transient produces great amount of important information.

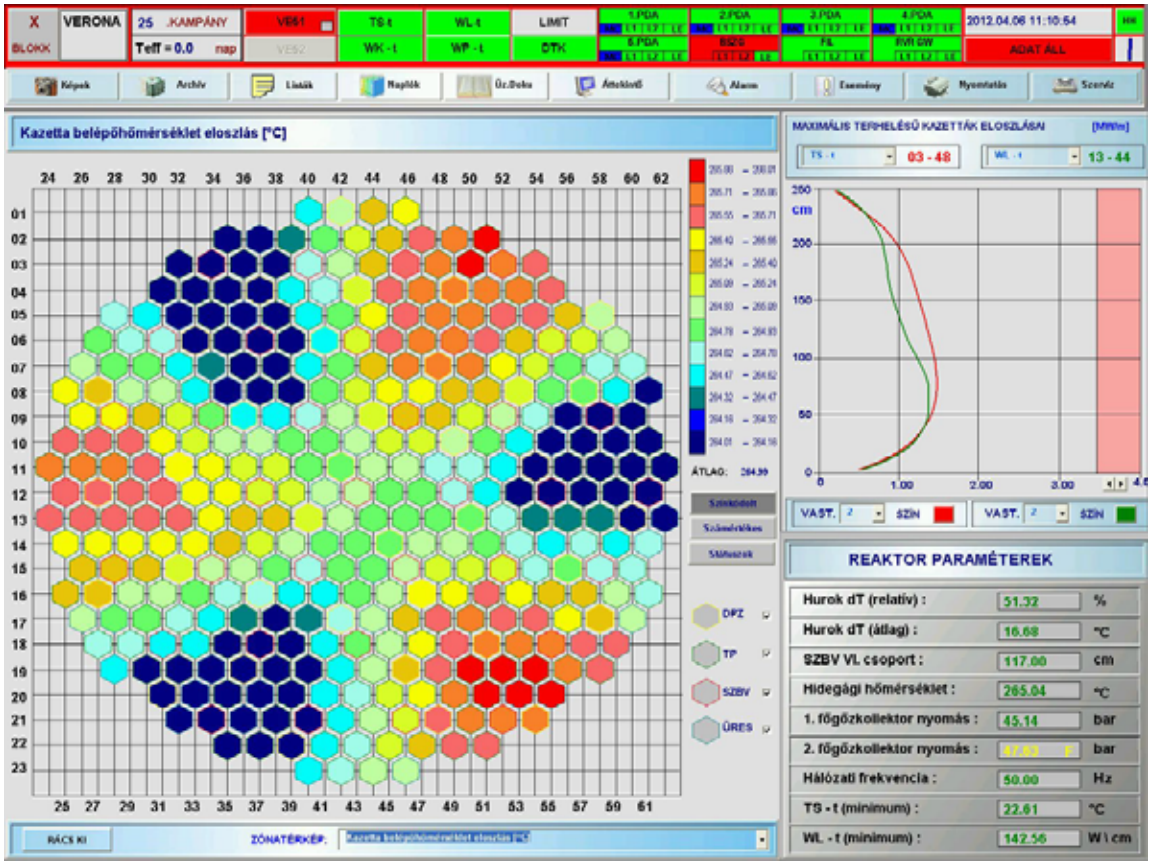
There are very well elaborated models and programs verified and validated by numerous and expensive experiments performed on dedicated experimental facilities. These models are used mostly to prepare the Safety Report of the given plant for the national authorities. The basic problem of the comparison of a training simulator models with them that the training simulator is based on the “best estimate” philosophy, while the programs mentioned above and used to evaluate the actual safety are based on the “worst case” methodology and scenarios. This difference becomes significant when the models have to handle uncertainties. “Best estimate” models usually take the estimated mean value, whereas the “worst case” models take the value which “hurts” most to the safety of the nuclear power plant.

Sometimes simple events help a lot. Turbine trips are not very frequent events, and it is not allowed to operate with only one turbine available. Sometimes, however, it is a simple electrical overload what happens and the protections disconnect the generator in question from the grid. If the operators are able to re-start the turbine and recover the full power soon, they do not shut down the plant. Meanwhile precious data can be collected.

The operation with one turbine is asymmetric. Three steam generators feeding the steam header of the operating turbine remain in a quite similar condition, but the other steam header to which the other three generators are connected remains without turbine. This steam header can feed the other one, but because of the connection losses the pressure of this header becomes higher by 2.5 bars than the other one. Correspondingly, the water temperature on the secondary side of them becomes higher by 2°C. Obviously, those parts of the core fed with coolant of higher temperature reduce somewhat the power (due to the negative temperature feed-back) therefore the output temperatures of these sectors are less warmer than the inlet difference of 2°C.

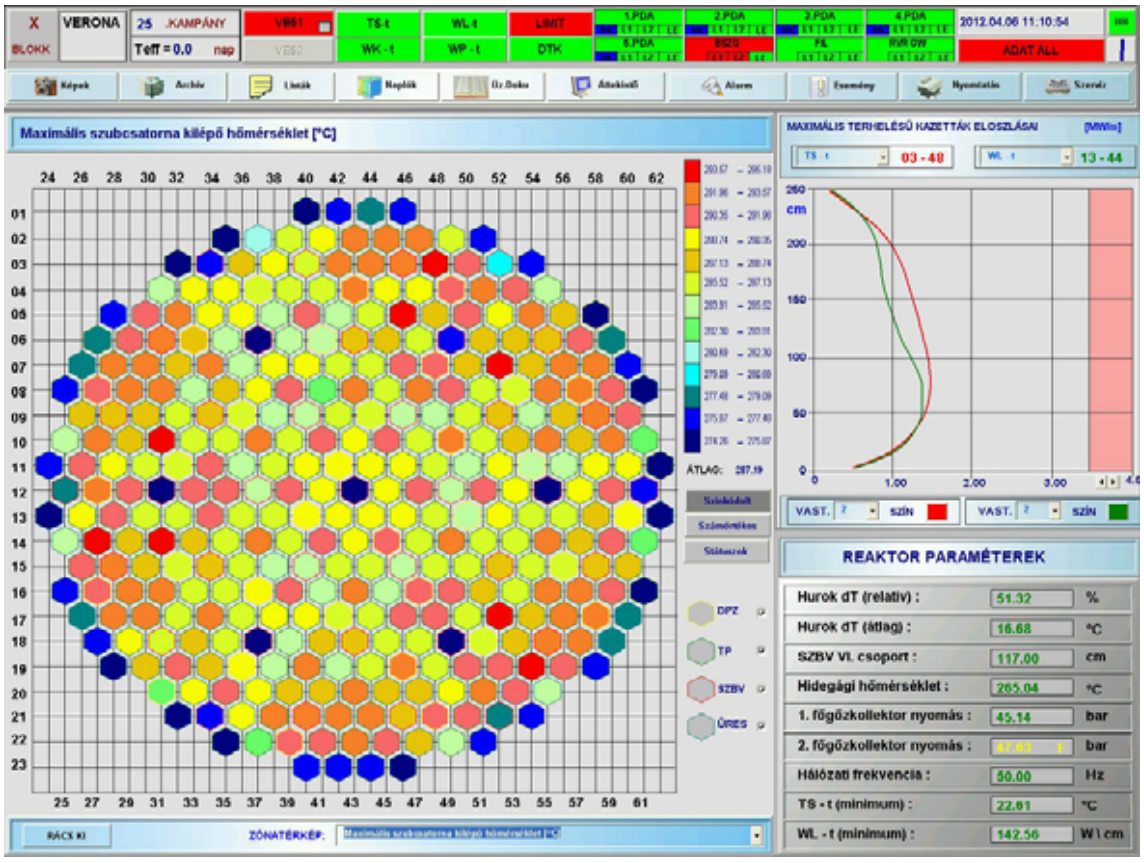


Pic. 4. Asymmetric operation with one turbine



Pic. 5. One steam turbine is out, and three from the six loops have higher cold loop (i.e. reactor inlet) temperatures

The reactor power controller had to reduce the thermal power of the core to about 50% (51.2% on the VERONA display below), and therefore all the seven control rods are lowered somewhat; the corresponding part of the fuel elements move out and down from the core and the outlet temperature of these channels decrease. (VERONA: in-core surveillance system for our plant.) All these effects can be studied on the Pic. 6.



Pic. 6. Fuel element outlet temperatures shown by the VERONA in-core surveilance system

CONCLUSIONS

After accomplishing a long project, we are able to simulate all important processes of our pressurized-water power reactor practically in all necessary states and in all necessary detail in order to train our operators to the upcoming new fuel and the corresponding new procedures. It is very important that the extremely fast-growing computer power can be used not only to enhance the quality of the GUI – as it is nowadays unfortunately usual - but it is used for more detailed and correct simulation of sophisticated systems.

The real trouble we have always to encounter: thanks to the growing computer power, even if we are trying to formulate as detailed models as possible, it is very difficult to obtain useful measurement data from the real plant to compare our models with.

The standard instrumentation existing on the plant - being fully satisfactory to control the operation and ensure safety – is usually inadequate to record data for comparison with the results of 3D calculations made using elaborated model systems. Lack of consistent data – this is the main problem for our V&V procedures. Up to now we do not see how to overcome this problem.

REFERENCES

Janosy J.S., 2010: Simulator Upgrade – Real-Time 3D Nuclear Reactor Simulation, WAMS 2010, May 5-7, Busios, Brasil, ISBN 978-85-285-0135-3, pp. 365-370

Janosy, J.S., 2003: Modeling and Simulation of Nuclear Energy in Eastern Europe. Business and Industry Simulation Symposium, 2003 Advanced Simulation Technologies Conference, Orlando, Florida, March 30 - April 03, 2003

Janosy, J.S., 2007: Simulation Aided Instrumentation and Control System Refurbishment at Paks Nuclear Power Plant. First Asian International Conference on Modeling and Simulation, AMS 2007, 27-30 March 2007, Phuket, Thailand, ISBN 0-7695-2845-7

Janosy, J.S., 2008: Simulators are the key for large-scale Instrumentation and Control System Refurbishment Projects. Keynote speech, Second Asian International Conference on Modeling and Simulation, AMS 2008, May 12-15, 2008, Kuala Lumpur, Malaysia.

Keresztúri, A., Hegyi, Gy., Maráczy, Cs., Panka, I., Telbisz, M., Trosztel I., and Hegedűs, Cs.: 2003: Development and validation of the three-dimensional dynamic code - KIKO3D, Annals of Nuclear Energy 30 (2003) pp. 93-120.

Házi, G., Mayer G., Farkas, I., Makovi, P., and El-Kafas, A.A.: 2001: "Simulation of a small loss of coolant accident by using RETINA V1.0D code", Volume 28, Issue 16, November 2001, Pages 1583-1594

Házi, G., Keresztúri, A., Farkas, I., Mayer, G., Hegyi, Gy., and Panka, I., 2002: “First experience with a six-loop nodalization of a VVER-440 using a new coupled neutronic-thermohydraulics system KIKO3D-RETINA V1.1D” Annals of Nuclear Energy, Volume 29, Issue 18, December 2002, Pages 2235-2242

ADVANCED DESIGN OF INDUSTRIAL MIXERS FOR FLUID FOODS USING COMPUTATIONAL FLUID DYNAMICS

Davide Marchini[(a)], Federico Solari[(b)], Mattia Armenzoni[(c)], Roberto Montanari[(d)], Marta Rinaldi[(e)], Eleonora Bottani[(f)], Gino Ferretti[(g)] and Giuseppe Vignali[(h)]

[(a), (c)] SITEIA Interdepartmental Centre, University of Parma, viale G.P.Usberti 181/A, 43124 Parma (Italy)
[(e)] CIPACK Interdepartmental Centre, University of Parma, viale G.P.Usberti 181/A, 43124 Parma (Italy)
[(b), (c), (f), (g), (h)] Department of Industrial Engineering, University of Parma, viale G.P.Usberti 181/A, 43124 Parma (Italy)

[(a)]davide.marchini@unipr.it; [(b)]federico.solari@unipr.it, [(c)]mattia.armenzoni@unipr.it, [(d)]roberto.montanari@unipr.it, [(e)]marta.rinaldi@unipr.it, [(f)]eleonora.bottani@unipr.it, [(g)]gino.ferretti@unipr.it, [(h)]giuseppe.vignali@unipr.it

ABSTRACT

This work focuses on discontinuous (batch) vertical fluid mixing systems for food fluids with particles. The purpose was to identify the main structural and geometrical parameters that can influence the mixing process, in order to obtain useful indications for the design of mixing systems. The products examined are Newtonian multiphase fluids, with different viscosity values.

A properly designed mixing process has to provide two main results: a satisfactory homogeneity of the product and the preservation of the integrity of solid particles. Different mixer designs were thus analyzed to identify the structural factors which have the greatest impact on the above requirements. In particular, a Computational Fluid Dynamics (CFD) software was used to carry out a series of simulations, from which some Key Performance Indicators (KPI) related to the effectiveness of the different mixer configurations were derived. From those KPIs, some main conclusions were derived, that can be useful to choose the appropriate design solutions for fluid mixers.

Keywords: food fluids, batch mixers, design optimization, computational fluid dynamics

1. INTRODUCTION

Mixing, in many industrial processes, is used for homogenization, i.e. to mix together multiple components into a homogeneous mixture. In particular, when fluids containing suspended particles are processed, the action of the mixer is to transfer a given amount of energy to the material, in order to develop it to an optimum, i.e. to bring it at a uniform inter dispersion of the two phases. The way this energy is transferred is also crucial: in fact, it has to keep the whole internal volume of the mixer in motion, avoiding, at the same time, damages to the solid particles.

Generally speaking, when talking about mixing, it should be noted that there are many different types of mixing systems, which can differ from one another as regards the operation mechanism (continuous or discontinuous), the shape and dimensions of tanks and impellers and their reciprocal orientation (vertical, horizontal or oblique orientation of the impellers axis) and disposition (centered or decentered impellers). In this paper, we focus on the optimization of a batch vertical mixer, and aim to determine whether (and to what extent) the structural characteristics of a mixer may impact on its effectiveness.

Batch mixers are simple discontinuous mixers, consisting of a containing vessel in which the fluid is kept in motion by means of rotating agitators. Batch mixers are very common in food industries, and are often used in different stages of the transformation process. Moreover, in many real industrial processes, homogenization is of paramount importance and may be a limiting factor for the success or failure of processes (Rahimi, 2005). Consider, for instance, the importance of homogeneity of solid and liquid contents for multiphase fluids, in which the solid fraction should be kept as constant as possible. If this does not happen, serious problems may occur during filling. Companies would not be able to ensure the constancy between a container and the other of the bottled product, making the process unsatisfactory.

In the light of the importance of this issue, several studies have been carried out, aimed at analyzing the mixing process from different points of view: mixing efficiency, mixing times, and power consumption have been investigated theoretically and experimentally (Godfrey, 1986; Ottino, 1989; Tanguy et al., 1992, 1996). Moreover, many approaches based on numerical modeling (Alexopoulos et al., 2002, Daskopoulos and Harris, 1996 and Rahimi et al., 2000; Jongen, 2000) and CFD simulation (Patwardhan and Joshi, 1999; Sahu et al., 1999; Rahimi, 2005) have been applied to the study of mixing and have proved successful in predicting the course of homogenization. A better knowledge of the way the fluid is mixed may improve the process outcomes in many ways, e.g., in terms of yield or efficiency of the process. Thus, it would be useful to be able to predict mixing by using modern technologies such as numerical models.

The work builds upon this consideration, and is based on the modelling of mixing through CFD simulation. Nonetheless, we also tried to introduce an innovative method for the evaluation of the effectiveness of mixing plants. Mixers were assessed using four new parameters(that will be detailed later in

the paper), trying to take into account all the aspects that may be indicative of a homogeneous mixing and of the preservation of pieces integrity.

2. MATERIALS AND METHODS

2.1. Basic design and structural variables of the mixer

The mixer considered in this study basically consists of a simple cylindrical vessel in which the fluid is kept in motion by means of a vertical impeller. The capacity of the tank has been set at 600 l, and is kept unchanged. Conversely, the tank configuration has been changed during the study. Such a plant is extremely flexible from a design point of view, given the possibility to act on many operating leverages. In order to understand which of these leverages have the greatest impact on the performance of the mixing process, some plant operators were preliminary asked about the importance of some structural characteristics of the mixer and their impact on mixing performance. Therefore, the tests were oriented to the analysis and the comparison of different batch mixers, which differed from each other as regards three main characteristics, related to the geometrical structure or to the properties of the treated product:

1. Rotor disposition inside the tank
2. Aspect Ratio
3. Viscosity of the analyzed fluid.

Each variable has been discretized on two levels (i.e., "high" and "low"), according to the description below, so that a total of 8 configurations was analyzed.

1. In each configuration the rotor is vertically disposed, but it can assume two different dispositions, i.e. centered or eccentric. This means that the impeller can be coaxial with the tank or at a certain distance from the tank axis. In the first case, the eccentricity of the two elements (i.e. the ratio between the radius of the tank and the distance of the axis of tank and rotor) is obviously null, while for all analyzed mixers with decentred impeller, the eccentricity was fixed at about 0.43. Therefore, the "Rotor disposition" variable could score 0 or 0.43.

Eccentricity is a very important factor for mixers, because it is connected with the shape of the impeller blades and with the presence of stators on the inner wall of the tank. In particular, with centered rotor, the tank was considered equipped with stators, while in the case of decentralized impeller, stators were not inserted in the geometry. Stators are fixed metal parts inside the tank, which should deflect the flow of product in order to create relative motions that help the mixing process. The position of the rotor inside the tank also affects the shape of blades. Figure 1 shows and compares two possible configurations of the mixer (i.e., with axial and decentralized rotor).

With eccentric impeller, mixers are equipped with 5 superposed propellers, all equal in shape, and whose dimensions are limited by the proximity of the wall. The task of creating recirculation motion inside the system is executed thanks to the asymmetry of the geometry. On the other hand with centered rotor the design of the blades is more complex: there are only 4 propeller, equal in pairs.

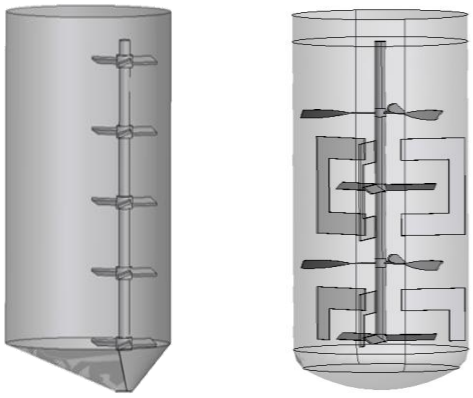


Figure 1 - The two configurations analyzed: axial rotor and stators (a) and decentralized rotor (b)

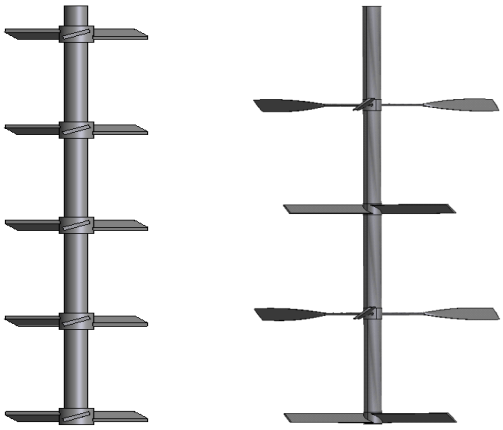


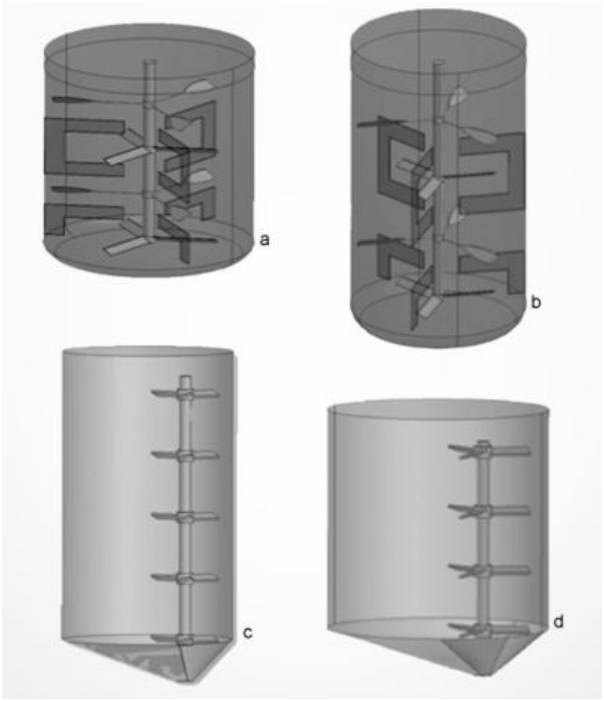
Figure 2 - Detail of the two different impellers of mixers with axial and decentralized rotor

Two of them are simple propeller, while the others are constituted by a kind of oars disposed radially. The internal blades should create a down flow, the external ones an up flow, in order to facilitate the recirculation.

2. The Aspect Ratio is a coefficient which describes the shape of the tank, and is computed by dividing its height by its diameter. An Aspect Ratio greater than 1 means that the tank is higher than wide and vice versa. Two configurations, having Aspect Ratio of respectively 0.9 and 1.55, were taken into account in this study. In the first one, the tank has a diameter of 931 mm and an height of 838 mm; in the second one, the diameter is 776 mm and height 1200 mm (Figure 3).

3. Two different fluids were considered, one with a high viscosity value and the other with a significant

lower one. To get practical feedbacks, the more viscous fluid owns approximately the viscosity of apricot puree, while the fluid with lower viscosity is similar to water at 30°C.



a = Aspect Ratio 0.9, centered impeller with stators
b = Aspect Ratio 1.55, centered impeller with stators
c = Aspect Ratio 1.55, eccentrical impeller
d = Aspect Ratio 0.9, eccentrical impeller

Figure 3 - The four geometrical configuration analyzed for a tank with a constant capacity of 600 l

As said previously, the study consists of eight simulations which were performed by varying 3 input variables on 2 levels:

- Aspect ratio: Level 1: 0.9; Level 2: 1.55;
- Rotor position: Level 1: 0 (centered) Level 2: 0.66 (eccentric)
- Viscosity: Level 1: 8.6e-4 Pa·s; Level 2: 10 Pa·s.

2.2. Simulation settings

The most relevant part of a simulation model for a mixing process is the impeller modelling. In this study, the CFD simulations were carried out by means of the commercial software Tdyn Multiphysics. It has a mesh deformation module that includes all the necessary capabilities to solve problems with mesh updating techniques. This module includes several mesh updating techniques and arbitrary Lagrangean Eurelian Algorithms (ALE) for solving systems of equations. In order to solve problems with moving parts, it exploits the "sliding mesh" or "sliding grid" method: the flow domain is divided into two cylindrical, non-overlapping sub-domains, each gridded as a separate block. The grid in the impeller region rotates together with the impeller, while the grid in the tank remains stationary. the two mesh crawl over one another at the cylindrical interface. Brucato et al. (1998) demonstrate that the sliding-mesh method provided good agreement with experimental data for the mean flow field.

At the sliding interface, a conservative interpolation is used for both the mass and momentum equations, using a set of fictitious control volumes. In order to reduce the errors related to the interpolation and thus to obtain good convergence of the simulation, it is necessary that the grid on the interface surface is sufficiently fine. In this study, by setting a maximum size of the elements in this area of 2 cm, good convergences of the simulation were obtained.

Another critical area is the region close to the impeller, where transfers of motion from the mechanical organ to the fluid take place. It is therefore necessary to set a condition of no-slip wall on impeller surfaces and produce a dense mesh in order to reconstruct the fluid boundary layer. To achieve this, a maximum size of the elements of 2 mm is assigned to these surfaces. Moreover, the elements size is further reduced in the direction normal to the surface (0.1 mm), in order to drop at least two layers of the grid inside the boundary layer. No-slip boundary conditions are also used on the tank walls and on the baffles. No experimental data is prescribed in the outflow of the impeller. All fluid motion strictly arises from the rotation of the impeller blades.

Spalart-Allmaras turbulence model was used during the simulation. This model solves directly a transport equation for the eddy viscosity and it is quite popular because of its reasonable results for a wide range of flow problems and its numerical properties (Deck et al., 2002). It is a one equation model for turbulent flows with integration to the wall. The aim of this model is to improve the predictions obtained with algebraic mixing-length models to develop a local model for complex flows, and to provide a simpler alternative to two-equations turbulence models. Contrarily to two-equations turbulence models, it does not require an excessively fine grid resolution in wall-bounded flows, and it shows good convergence in simpler flows (Tdyn Turbulence Handbook). Those characteristics allow to reach a good compromise between the precision of the results and the computation time, even in large size contexts as the industrial ones.

Time dependent simulations were performed with a characteristic time of the simulations of $5 \cdot 10^{-3}$ s.

The purpose of this work is to develop a method that can give indications on the performance of the mixing process, in terms of homogeneity, effectiveness and damage of the particles. However, the behaviour of solid-liquid mixtures is very complex, as well as their simulation through CFD software. In this work we will develop an innovative method for evaluate the performance of a mixer without including within the simulation the solid particles: the fluid dynamic simulations were conducted considering fluid without particles, and the results were investigated in order to provide hypotheses on the behaviour of the mixer in the

presence of particles. Obviously, the lower is the solid phase concentration the more reliable are the results obtained.

2.3. Performance measurement

The final objective of a mixer for fluids with particles is to obtain a homogeneous dispersion of the solid phase within the liquid phase . To achieve that, the mixer must grant the following actions:

- Avoid the stratification of the solid phase;
- Ensure the suspension of the solid phase within the liquid medium.
- Ensure the homogeneity of the suspension of the solid phase inside the tank;
- Avoid damage of suspended particles.

To have an indication about the overall performance of a mixer, all the above factors should be taken into account. A Global Performance Index (GPI) was then introduced to include four main KPIs of the mixer, namely:

- Velocity index;
- Mixing homogeneity index;
- Particles damage index;
- Mixing efficacy index.

To ensure comparability of various aspects, each KPI was expressed in percentage. In particular:

- Velocity index: calculated as the percentage of nodes (i.e. the volume) where the speed is greater than a threshold value.
- Mixing homogeneity index:

$$MHI = \frac{\sigma}{\sigma_{max}}$$

  The maximum standard deviation value (σ_{max}) was calculated assuming a fictitious distribution in which half the nodes has zero velocity and remaining ones have the maximum speed detected inside the tank.
- Particles damage index: calculated as the percentage of nodes where the impact coefficient is lower than a threshold value.
- Mixing efficacy index: calculated as the percentage of vertical kinetic energy to the total energy supplied to the fluid.

The four KPIs were then reported on a Kiviat (or radar) chart; this is a graphical method to display data relate to multiple variables in the form of a two-dimensional graph of several (4 variables in this case), represented on multiple axes with the same origin. In Figure 4 the radar chart for the above KPIs is represented: clockwise, the axes show the Mixing homogeneity index, the Particles damage index, the Velocity index, and the Mixing efficacy index.

By joining the four points obtained by plotting the KPI values on the axes, a rhombus is obtained. The ratio between the area of the rhombus divided and the maximum achievable area corresponds to the GPI.

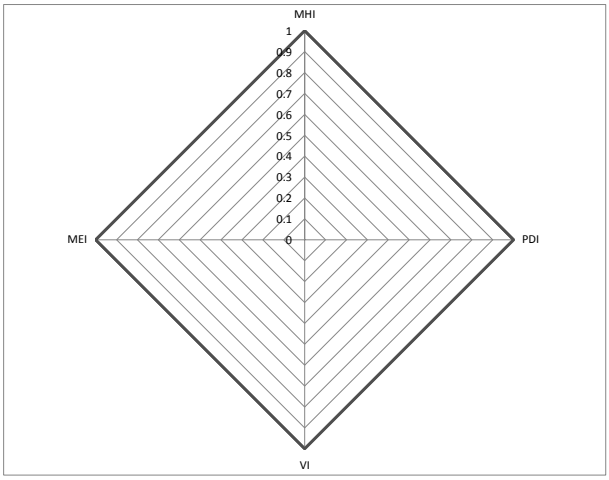


Figure 4: Kiviat chart used to represent the global performance of the mixers

3. RESULTS AND DISCUSSIONS

3.1.1. Velocity distribution

This index, as previously stated, represent the fraction of the internal volume of a mixer, whose velocity absolute value is higher than a certain threshold value. Such a threshold value must be calibrated so that it is representative of the minimum speed required at each point of the fluid so that it may be considered in a good state of mixing. The chart in Figure 5 shows the velocity distribution in the eight cases considered.

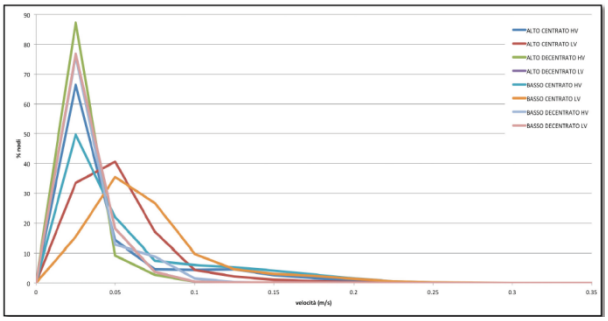


Figure 5: velocity distributions for the configurations analyzed

A better speed distribution can be observed for the mixers with centered rotor; in fact, the curves related to these configurations are particularly shifted to the right, i.e. towards higher values of speed.

In mixers with decentralized rotor, most of the fluid (75-85%) moves with very low speed values. Greater values are observed for mixers with centered rotors and in particular for the configuration characterized by low viscous product ($8.6 \cdot 10^{-4}$ Pa·s).

Table 1 shows the percentages of volume with a fluid velocities higher than a reference value (v_{ref}).

Table 1: percentages of volume with $v>v_{ref}$

Aspect ratio	**high**	**high**	**high**	**high**	**low**	**low**	**low**	**low**
Rotor position	**centered**	**centered**	**decentered**	**decentered**	**centered**	**centered**	**decentered**	**decentered**
Viscosity	**high**	**low**	**high**	**low**	**high**	**low**	**high**	**low**
VI (v_{ref}=0,025 m/s)	33.52%	66.54%	12.75%	23.12%	50.36%	84.65%	24.03%	23.12%
VI (v_{ref}=0,05 m/s)	19.28%	25.99%	3.55%	5.05%	28.45%	49.22%	11.34%	5.05%
VI (v_{ref}=0,1 m/s)	10.21%	4.57%	0.34%	0.75%	14.85%	12.77%	0.86%	0.75%

Speed reference values (v_{ref}) in Table 1 were chosen for illustrative purposes in order to make a comparison between the different mixer configurations. In practical cases, the designer will refer to the minimum speed which the fluid must have in order to maintain the particles in suspension and will choose the configuration where the whole fluid moves with a speed higher than the reference one (percentage 100%).

Table 1 shows that mixers with the rotors in centered position guarantee a better motion of the fluid.

3.1.2. Mixing homogeneity index

The velocity distribution standard deviation (σ) indicates how velocity is homogeneous inside the mixer and then indicates how homogeneous the mixing process is. Table 2 shows the standard deviation values calculated:

Table 2: velocity distribution variance for each configuration

Aspect ratio	**high**	**high**	**high**	**high**	**low**	**low**	**low**	**low**
Rotor position	**centered**	**centered**	**decentered**	**decentered**	**centered**	**centered**	**decentered**	**decentered**
Viscosity	**high**	**low**	**high**	**low**	**high**	**low**	**high**	**low**
σ [m/s]	0.041	0.030	0.016	0.013	0.048	0.041	0.023	0.018
σ_{max} [m/s]	0.101	0.095	0.049	0.044	0.113	0.114	0.067	0.064
MHI	59.54%	68.02%	68.06%	71.16%	57.86%	64.08%	65.94%	71.74%

The configurations with the rotors in decentered position guarantee a more homogeneous process. This result can be also inferred observing Figure 5: the velocity distributions in the mixers with a decentralized rotor position is highly concentrated around very low values. Thus, the homogeneity looks very good. In the case the system performance are examined from a more global point of view, it can be seen that this is due to the fact that almost all the liquid is at a standstill; this is why this kind of geometry with these speeds of rotation (10 rpm) is not able to adequately put the product in motion .

The above considerations suggest that the process homogeneity cannot be taken as the unique parameter in the choice of a mixer; conversely, a more detailed evaluation is required.

3.1.3. Particles' damage index

The mixing process of a fluid containing solid particles (i.e pieces of fruit, sacs, etc.) can lead to damage of the suspended particles due to possible collisions between the particles themselves or between the particles and the moving parts. In order to estimate the probability of having particles' damage, an Impact Coefficient was introduced. This coefficient was calculated at each node of the geometry of each mixer, applying the following formula:

$$C_{Impact} = \left|\frac{\partial v_x}{\partial x}\right| + \left|\frac{\partial v_y}{\partial y}\right| + \left|\frac{\partial v_z}{\partial z}\right| \quad (1)$$

It describes the acceleration imparted, in the three normal directions x, y and z, to a hypothetical solid particle which is located in a certain node at a certain time instant. The higher is this acceleration (or deceleration), the stronger will be the force acting instantaneously on the particle to change its state of motion, with a corresponding increase in its damage probability. In fact, it is supposed that a strong velocity variation can occur as a consequence of an impact. Figure 6 represents the trend of the Impact coefficient for the different mixer geometries:

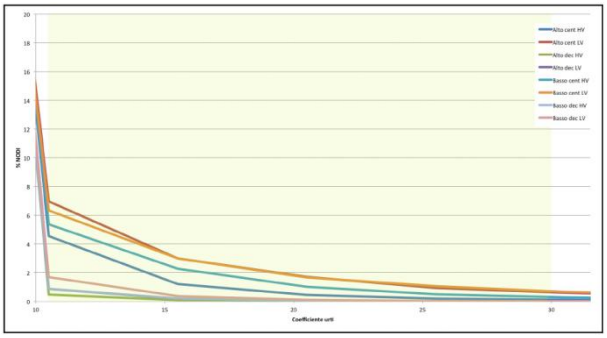


Figure 6: distribution of impact coefficient for the configurations analyzed

On the basis of this Impact coefficient, a proper KPI, named *Particles Damage Index (PDI)* was then computed, as already mentioned, as the percentage of nodes (i.e. volume) with a value of Impact Coefficient lower than a threshold value. In this study, the threshold value was set at 10 s-1; this value was chosen for illustrative purposes in order to make a comparison between the different configurations. In practical cases, the designer will choose this parameter as a function of

the fragility of the particles contained in the liquid phase.

The mixer configurations with a centered rotor position show higher PDI, because of the higher diameter and then the higher peripheral speed of the rotor.

3.1.4. Mixing efficacy index

Mixing efficacy is related to the ability of the mixer to prevent stratification of the particles that, in doing so, would tend to deposit to the bottom or to float on the free surface. To avoid stratification, it is necessary that most of the energy supplied to the fluid is directed vertically. The Mixing Efficacy Index (MEI), therefore, assesses the percentage of vertical kinetic energy against the total energy supplied to the fluid.

$$MEI = \sum_{i=1}^{n} v_{y;i}^2 \,/ \sum_{i=1}^{n} \left(v_{x;i}^2 + v_{y;1}^2 + v_{z;1}^2 \right) \qquad (2)$$

Where $v_{x;i}$, $v_{y;i}$ and $v_{z;i}$ are the three components of the velocity at each point *i* of the geometry. In particular $v_{y;i}$ is the vertical component of the velocity

Figure 7 represents the sharing of the kinetic energy among the horizontal and vertical components, for different mixer configurations.

A first analysis shows that the centered position of the rotor allows increasing the vertical component of the velocity, since the aspect ratio and the viscosity have less impact on the MEI.

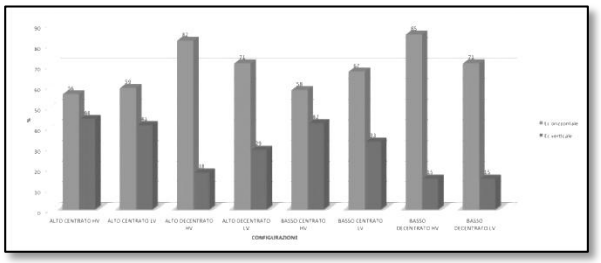


Figure 7: sharing of the kinetic energy among the horizontal and vertical components

From the results obtained, the radar charts for each configuration was derived as shown in Figure 8.

Table 3: KPIs of the configurations analyzed (8.33 rpm)

Aspect Ratio	**high**	**high**	**high**	**high**	**low**	**low**	**low**	**low**
Rotor position	**centered**	**centered**	**decentered**	**decentered**	**centered**	**centered**	**decentered**	**decentered**
Viscosity	**high**	**low**	**high**	**low**	**high**	**low**	**high**	**low**
MHI	59.54%	68.02%	68.06%	71.16%	57.86%	64.08%	65.94%	71.74%
PDI	97.73%	96.33%	99.88%	99.45%	95.20%	91.59%	99.73%	99.77%
VI	33.51%	66.54%	12.75%	23.13%	50.36%	84.65%	24.03%	23.13%
MEI	41.12%	43.66%	17.86%	29.08%	41.78%	33.19%	14.67%	29.08%
GPI	**32.30%**	**47.09%**	**23.79%**	**30.30%**	**37.06%**	**46.39%**	**25.73%**	**30.56%**

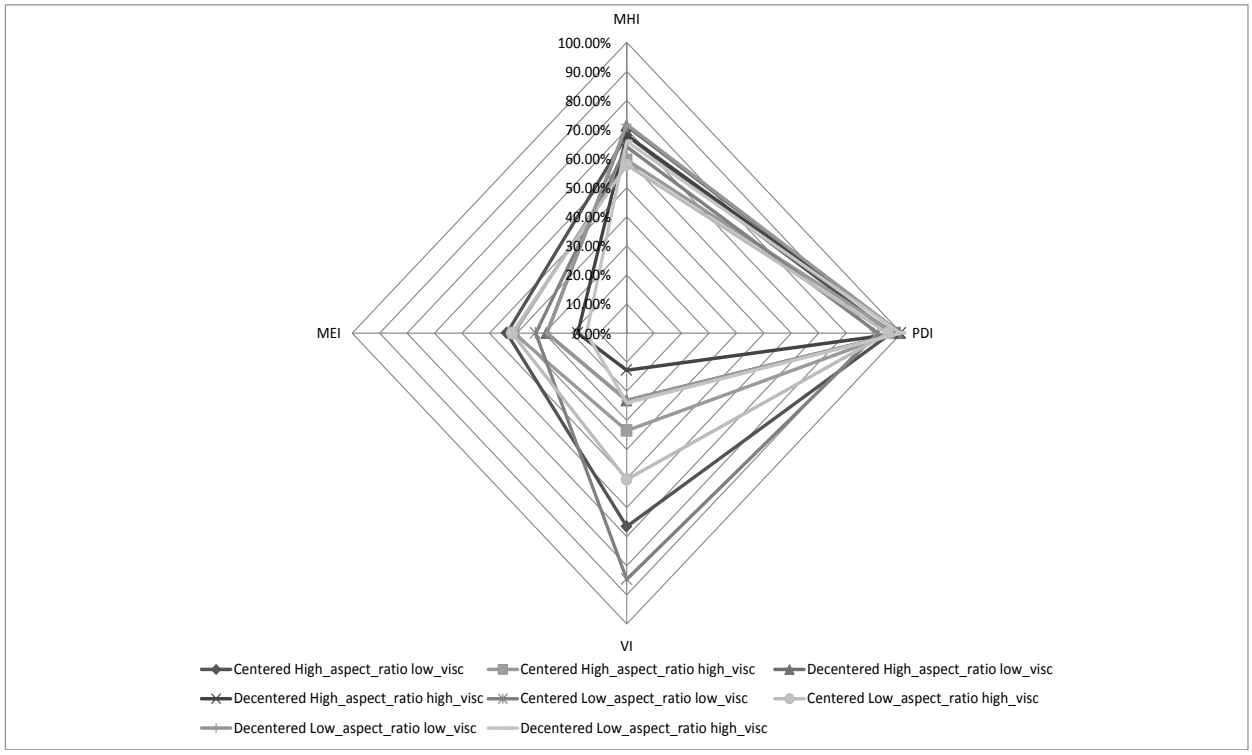


Figure 8: Kiviat chart for the different configurations analyzed

From the results obtained, it emerges that, in the configurations with a decentered rotor, a blade speed rotation of 50 degrees per second (i.e. 8.33 rpm) produces a velocity distribution within the fluid. In the best configuration (i.e., low aspect ratio and high viscosity fluid), only 24% of nodes exceeds 0.025 m/s, while in the worst configuration (high aspect ratio and high viscosity fluid), this percentage drops to 12.7%. In these configurations, in fact, the diameters of the rotors are much lower than those set in the configurations with centered rotor; consequently, the power involved and the energy supplied to the fluid are significantly lower. **Errore. L'origine riferimento non è stata trovata.** shows a comparison between the power required on the rotor shaft in decentered configurations and in the centered ones.

In order to improve the comparison between the performances of the mixers with decentered rotors, and those of the mixers with centered rotors, it is necessary that the energy supplied to the fluid is approximately the same. For this reason, the four analyses related to the mixers with decentered rotor were repeated with a different value of speed (25rpm).

The results of these analysis are summarized in Table 4.

Table 4: KPIs of the configurations with the rotors in decentered positions (25 rpm)

	Decentered High_Aspect_Ratio Low_Visc_25rpm	**Decentered High_Aspect_Ratio High_Visc_25rpm**	**Decentered Low_Aspect_Ratio Low_Visc_25rpm**	**Decentered Low_Aspect_Ratio High_Visc_25rpm**
PDI	98.13%	98.22%	95.91%	96.70%
VI	60.21%	34.38%	80.29%	53.29%
MEI	28.22%	14.21%	33.42%	12.87%
MHI	70.81%	68.62%	72.75%	66.77%
GPI	**41.39%**	**28.95%**	**49.48%**	**32.89%**

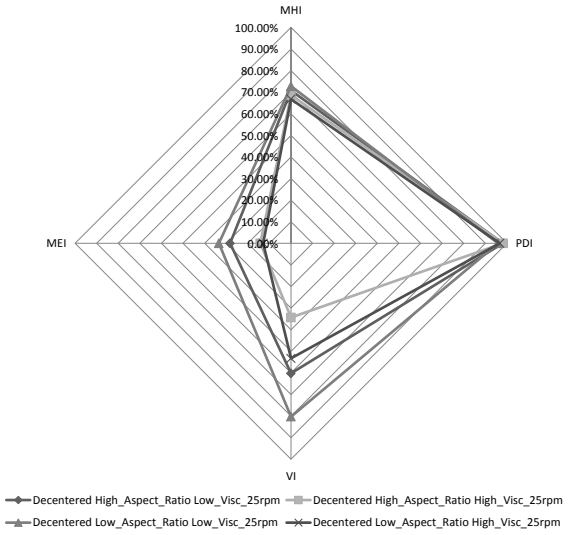


Figure 9: Kiviat chart for the four configurations analyzed with decentered rotors.

The highest GPIs is obtained when the product has low viscosity (8.6 e-4 Pa*s); the aspect ratio turns out to have a limited influence on the GPI.

This kind of configuration, however, provides inadequate performance when mixing high viscosity fluids .

Finally, it should be noted that to achieve the same performance realized by the mixer with the rotor in centered position using a mixer with the rotor in decentered position, the rotation speed should be increased of a factor of 3.

3.2. Factorial analysis

A full factorial analysis was conducted on the results obtained, in order to assess the significance of the influence of the three input variables (position of the rotor, viscosity, aspect ratio), on the overall performance index (GPI) and on the individual KPIs.

A commercial software (Design Expert®) was used to carry out the factorial analysis. The results, in terms of the effect of individual factors on the GPI and on the other KPIs, are reported below.

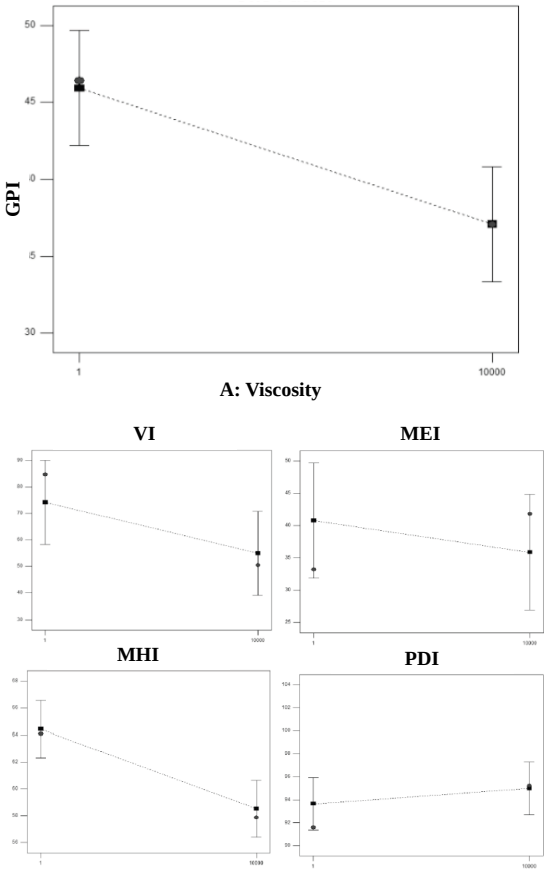


Figure 10: influence of viscosity on process KPIs

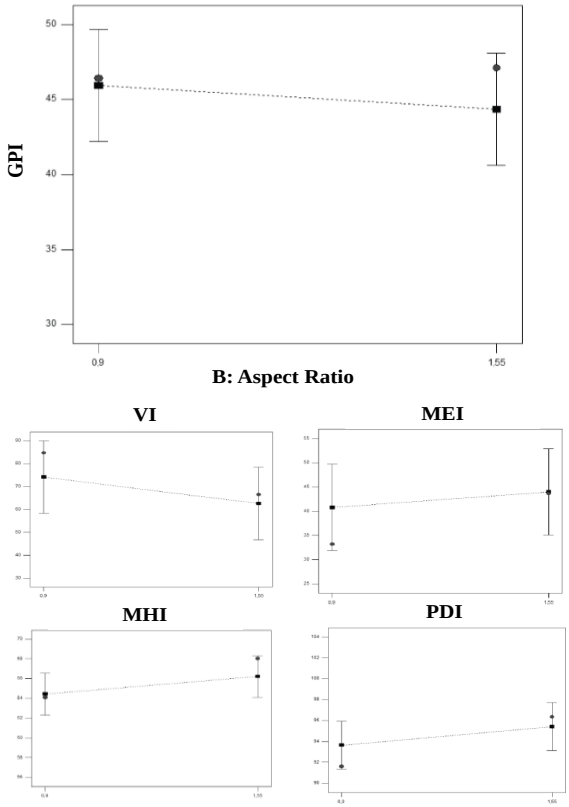


Figure 11: influence of aspect ratio on process KPIs

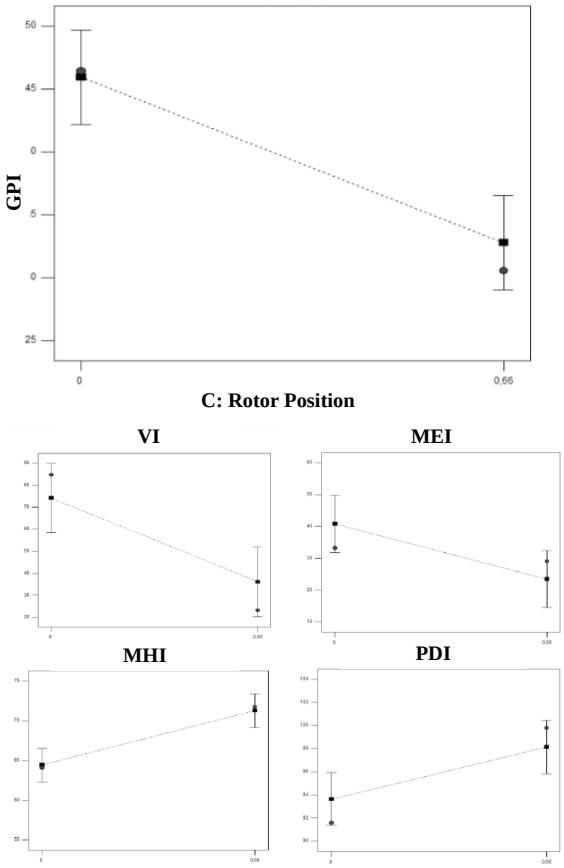


Figure 12: influence of rotor position on process KPIs

Through the observation and the analysis of the graphs, the influence of each variable on the KPIs could be evaluated. In particular, the higher the slope of the line, the higher the influence of the variable on the KPI measured.

It can be seen that PDI is influenced to a limited extent by all the factors. Conversely, MHI and VI seem to be very influenced by both rotor position and viscosity, while the influence of aspect ratio seems to be lower. Overall, the aspect ratio turned out to be a parameter with little influence on all response variables. Conversely, MEI resulted to be significantly influenced by the rotor position.

Finally, the ANOVA analysis performed on the results confirmed that the GPI was significantly influenced by both the viscosity and the rotor position, while the influence of aspect ratio turned out not to be significant

Table 5 - Results of the ANOVA analysis to understand what are the factors that impact more on KPIs considered

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Model significant	506.6	3	168.87	23.18	0.0054
A-Viscosity	*156.56*	*1*	*156.56*	*21.49*	*0.0098* ***SIGNIFICANT***
B-Aspect Ratio	*5.02*	*1*	*5.02*	*0.69*	*0.4532 NOT SIGNIFICANT*
C-Rotor position	*345.03*	*1*	*345.03*	*47.36*	*0.0023* ***SIGNIFICANT***

4. CONCLUSIONS

In this study an innovative approach for the evaluation of the performances of discontinuous vertical fluid mixing systems for food fluids containing particles was developed. In particular four KPI's were introduced in order to evaluate the capability of the mixer to prevent both particles stratification and damage and to guarantee a homogeneous process. The stochasticity linked to the shape of the particles, their distribution, and their interactions, prevents from conducting a reliable CFD simulation of the real process conditions. Some approximation is therefore necessary. In this study it was decided to neglect the solid phase and the CFD simulation were conducted considering only the liquid phase; the results obtained were then analyzed in order to provide hypothesis on the behavior of the particles. The results will be the more reliable the lower the concentration of the solid phase is; a validation phase is already planned in order to verify the accuracy of the results obtained.

This new method was applied to the analysis of different configurations. It was demonstrate how it can be useful in the design phase of a new mixing plants or in the optimization of an existing one.

REFERENCES

M. Rahimi, The effect of impellers layout on mixing time in a large-scale crude oil storage tank - Journal of Petroleum Science and Engineering 46 (2005) 161–170.

A.H. Alexopoulos, D. Maggioris and C. Kiparissides, CFD analysis of turbulence non-homogeneity in mixing vessels—a two-compartment model - Chem. Eng. Sci. 57 (2002) 1735–1752.
P. Daskopoulos and C.K. Harris, Three dimensional CFD simulations of turbulent flow in baffled stirred tanks—an assessment of current position - Chem. Eng. Symp. Ser. 140 (1996) 1–13.
S. Deck, P. Duveau, P. D'Espiney, P. Guillen, Development and application of Spalart–Allmaras one equation turbulence model to three-dimensional supersonic complex configurations. Aerospace Science and Technology, 6 (2002), 171–183
M. Rahimi, P.R. Senior and R. Mann, Visual 3-D modeling of stirred vessel mixing for an inclined-blade impeller - Trans. Inst. Chem. Eng. 78 (2000) 348–353.
T. Jongen, Characterization of batch mixers using numerical flow simulations – AIChE Journal Vol. 46 No. 11 (2000) 2140-2150.
A.W. Patwardhan and J.B. Joshi, Relation between flow pattern and blending in stirred tanks - Ind. Eng. Chem. Res. 38 (1999), pp. 3131–3143.
A.K. Sahu, P. Kumar, A.W. Patwardhan and J.B. Joshi, CFD modeling and mixing in stirred tanks - Chem. Eng. Sci. 54 (1999), pp. 2285–2293.
M. Rahimi, The effect of impellers layout on mixing time in a large-scale crude oil storage tank – Journal of Petroleum Science and Engineering 46 (2005) 161-170.
A. Brucato, M. Ciafalo, F. Grisafi and G. Micale, Numerical prediction of flow fields in baffled stirred vessels—a comparison of alternative modelling approaches - Chem. Eng. Sci. 53 (1998) 3653–3684.

AUTHORS BIOGRAPHY

Davide MARCHINI is scholarship holder in Industrial Engineering at the University of Parma (Interdepartmental Center Siteia.Parma). He got a master degree in Mechanical Engineering for the Food Industry, discussing a thesis titled: "Advanced design of a UV reactor for water treatment using computational fluid dynamics". He attended the 10th International Conference on Modeling and Applied Simulation (Rome, 12-14 September 2011), presenting the paper titled "Advanced design of a static dryer for pasta with simulation tools", and the 2011 EFFoST Annual Meeting (Berlin 9-11 November 2011), presenting the paper titled "Advanced design of a UV reactor for water treatment using computational fluid dynamics".

Federico SOLARI is a PhD Student in Industrial Engineering at the University of Parma; master degree in Mechanical Engineering of the Food Industry, dissertation of the thesis: "Analysis and design of a plant for the extraction of volatile compounds from aqueous matrix"; Attending many international conferences, he's author or coauthor of 7 international papers. His research activities mainly concern industrial plants and food process modeling and simulation, with a particular focus on the CFD simulation for the advanced design of food plants.

Mattia ARMENZONI is a scholarship holder in Industrial Engineering at the University of Parma (Interdepartmental Center Siteia.Parma). He got a master degree in Mechanical Engineering for the Food Industry, discussing a thesis titled: "Advanced design of a static dryer for pasta with simulation tools". He attended the 10th International Conference on Modeling and Applied Simulation (Rome, 12-14 September 2011) presenting the paper titled "Advanced design of a static dryer for pasta with simulation tools", and the 2011 EFFoST Annual Meeting (Berlin 9-11 November 2011) presenting the paper titled "Advanced design of a UV reactor for water treatment using computational fluid dynamics".

Roberto MONTANARI is full professor of Mechanical Plants at the University of Parma (Italy). He graduated (with distinction) in 1999 in Mechanical Engineering at the University of Parma. His research activities mainly concern equipment maintenance, power plants, food plants, logistics, supply chain management, supply chain modelling and simulation, inventory management. He has published his research in approx. 60 papers, which appear in qualified international journals and conferences. He acts as a referee for several scientific journals and is editorial board member of 2 international scientific journals.

Marta RINALDI is a PhD Student at the University of Parma, and Scholarship Holder in Industrial Engineering at the Interdepartmental Center CIPACK. She got a master degree in Management Engineering with a thesis titled "Analysis and evaluation of energy efficiency of a shrinkwrap-packer". Her main fields of research are discrete event simulation and simulation of industrial plants.

Eleonora BOTTANI is Lecturer (with tenure) in Mechanical industrial plants at the Department of Industrial Engineering of the University of Parma (Italy). She graduated in 2002 in Industrial Engineering and Management at the University of Parma, and got her Ph.D. in Industrial Engineering in 2006. Her research activities concern logistics and supply chain management issues, encompassing intermodal transportation, development of methodologies for supplier selection, analysis and optimization of supply chains, supply chain agility, supply chain modelling and performance analysis, and, recently, the impact of RFID technology on the optimization of logistics processes and supply chain dynamics. She is the author or coauthor of more than 80 scientific papers, referee for more than 40 international scientific journals, editorial board member of 2 scientific journals and Associate Editor for one of those journals.

Gino FERRETTI is full professor of Mechanical Plants at the University of Parma (Italy). He graduated in Mechanical Engineering in 1974 at the University of Bologna, where he served as assistant professor for the courses of “Machines” and “Mechanical Plants”. He worked as associate professor at the University of Padua and as full professor at the University of Trento. In 1988, he moved to the Faculty of Engineering, University of Parma, where at present he is full professor of Mechanical Plants. His research activities focuses on industrial plants, material handling systems, and food processing plants, and have been published in numerous journal and conference papers.

Giuseppe VIGNALI: graduated in 2004 in Mechanical Engineering at the University of Parma. In 2009, he received his PhD in Industrial Engineering at the same university, related to the analysis and optimization of food processes. Since August 2007, he works as a Lecturer at the Department of Industrial Engineering of the University of Parma, and, since the employment at the university, he has been teaching materials, technologies and equipment for food packaging to the food industry engineering class. His research activities concern food processing and packaging issues and safety/security of industrial plant. Results of his studies related to the above topics have been published in more than 20 scientific papers, some of which appear both in national and international journals, as well in national and international conferences. He acts also as a referee for some international journals, such as Prevention Today, Facilities.

AUTONOMOUS CONTROL IN EVENT LOGISTICS

Harjes, F.[a], Scholz-Reiter, B.[b]

[a,b]BIBA – Bremer Institut für Produktion und Logistik GmbH an der Universität Bremen, Hochschulring 20, 28359 Bremen

[a]haj@biba.uni-bremen.de, [b]bsr@biba.uni-bremen.de

ABSTRACT

The conflict between economic interests and order-related requirements complicates the scheduling and control of orders in the field of event logistics. Often, dynamic influences, such as rush orders, thefts or damage to material and equipment, require an adaptive replanning of events, resources and transport routes. In these cases, it is very difficult to determine the optimal trade-off between the utilization of transport devices, the adherence to due dates and customer wishes. This paper introduces a concept for the implementation of autonomous control in event logistics. At this, the focus lays on the three key aspects of autonomously controlled systems; the modelling process, the object representation and the structuring of the communication processes. A use case illustrates the starting points of the implementation.

Keywords: event logistics, autonomous control, decentralised decision making, modelling and simulation

1. INTRODUCTION

The management of public events, such as concerts, company anniversaries and so on, involves high customer requirements concerning the adherence to due dates, the flexibility, cost-effectiveness and technical reliability. This applies both for the scheduling and control of events and the related logistic processes. The efficient execution of these processes often implies a conflict between the order- or event-oriented requirements and economic motives. At this, the optimal utilization of transport capacities reduces the mobility of equipment, such as stages, speakers, headlights and so on and therefore complicates a dynamic replanning.

These aspects further increase, when a close temporal sequence of events makes a return of the equipment to storage impossible. This results in a manual disposal of equipment for subsequent events directly at a venue. In combination with dynamic effects, such as damages or thefts, the consequences are multitude, often inefficient and underemployed transports with corresponding costs and time exposure.

From a scientific point of view, this problem constitutes a combination of an event-oriented scheduling (Gudehus 2006) and a Dynamic Multi Vehicle PDPTW (Pick-up and Delivery Problem with Time Window) (Parragh 2008).

With regard to these two sub-problems, the current state of the art may be summarised as follows. In the field of event-oriented scheduling, exiting approaches are already able to handle the problem of resource allocation within logistic networks satisfactorily with regard to the dispatching of articles, the vehicle utilization and -order (Gudehus 2006). However, most of the available approaches consider central planning processes and the related structures of information acquisition and processing. The desired application in the field of event logistics requires a scheduling procedure that is able to cope with dynamically changing conditions and constrains in decentralised structures.

The question for the optimal or best possible route defines a NP-hard problem, which is often referred to as the Traveling Salesman Problem (TSP) in operations research(Applegate 2006). It is possible to find optimal solutions for a slight TSP, but the required computing time often limits the practicality. Therefore, heuristics came into operation to find approximately optimal solutions for larger cases of application. Unfortunately, they are not able to guarantee the optimality of the solution(Applegate 2006).

Further enhancements of the TSP consider multiple vehicles, time windows, and an incomplete list of destinations at the departure as well as restrictions of the transport capacity. These problem class is called a Vehicle Routing Problem (VRP)(Parragh 2008). Depending on the kind of the considered restrictions, it is possible to distinguish between different versions of the problem, such as the dynamic VRP(Larsen 2000). In order to stay applicable in practice, these versions often consider a limited number of restrictions and/or target functions(Fabri 2006; Gendreau 2006).

As the scheduling and control of events require a consideration of both the real-world dynamic and manifold individual restrictions and target functions for the logistic objects, the adaption of the existing heuristics is difficult. The application of methods from the field of autonomous control seems to be a promising approach to cope with this problem.

The paradigm of autonomous control denotes a decentralised decision making of autonomous logistics objects in heterarchic structures (Windt 2007). At this, the central planning and disposal shifts to a distributed and flexible proceeding, where the decision-making falls to single objects such as transport vehicles, goods, and so on.

This paper focuses on a concept regarding the implementation of autonomous control in the scheduling and transport processes in event logistics. The objective is an autonomously controlled system that optimizes the resource allocation and makes the considered logistic processes more robust at the same time. The conception and operation of the system are based on specialised modelling and simulation techniques for autonomously controlled processes. A SME (small or medium enterprise) from the field of event logistics serves as a use case.

The structure is as follows. Section 2 gives an overview of autonomous control in general and introduces the basics of event logistics. Section 3 deals with the use case, the considered processes and the related weak points. The following section 4 describes the implementation concept, before the paper finishes with a short summary and outlook in section 5.

2. BASICS

2.1. Autonomous control

Today`s logistic processes face an increase in dynamic and complexity(Scholz-Reiter 2004). As established production planning and control systems reach their limits, new concepts on the basis of technologies such as RFID (Radio Frequency Identification) or GPS (Global Positioning System) came into focus. Autonomous control combines these technologies and related methods to shift from a centralised planning and control to a decentralised decision-making of autonomous objects(Windt 2007).

Within autonomously controlled systems, single objects have the ability and possibility for independent decisions. For this purpose, every object is equipped with the necessary technical requirements to detect its own position, to interact with other objects within or outside the system and to make own decisions following individual targets(Windt 2008). The main objective of autonomous control is the improvement of the overall system`s robustness by enabling a flexible and distributed handling of dynamic and complexity(Windt 2007).

2.2. Event logistics

In this paper, event logistics comprises the logistic processes related to the planning and execution of company anniversaries, concerts, festivals, public performances, fairs, and so on. Often, event logistics is embedded into several phases of event management (Allen 2008). Typically, the event execution follows a multi-stage planning of organisational and artistic aspects. Figure 1 depicts an exemplary overview of the procedure that is inspired by an example use case.

The procedure consists of five phases, ranging from the rough planning after the order receipt to the concrete event execution directly at the venue. In phase one, the determination of the event parameters and the related services takes place. In general, a project meeting with the customer specifies the individual wishes, while an inspection of the event location contributes technical and local restrictions.

In phase two, the event management company develops a concept, including the required equipment, services and logistics. If the customer agrees, the order is finally confirmed and the detailed planning begins (phase three).

Phase four comprises the execution of the logistic planning including the personnel and equipment allocation as well as possible leasing orders. Finally, the last phase five covers the event accomplishment from the warehouse exit over the assembly and dismantling of the equipment back to the warehouse entry. The final billing completes the process. For reasons of simplification, this phase is not further mentioned.

The processes, relevant to this paper, mainly take place in the phases three and five. Generally, they comprise the transport planning and execution for equipment between one or more warehouses, belonging to the event organiser or a subcontractor, and the venue.

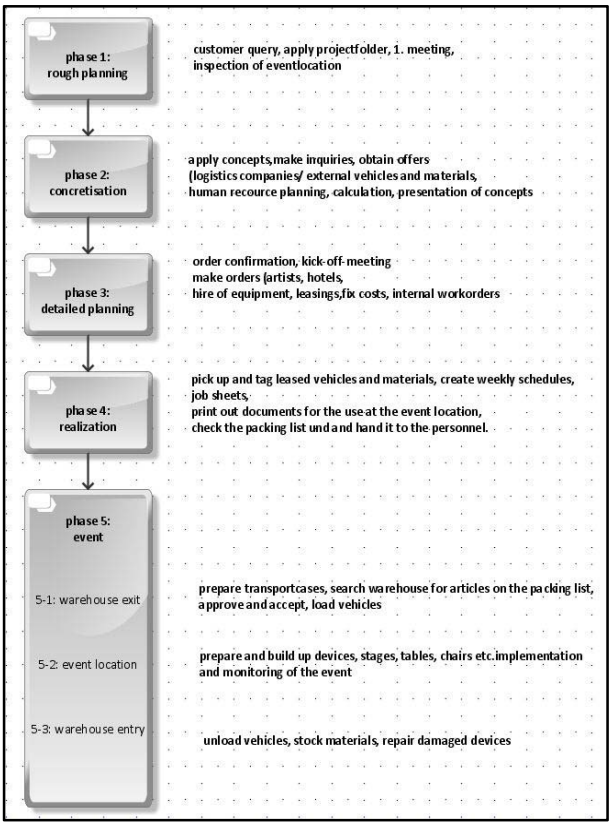


Figure 1: Exemplary Procedure of Event Management (own depiction)

3. USE CASE

In the following, a use case shall act as an example scenario for the application of autonomous control in event logistics. The use case considers a full-service-agency from the branch of event marketing. With 60 employees and an annual turnover of ca. 7 million €, the agency constitutes a typical SME (small or medium enterprise). The main business segment is the letting of event related equipment, reaching from chairs and cloak hangers over stage elements up to electronic devices.

The related services comprise the provision, construction and dismantling of equipment at the venues, including the logistics. For the latter, a car pool consisting of a lorry and several compact vans comes into operation. Further, the agency operates a central storage for the equipment. If required, additional vehicles and equipment are hired.

Figure 2 gives an overview of an event accomplishment. In general, every order represents an event and is linked to a material list, depending on the event`s requirements. In the following, the required equipment is put together in accordance to the list and allocated to a suitable transport vehicle. After the route planning, the transport leaves the storage and delivers the ordered equipment directly to the venue. Subsequently, the unloading and construction takes place. After the event, the procedure happens similar in the opposite direction.

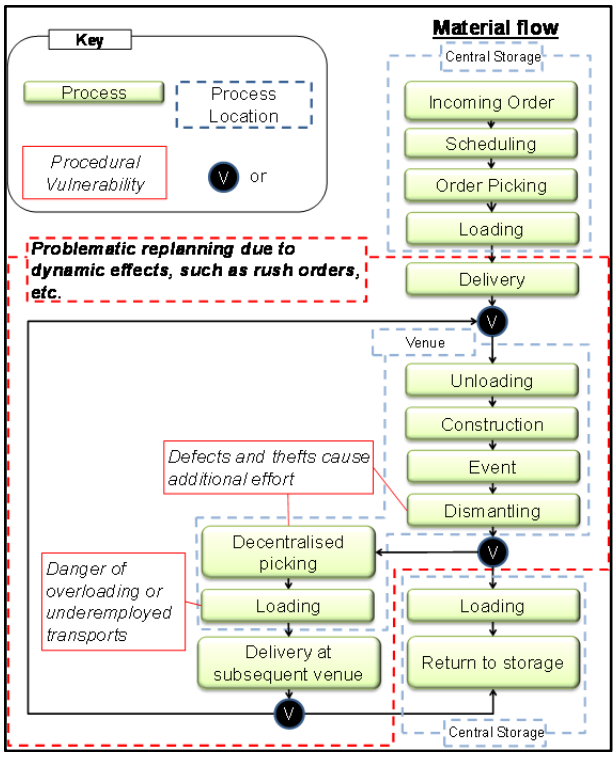


Figure 2: General Process Flow

Subsequent events or rush orders often require a deviant proceeding. If the following event begins with a close temporal distance, a decentralised planning becomes necessary. Figure 3 depicts two such cases of decentralised (re-)planning (cases 2 and 3) as well as a single event (case 1). In case 2, the available equipment has to be split up into one transport to the storage and one to the subsequent venue. This either results in two underemployed transports or implies the inclusion of the subsequent venue as a stopover into the way back direction. The latter implies a transport of equipment that is unnecessary for the specific event. Additionally, this material is not available for other events during this time and dynamic effects, such as thefts or defects, further complicate the proceeding. Case 3 depicts an uncomplicated sequence of events which have no time conflict and require the same set of equipment.

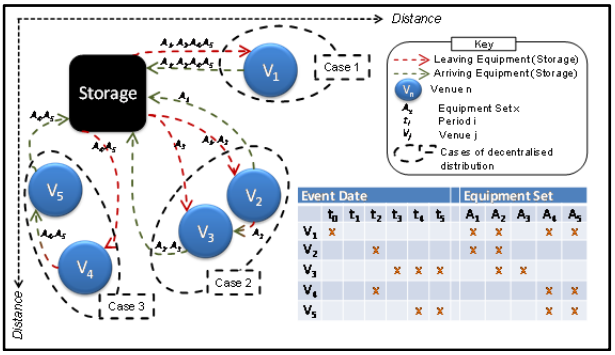


Figure 3: Cases of Decentralised Planning

Currently, a central planning system is responsible for the allocation of all resources. Especially the rush orders and the related decentralised replanning at a venue for subsequent events brings this system to its limits. This results in unstable planning and scheduling processes.

Besides the dynamically occurring external influences, further internal problems regarding the data transparency and availability reduce the process reliability. In the current state, RFID-gates gather loading processes only, when transport vehicles leave or enter the storage. Loading or unloading at the different venues is not recorded automatically, so that the position of equipment is often unclear, until it returns to storage again.

Summarized, the weak points of the current processes are the insufficient handling of dynamic effects due to the centralised approach and the lack of up-to date information concerning the position of the equipment after leaving the storage.

4. CONCEPT

In order to cope with the problems, this paper suggests the implementation of autonomous control for the logistic processes of the example SME. Autonomous control in logistics systems is characterised by the ability of logistic objects to process information, to render and to execute decisions on their own(Windt 2008). Due to this definition, autonomously controlled processes require the representation of the involved entities as autonomous objects with their belonging knowledge, abilities and objectives.

Therefore, the identification of the relevant objects is the first step of the implementation. With regard to

the processes of the example SME, the car pool and the event equipment are from central interest. In the current centralised planning system, both classes of objects only constitute allocable resources. Within the planned autonomous control system, all those objects are capable to act both independently and in cooperation.
As this approach differs from the traditional perspective, the modelling process requires an adopted methodology. The design of the presented work bases on the Autonomous Logistics Engineering Methodology (ALEM). ALEM is a multipart framework for modelling autonomously controlled processes in logistic systems(Scholz-Reiter 2009). It comprises three components, each covering a special aspect of the modelling process.

ALEM-N (ALEM-Notation) defines a view concept for the representation of specific aspects of the modelled logistic system. It further provides the notational elements and their meaning within the framework. ALEM-P (ALEM-Procedure) describes the steps of the modelling process and acts as a guideline for the analysis and specification of the intended logistic system. ALEM-T (ALEM-Tool) combines both components into a software tool and adds a reference that enables a reuse of existing models(Scholz-Reiter 2009).

The result of the modelling process with ALEM is a system representation, where every object is defined as class with belonging features, abilities and knowledge. Figure 4 shows a simplified description of the transport vehicles as autonomous objects.

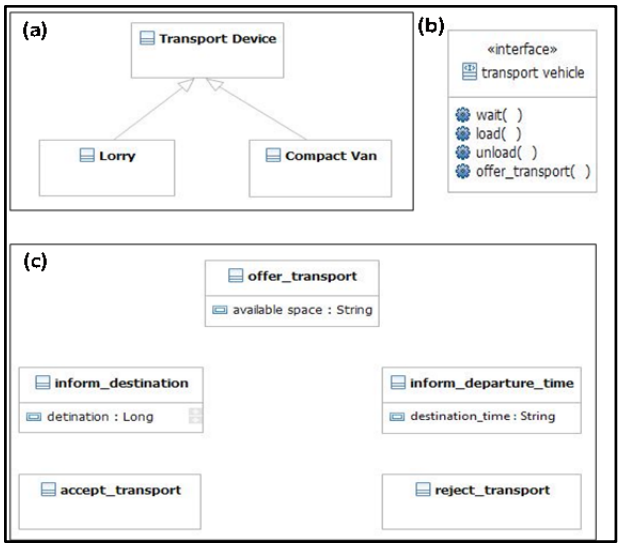


Figure 4: Transport Vehicles as Autonomous Objects (simplified)

Part (a) of the figure depicts the ALEM class view, where transport vehicle defines a general class. The classes lorry and compact van define more specific derivations. Part (b) shows the general abilities that all transport vehicles share. For example, all vehicles can wait for an order, load or unload equipment or offer a transport. The communication view on part (c) defines an excerpt of the messages, transport vehicles can exchange within negotiations with other objects. The modelling of other objects follows a similar procedure.

The second step of the implementation focusses on the technical aspects of the distributed decision-making process. In general, two approaches exist. They differ in the underlying technical infrastructure. The first possibility is to equip every single object with the required technology for positioning, communication and decision making. This procedure is suitable, if all objects are physically large enough and the amount of objects to equip is not to large. If the objects are too small to include the required devices or so numerous, that an individual equipping would be very sophisticated and expensive, a client-server architecture could be suitable.

This architecture divides the information acquisition and the decision-making processes spatially. The information acquisition takes place in a decentralised manner, while a multi-agent based simulation (MAS) on a central server takes over the decision-making processes. At this, every object is capable to locate itself and to send the related information to the central server. Within the MAS-system, every software agent represents one individual autonomous object. The agent performs the decision-making for this object depending on the locally acquired information and the corresponding objectives. He is further able to negotiate and cooperate with the other agents.

For the use case, an adapted version of the client-server infrastructure comes into operation. As the event-related processes often take place in a hurry, the treatment of the equipment is commonly not very careful. Thus, adding sensitive devices to every single autonomous object is not advisable, although the majority of the objects is large enough and they are generally not to numerous. As a middle course, the transport devices act as a kind of information hub. They will be equipped with a GPS-system for positioning, an RFID-reader for the identification of the loaded objects, and an UMTS-device (Universal Mobile Telecommunications System) for the data transfer. Correspondingly, every autonomous object carries a RFID-transponder that allows a clear identification. By this means, the transport vehicles ensure that the required positioning information for the software agents is available at all times. The objectives and therefore the target function for every object are derived from the order database. Relevant is for example the due date of the event or transport-related information, such as weight and size.

For the implementation of the MAS-system, the PlaSMA approach (Platform for Simulations with Multiple Agents) finds a use. PlaSMA is a method for the evaluation of autonomously controlled logistic processes by simulations of multi-agent systems (Warden 2010). Technically, PlaSMA bases on the Java Agent Development Framework (JADE) (Applegate 2006). Within PlaSMA, an individual software agent represents every autonomous system element and

interacts representatively with other agents in the system. At this, the agent`s decisions base upon available situation specific information and follow an individual target function (Gehrke 2010).

In order to use the PlaSMA-Simulation as a part of a control method, the simulation has to run permanently after the initialisation. During the initialisation phase, a world model defines the physical basics of the simulation. This comprises information about the relevant physical properties and elements, such as streets, places, distances and so on. This information is later used for the route planning. Further, number and kind of transport devices and event equipment enhances the initial database and determines the generation of the required agents.

Within the running simulation, it is possible to add new elements, such as additional equipment, transport devices, orders and so on dynamically. Only changes related to the physical world model require a restart of the simulation. This is for example the case for new streets.

Following the principles of autonomous control, the results of the simulation agents' negotiations constitute planning decisions. As the PlaSMA approach implements a time-discrete procedure, it is necessary to write back the corresponding data sets periodically, so that the corresponding processes for the execution can take place.

The third step of the implementation focusses on the route planning and the corresponding communication procedures within the logistic system. To enable individual routing decisions for every object, the DLRP (Distributed Logistic Routing Protocol) comes into operation. The DLRP focuses on the autonomous routing of logistic objects through dynamic logistic systems (Rekersbrink 2009). Its fundamental functionality is derived from established data routing protocols in decentralised communication networks, such as the internet or cell phone networks(Scholz-Reiter 2006). For this, the protocol provides communication standards and procedures for the collaboration between transport vehicles, commodities and logistic hubs(Rekersbrink 2009).

5. SUMMARY AND OUTLOOK

This paper introduces the technical aspects of a concept for the implementation of autonomous control in event logistics. At this, the current contribution focusses on the modelling process, the representation of the involved objects as autonomous entities and the structuring of the communication and cooperation between these objects.

As the concept currently addresses the technical implementation, future work will concentrate on the underlying methods for planning and scheduling in detail. At this, the control strategy for the objects will be from central interest.

In combination, the technical and methodical aspects aim to the evaluation of the general applicability of autonomous control for the dynamically influenced dispatching of circulation rental articles. Furthermore, a main motivation is to improve the performance and robustness in dynamic logistic systems with manifold restrictions and changing transport nodes.

ACKNOWLEDGMENTS

This research is supported by the German Research Foundation (DFG) as part of the Collaborative Research Centre 637 "Autonomous Cooperating Logistic Processes– A Paradigm Shift and its Limitations" at the University of Bremen.

REFERENCES

Allen, J., O`Toole, W., Harris, R., McDonnell, I. (2008). Festival and special event management. Milton, Australia, John Wiley & Sons Australia, Ltd.

Applegate, D., Bixby, R., Chvátal, V., Cook, W. (2006). The Traveling Saleman Problem: A Computational Study. Princeton, Princeton University Presss.

Fabri, A., Recht, P. (2006). "On Dynamic Pickup and Delivery Vehicle Routing with Several Time Windows and Waiting Times." Transportation Research Part B 40(4): 335-350.

Gehrke, J. D., Herzog, H., Langer, H., Malaka, R., Porzel, R., Warden, T. (2010). "An Agent-based Approach to Autonomous Logistic Processes - Colloborative Research Centre 637: Autonomous Cooperating Logistic Processes." Künstliche Intelligenz 24: 137-141.

Gendreau, M., Guertin, F., Povton, J.-Y., Séguin, R. (2006). "Neighborhood Search Heuristics for a Dynamic Vehicle Dispatching Problem with Pick-Up and Deliveries." Transportation Research Part C 14(3): 157-174.

Gudehus, T. (2006). Dynamische Disposition. Berlin, Springer Verlag.

Larsen, A. (2000). The Dynamic Vehicle Routing Problem Dissertation, Technical University of Denmark

Parragh, S., Doerner, K., Hartl, R. (2008). "A survey on pickup and delivery problems. Part II: Transportation between pickup and delivery locations." Journal für Betriebswirtschaft 58: 81-117.

Rekersbrink, H., Makuschewitz, T., Scholz-Reiter, B. (2009). "A distributed routing concept for vehicle routing problems." Logistics Research 1(1): 45-52.

Scholz-Reiter, B., Hildebrandt, T. (2009). ALEM-T: A Modelling Tool for Autonomous Logistic Processes. 40th CIRP Seminar on Manufacturing Systems, Liverpool, Department of Engineering, University of Liverpool.

Scholz-Reiter, B., Kolditz, J., Hildebrandt, T. (2009). "Engineering autonomously controlled logistic systems." International Journal of Production Research 47(6): 1449-1468.

Scholz-Reiter, B., Rekersbrink, H., Freitag, M. (2006). Internet routing protocols as an autonomous

control approach for transport networks. 5th CIRP international seminar on intelligent computation in manufacturing engineering.
Scholz-Reiter, B., Windt, K., Freitag, M. (2004). Autonomous Logistic Processes – New Demands and First Approaches. 37th CIRP International Seminar on Manufacturing Systems, Budapest, Hungary, Computer and Automation Research Institute, Hungarian Academy of Science.
Warden, T., Porzel, R., Gehrke, J., Herzog, O., Langer, H., Malaka, R. (2010). Towards Ontology-based Mulitagent Simulations: The Plasma Approach. 24th European Conference on Modeling and Simulation (ECMS 2010), European Council for Modeling and Simulation.
Windt, K., Böse, F., Phillip, T. (2008). "Autonomy in production logistics: Identification, characterisation and application." Robotics and Computer-Integrated Manufacturing 24(4): 572–578.
Windt, K., Hülsmann, M. (2007). Changing Paradigms in Logistics - Understanding the Shift from Conventional Control to Autonomous Cooperation & Control The Impact of Autononomy on Management, Information, Communication, and Material Flow. K. Windt, Hülsmann, M. Berlin, Springer: 4-16.

AUTHORS BIOGRAPHY

Dipl.-Inf. Florian Harjes, born in 1981, is a scientific research assistant at the Bremer Institut für Produktion und Logistik GmbH (BIBA) at the University of Bremen.

He received a diploma in computer science from the University Bremen in 2008, where he pursued his thesis “Exact synthesis of multiplexor circuits” at the same year. During this time, he developed a tool for the automated synthesis of minimal multiplexor circuits for a corresponding Boolean function.

In BIBA, **Dipl.-Inf. Florian Harjes** was in charge of long time simulations of neural networks and the development of a hybrid architecture for the continuous learning of neural networks in production control between 2009 and 2012. Since the beginning of 2012 he works on a project regarding the autonomously controlled dispatching of rental articles in the field of event management.

Prof. Dr.-Ing. Bernd Scholz Reiter is managing director of the Bremer Institut für Produktion und Logistik GmbH at the University of Bremen (BIBA) and head of the research center “Intelligent Production and Logistics Systems (IPS)”.

Born in 1957, he studied Industrial Engineering and Management at the Technical University of Berlin. After his doctorate in 1990 he was an IBM World Trade Post-Doctoral Fellow at the IBM T.J. Watson Research Center, Yorktown Heights, NY, USA, in Manufacturing Research until the end of 1991. Subsequently, he worked as a research assistant at the Technical University of Berlin and in 1994 was appointed to the chair of Industrial Information Technology at the Brandenburg Technical University of Cottbus. From 1998 to 2000, he was head and founder of the Fraunhofer Application Center for Logistics Systems Planning and Information Systems in Cottbus, Germany. Since 2000 he heads the chair of Planning and Control of Production Systems in the Department of Manufacturing Engineering at the University of Bremen. Since 2002 he also serves as Managing Director of the BIBA – Bremer Institut für Produktion und Logistik GmbH (BIBA) at the University of Bremen. At the BIBA, Prof. Scholz-Reiter works in applied and industrial contract research.

Between July 2007 and 2012 **Bernd Scholz-Reiter** was Vice President of the German Research Foundation (DFG). Since August 2011 he serves as fellow in the International Academy for Production Engineering (CIRP).

Prof. Scholz-Reiter is a full member of the German Academy of Engineering Sciences and of the Berlin-Brandenburg Academy of Sciences, Associate Member of the International Academy for Production Engineering (CIRP), member of the Scientific Society of Manufacturing Engineering, member of the European Academy of Industrial Management and a member of the Advisory Commission of the Schlesinger Laboratory for Automated Assembly at the Technion - Israel Institute of Technology, Haifa, Israel.

SUPPLY CHAIN SIMULATION: A STUDY ON REORDER POLICIES FOR PERISHABLE FOOD PRODUCTS

Marta Rinaldi[(a)], Eleonora Bottani[(b)], Gino Ferretti[(c)], Mattia Armenzoni[(d)], Davide Marchini[(e)], Federico Solari[(f)], Giuseppe Vignali[(g)] and Roberto Montanari[(h)]

[(a)] CIPACK Interdepartmental Centre, University of Parma, viale G.P.Usberti 181/A, 43124 Parma (Italy)
[(b), (c), (f), (g), (h)] Department of Industrial Engineering, University of Parma, viale G.P.Usberti 181/A, 43124 Parma (Italy)
[(d), (e)] SITEIA Interdepartmental Centre, University of Parma, viale G.P.Usberti 181/A, 43124 Parma (Italy)

[(a)]marta.rinaldi@unipr.it, [(b)]eleonora.bottani@unipr.it, [(c)]gino.ferretti@unipr.it, [(d)]mattia.armenzoni@unipr.it, [(e)]davide.marchini@unipr.it, [(f)]federico.solari@unipr.it, [(g)]giuseppe.vignali@unipr.it, [(h)]roberto.montanari@unipr.it

ABSTRACT

In this paper, we analyze three traditional reorder policies (i.e., EOI, EOQ and S,s) applied to 5 different products, in a 2-echelon supply chain. A particular attention is placed to fresh products with limited shelf life, and to the suitability of applying the inventory management policies to those products.

An *ad hoc* simulation model, reproducing the reorder process of the two supply chain players, is developed under MS Excel™ and used to simulate the low of the different products along the supply chain, according to the different reorder policies. From the simulation, the minimum cost setting is derived for all policies, together with additional performance parameters (e.g. the throughput time of items along the supply chain), which allow assessing the suitability of a reorder policy for a given product.

From the simulation outcomes, some guidelines are derived for the optimal inventory management of perishable products. Since the supply chain and product data are derived from a real scenario, it is expected that our outcomes and guidelines are of practical usefulness to inventory managers.

Keywords: inventory management, inventory policies, perishable products, simulation model.

1. INTRODUCTION

A main goals of supply chain management is to maximize customer satisfaction, at the same time optimizing demand planning and management, resource use, integration between supply and demand, and stock levels. As regards this latter point, stocks are required at any level of the supply chain, to provide a buffer between uncertain supply and demand. A proper inventory management policy is thus expected to provide uninterrupted material and product flow throughout the supply chain at the minimum cost (Waters, 2003). Inventory management has obvious impact on the supply chain efficiency, since it generates several cost components, namely the purchasing cost of items, the order cost, the inventory holding cost and the stock-out cost (Bottani and Montanari, 2011).

Inventory management models can focus either on a single-period problem, which is also known as the newsvendor problem, or on a multi-period problem. In the former case, the goal is to find the order quantity which maximizes the expected profit in a single period probabilistic demand framework (Abdel-Malek and Montanari, 2005a,b). For multi-period problems, which are the focus of this paper, specific policies, such as economic order quantity (EOQ) or economic order interval (EOI), were developed with the purpose of achieving a proper balance between the different cost components (Waters, 2003). More recently, the (S,s) policy, i.e. a periodic review policy with re-order point and order-up-to level, has been introduced as a combination of the inventory control policies mentioned above, and is currently adopted in many contexts (Silver et al., 2009). According to that policy, the stock of an item is examined at periodic review intervals, and, if the inventory position is found to be lower than reorder point *s*, an order is placed. The quantity ordered should allow raising the current stock to the order-up-to level *S*. For computational purpose, the inventory position of an item consists of the on-hand inventory plus the ordered quantities, excluding backordered quantities.

The role of inventory management in matching supply and demand and getting the right product in the right place at the right time is particularly crucial for perishable products (Deniz et al., 2004). Indeed, for those products, the economic value deteriorates significantly over time, due to the limited product shelf life (Blackburn and Scudder, 2009). This could generate further costs of shrinkage, spoilage or obsolescence (Deniz et al., 2004).

Research related to inventory management mainly focuses on the optimal determination of the control parameters of inventory policies, or on their optimality under particular operating conditions. Examples of such

studies include Schneider and Ringuest, (1990), and Silver et al., (2009), for the (S,s) policy, or Ferguson et al., (2007) and Goyal (1985) for the EOQ and EOI policies respectively. Some studies for the continuous review perishable inventory models are Weiss (1980), Schmidt and Nahmias (1985), Ravichandran (1995) and Liu and Lian (1999), while available reviews of inventory models have been performed by Raafat (1991) and Nahmias (1982).

However, the studies mentioned above suffer from two main limitations. First, they concern a specific inventory management policy (i.e., either EOI, EOQ or S,s) , while comparisons among those policies are rarely available in literature. Second, the inventory management policy is examined with respect to a specific supply chain player, without considering a whole supply chain.

Our goal with this study is to compare the use of different inventory policies to manage the stock level in a supply chain of perishable goods, with the ultimate purpose of identifying the optimal policy as a function of the product examined. The policies considered and compared are EOI, EOQ and (S,s), which are the traditional reorder policies proposed in literature. We examine a supply chain composed of a distribution center, a retail store and the final customer. The reorder process of 5 perishable products, with different characteristics, is reproduced and optimized, by means of a simulation model developed under MS Excel™.

The remainder of the paper is organized as follows. In the next section, we provide an overview of the supply chain examined and of the data collection phase. In section 3, we describe the development of the simulation tool and we detail the simulations The main results, in terms of the optimal inventory management policy as a function of the product considered and of the supply chain player, are proposed and discussed in section 4. In the last section, we summarize the key findings of the study and indicate future research directions.

2. THE SUPPLY CHAIN EXAMINED

The present study examines a real supply chain, composed of a distribution center, a retail store and the final customer (see Figure 1).

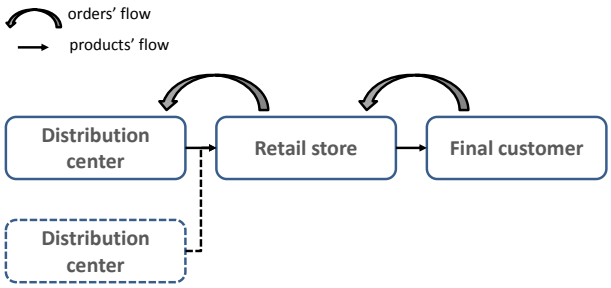


Figure 1: the supply chain examined

The distribution center considered is a warehouse of a main Italian retailer, and is located in northern Italy, near Reggio Emilia. It is specialized in the distribution of perishable products, such as fish, fruits, vegetables, milk derivatives and dairy products. Overall, the distribution center handles approx. 800 different items per day.

The retail store is a hypermarket, located near Mantova (Italy). It receives fresh products from the distribution center mentioned above, while the remaining product categories (e.g. grocery, beverage, health and beauty care, frozen foods, paper) are supplied by a different distribution center (dashed line), which is not considered in this study.

With respect to the perishable products, the current inventory management of the retail store and of the distribution center is as follows. The retail store places orders to the distribution center according to a proposal formulated by the company's information system, taking into account the final customer's demand. Typically, the product ordered are supplied within one or two days. The retail store does not own a refrigerated warehouse, so that the product received are immediately located on the store shelves. In the case the fresh product is not immediately sold, and its shelf life is expiring, the distribution center will pick up the product from the retail store, to send it to an alternative channel before it expires.

The distribution center receives the orders from several retail stores (besides the one considered in this study) and fulfills them primarily by using the available stock. Moreover, orders are collected to derive an aggregated demand, which will be processed by the warehouse management system to compute the product quantity to be ordered to manufacturers. Product ordered are available within some days; the specific lead time varies depending on the product considered.

Both the retail store and the distribution center were visited with, and people in charge were interviewed to derive the relevant pieces of information related to the current inventory management process. Moreover, store's managers were asked to indicate 5 perishable products, with different characteristics, whose reorder process could be investigated and optimized through simulation. The products chosen for the analysis, as well as their relevant data, are provided in Table 1. The following nomenclature is used for products: 1-Fresh milk; 2-Mozzarella cheese; 3-Yoghurt; 4-Plum pulp; 5-Royal jelly.

As can be seen from Table 1, for each product we collected data related to the total product shelf life, which is shared between the distribution center (for approx. 1/3) and the retail store (for 2/3). The "residual shelf life for picking up" represents the residual shelf life of the product which is collected from the retail store, to be sent to different channels.

Table 1: the products investigated

Product	**1**	**2**	**3**	**4**	**5**
Total shelf life [days]	7	19	35	365	720
Shelf life for distribution center [days]	2	5	8	80	158
Shelf life for retail store [days]	3	9	15	163	322

Residual shelf life for picking up [days]	2	3	3	5	10
Procurement lead time for distribution center [days]	0	1	1	5	8
Number of deliveries per week [delivery/week]	6	5	4	1	1
Procurement lead time for retail store [days]	1	2	2	2	2
Maximum number of items on the shelf [items]	70	54	21	12	24
Number of items per case [items/case]	6	-	6	-	-

Further data collected from the interviews with the distribution center and retail store's managers refer to holding cost, stock-out cost and order cost of the product examined. They are proposed in Table 2. As can be seen from that table, the order cost is the same for all products considered, since it only depends on the cost of manpower dedicated to the order process, while it is not affected by the product type. The stock-out cost and the cost of holding stock, instead, are specific for the product considered in the case of the retail store, while the same cost is assumed for the distribution center. Indeed, the cost of holding stock for the distribution center is specific for a given product category, since the product category may generate different cost components (e.g. energy, facility maintenance, or plant amortization). As the product considered in this study belong to the same category (i.e., fresh products), the same holding cost results. Finally, the stock-out cost is derived by multiplying the cost of stocks by a suggested factor of 20.

Table 2: unitary costs used in the simulation

Product	**1**	**2**	**3**	**4**	**5**
stock-out cost for distribution center [€/case/day]	0.44	0.44	0.44	0.44	0.44
stock-out cost for retail store [€/item/day]	2.38	2.66	3.08	3.48	2.32
cost of holding stock for distribution center [€/day/case]	0.022	0.022	0.022	0.022	0.022
cost of holding stock for retail store [€/day/item]	0.0119	0.0133	0.0154	0.0174	0.0116
order cost for distribution center [€/order]	6.8	6.8	6.8	6.8	6.8
order cost for retail store [€/order]	0.954	0.954	0.954	0.954	0.954

3. SOFTWARE MODELLING

3.1. Overview

As previously mentioned, multi-period inventory policies for perishable goods (i.e., EOI, EOQ and S,s) were studied using a simulation model, developed under MS Excel™.

Specifically, discrete event simulation has been used to reproduce the reorder process of the supply chain, in order to obtain a sufficient amount of observations on the performance of the system.

The simulation model consists of MS Excel™ file which reproduces the flow of orders related of a given product, according to a specific reorder policy. Both supply chain players (i.e., retailer and distributor) are considered in the file. For each product, we simulated the use of the three reorder policies, except for fresh milk. Indeed, from the interviews carried out with the store representatives, it emerged that the EOQ policy does not seem to be suitable to be used for that product, given its very short shelf life. Overall, we thus have 3 (policies) x 5 (products) – 1 = 14 MS Excel™ simulation files. Each simulation file reproduces, by means of as many spreadsheets, the two supply chain players.

3.2. Reorder policy settings

We start by simulating the flow of orders and products of the retail store, starting from the (known) customer's demand. For each product, this latter was generated as a random variable, whose parameters were suggested by the retail store representatives on the basis of the daily sales of the product. The simulation of retail store and distribution center were performed with different settings of the reorder policies. More precisely, the reorder policies considered are characterized by different operating leverages, as indicated below.

- For the EOQ policy, operating leverages are:
  - EOQ, i.e. the fixed quantity of product ordered to order [units]; and
  - order point (OP), i.e. the level of inventory at which the supply chain player will make an order to suppliers [units];
- For the EOI policy, operating leverages are:
  - EOI, i.e. the time interval between two subsequent orders [days];
  - order up to level (OUTL), i.e. the level of stock to recover when ordering [units].
- For the (S, s) policy, operating leverages are:
  - s, corresponding to the order point of the policy [units];
  - S, corresponding to the order up to level of the policy [units]; and
  - ΔT, i.e. the time interval between two subsequent controls of the stock level [days].

The basic idea of the simulation was to express one of the operating leverages of each policy as a function of the required level of safety stocks (k); in this way, a direct relationship between the total costs and achieved level of customer service was obtained and analyzed. The operating leverages that we chose to express as a function of k are OP for the EOQ policy, OUTL for the EOI policy, and s for the (s, S) policy.

The mathematical relationships which expresses one operating leverage as a function of the customer service *k* are shown below:

$$OP = LT \times \bar{d} + k\sigma\sqrt{LT} \quad (1)$$

$$OUTL = \bar{d} \times (EOI + LT) + k\sigma\sqrt{EOI + LT} \quad (2)$$

where LT is the order lead time [days], $\bar{d}$ is the average demand of the product [units/day] and σ is the standard deviation of the demand [units/day]. The formula for the OP is used for both the EOQ and (S,s) policies.

Through the simulation, a range of values has been assigned to *k* and to the remaining operating leverage(s) of each policy. Hence, during this preliminary simulations, the model was exploited to examine different settings of the operating leverages of each policy, with the purpose of identifying the *minimum cost* setting of each policy. Further input data used in the model were the order lead time and the maximum capacity of the store shelf, which is assumed as the S in the (S,s) policy.

3.3. Simulation procedure

A specific procedure was followed during the simulations. In particular, for a given product, we first simulated the order flow of the retail store, according to the different reorder policies investigated. As mentioned, different settings were used for the parameters of the reorder polies, with the purpose of identifying the *minimum cost* setting. We then collected the main performance parameters of each policy, in terms of the total cost (TC, expressed in [€/units/day]) of the policy and the throughput time (TT, expressed in [days]) of the product at the retail store. The total cost takes into account the cost of holding stocks, stock-out and order for each product examined. In turn, the cost of holding stocks and the stock-out cost are simply determined starting from the unitary costs shown in Table 2 and on the daily stock level of the retail store, while the order cost is calculated considering the number of orders made. The TT is computed, for each supply chain player, as

$$TT = LT + \frac{\bar{\imath}}{\bar{d}} \quad (3)$$

being $\bar{\imath}$ the average stock level [units] of the product for the player considered. In this study, the TT is a relevant performance indicator, because the products simulated have different shelf life characteristics. Moreover, some of them have very limited shelf life, and thus they should reach the final customer is a short time.

On the basis of the outcomes (i.e., TC, TT and the minimum cost setting of the policy), we assessed the *suitability* of each policy for each product. By suitability, we mean that the resulting optimal setting should comply with the product characteristics (i.e., its shelf life) or to the store constraints (i.e., the amount of shelf space). For instance, the minimum cost setting of the EOQ policy for a give product could generate an EOQ which could not be compatible with the amount of shelf space the particular product is given. Under this circumstance, the optimal setting of the EOQ policy is not suitable to be used for that product. Whenever the minimum cost setting of a policy turns out to be unsuitable for a given product, this policy will not be considered for that product. Among the suitable policies, we finally chose the *optimal* one, on the basis of the TC it generates.

The order flow of the retail store resulting under the optimal policy is used as input (and, in particular, as the demand) for the simulations of the distribution center reorder process. Again, all the reorder policies are simulated for this player, with different settings, and the same procedure described above is repeated for the selection of the optimal policy.

4. RESULTS AND DISCUSSION

Following the procedure described in section 3.3, we obtained, through simulation, a *minimum cost* reorder policy for each actor (i.e., retail store and distribution center) and for each product considered (from product 1 to product 5). However, as remarked, the minimum cost solution could be not suitable for application to a given product because of its characteristics, such as shelf life or store constraints. The *optimal* policy, i.e. the minimum cost policy which is also suitable to be applied to a given product, is highlighted in bold in Tables 3 and 4, respectively for the retail store and the distribution center. For each policy, we provide the resulting total cost [€/day/unit], the cost composition (in percentage) and the operative leverages set to obtain the minimum cost scenario, according to the description in section 3.2. The remaining outcomes (e.g., the TT) are directly derived from the simulation.

As regards the notation used, it should be remarked that, for simplicity, results are referred to "units" for both the distribution center and the retail store. The meaning, however, is different, since, by "units" we mean product "items" in the case of the retail store, and product "cases" in the case of distribution centers.

4.1. Retail store results

As regards the retail store and starting from the product with the shortest shelf life (i.e., product 1), it can be seen from Table 3 that the EOI policy appears as the minimum cost one, but it is not suitable to be adopted for this product, since it does not meet the shelf life and the store constraints (i.e., the store facing): in particular, the store facing for product 1 is 70 items, while, adopting EOI, the resulting OUTL should be significantly higher (232 items). For the same product, the (S,s) policy (which is the *optimal* one, because of the incompatibility between EOI and the product characteristics) generates a significant increase in the total cost, compared to EOI. This is mainly due to high stock-out cost resulting under the (S,s) policy, and indicates that, under that policy, the retail store experiences significant stock-out situations.

Specifically, stock-out cost accounts for 34% of the total cost resulting under (S,s) policy. From the above outcomes, two possible approaches can be suggested for the optimal management of product 1 at the retail store.

- A first approach could be to manage product 1 adopting the EOI policy with k=1.5 and EOI=4 days, which is the *minimum cost* setting. Since such setting does not met the constraint of the store facing, the retail store should consider to increase the amount of shelf space for this product, raising it to approx. 230 items;
- Alternatively, product 1 could be managed according to the (S,s) policy, with k=1 and ΔT=1 day. This would lead to a higher total cost, but all the problem constraints are satisfied. The main issue with the (S,s) policy is that the retail store experiences numerous out of stock situations, meaning that the number of items available for purchase is too low. However, since product 1 has a very limited shelf life, the lack of the product on the shelf could also be acceptable, since it means that there will be no product shrinkage, and thus no cost for product disposal will arise. This could be an interesting business strategy that the retail store could consider. Under this scenario, sale losses should be avoided by offering alternative products to customers once product 1 is out of stock.

Looking at the mozzarella cheese (product 2), we can see that the most severe constraint is given by product shelf life (9 days), which should be lower than the TT. On the basis of this consideration, the EOI policy turns out to be the only suitable policy for that product, since both EOQ and (S,s), with minimum cost setting, generate excessive TT. However, it should be remarked that, no matter the policy, the way product 2 is currently managed by the retail store is probably inefficient: in fact, the store facing (54 items, as indicated in Table 1) is too high, compared to the product demand (approx. 1.60 items/day). With those settings, all the reorder policies simulated generate somehow inconsistent results, and can be hardly adapted to the real scenario.

Because of the same reason, i.e. the inconsistency between the shelf facing and the daily demand, only the (S,s) policy is suitable to be adopted for product 3 (yoghurt), since its minimum cost setting meets all the problem constraints. This policy, although optimal, generates higher total cost, compared to EOI or EOQ.

An opposite situation occurs for product 4 (plum pulp): in this case, the daily demand is high and the shelf facing (12 items) of the product is probably undersized. It is thus likely that this product experiences stock-out situations. For product 4, the minimum cost policy resulting from the simulation is EOQ, which is also compatible with the product characteristics and is thus the optimal one. Similar considerations hold for product 5 (royal jelly): EOQ turns out to be the minimum cost policy, and is also compatible with the product characteristics. It should be noted, in this regard, that both products 4 and 5 have less severe constraints in terms of shelf life (163 days and 322 days, respectively), thus all inventory management policies can be easily adapted to those products, via appropriate settings. As a result, no relevant incompatibilities between these products and the reorder policies emerge from Table 4.

Table 3: retail store results

	Product 1-MILK (Shelf life 3 days) (đ = 44.52 units/day)		
	EOQ	EOI	**(S,s)**
Average total cost [€/day/unit]		2.675 (stock out cost 0.60%) (cost of holding stock 36%) (order cost 63.40%)	**10.241 (stock out cost 34%) (cost of holding stock 0%) (order cost 66%)**
k		1.5	**1**
TT [day]		2.81	**1.749**
Operative leverage		EOI = 4 OUTL = 232	**ΔT = 1 S = 70 s = 48**
	Product 2-MOZZARELLA CHEESE (Shelf life 9 days) (đ = 1.60 units/day)		
	EOQ	**EOI**	(S,s)
Average total cost [€/day/unit]	0.589 (stock out cost 0%) (cost of holding stock 55%) (order cost 45%)	**1.145 (stock out cost 0.30%) (cost of holding stock 15%) (order cost 84.70%)**	0.635 (stock out cost 0%) (cost of holding stock 70%) (order cost 30%)
k	1.7	**1**	1
TT [day]	15.3	**9**	20.5
Operative leverage	EOQ = 40 OP = 6	**EOI = 7 OUTL = 17**	ΔT = 1 S = 54 s = 5
	Product 3-YOGHURT (Shelf life 15 days) (đ = 1.83 units/day)		
	EOQ	EOI	**(S,s)**
Average total cost [€/day/unit]	0.617 (stock out cost 0%) (cost of holding stock 32%) (order cost 68%)	0.736 (stock out cost 0.15%) (cost of holding stock 28%) (order cost 71.85%)	**0.815 (stock out cost 0%) (cost of holding stock 15%) (order cost 85%)**
k	1.9	1	**1**
TT [day]	10.34	10.68	**7.1**
Operative leverage	EOQ = 30 OP = 6	EOI = 13 OUTL = 30	**ΔT = 1 S = 21 s = 5**
	Product 4-PLUM PULP (Shelf life 163 days) (đ = 0.30 units/day)		
	EOQ	EOI	(S,s)
Average total cost [€/day/unit]	**0.274 (stock out cost 1%) (cost of holding stock 43%) (order cost 56%)**	0.382 (stock out cost 3%) (cost of holding stock 17%) (order cost 80%)	0.282 (stock out cost 0.25%) (cost of holding stock 46%) (order cost 53.75%)
k	**1**	1	1
TT [day]	**27.26**	16.08	29.79
Operative leverage	**EOQ = 6 OP = 2**	EOI = 1 OUTL = 2	ΔT = 3 S = 12 s = 2
	Product 5-ROYAL JELLY (Shelf life 322 days) (đ = 0.70 units/day)		
	EOQ	EOI	(S,s)
Average total cost [€/day/unit]	**0.339 (stock out cost 0%) (cost of holding stock 43%) (order cost 57%)**	0.465 (stock out cost 0%) (cost of holding stock 16%) (order cost 84%)	0.346 (stock out cost 1%) (cost of holding stock 43%) (order cost 56%)
k	**1**	1	1
TT [day]	**20.36**	11.59	20.86
Operative leverage	**EOQ = 24 OP = 3**	EOI = 1 OUTL = 3	ΔT = 3 S = 24 s = 3

4.2. Distribution center results

Results related to the distribution center are proposed in Table 5.

The main outcome from Table 5 is that, for 3 out of 5 products simulated, none of the reorder policies considered turns out to be suitable for implementation. The products for which we were unable to find a suitable reorder policy are milk (product 1), mozzarella cheese (product 2) and yoghurt (product 3). The main reason for unsuitability is that the distribution center TT is always higher than the product shelf life, which, as already observed, is particularly short for those products. To reduce the TT, at the same time avoiding

product expiry, from a practical perspective it can be suggested that products 1, 2 and 3 should be supplied to the retail store through direct deliveries.

The remaining products, as already remarked, are less problematic in terms of their shelf life; thus, all the reorder policies simulated generate an acceptable scenario and the optimal policy is simply the minimum cost one. In particular, for both products 4 and 5, the optimal solution is given by the (S,s) policy.

Table 5: distribution center results

	Product 1-MILK (Shelf life 2 days) (d = 4.45 units/day)		
	EOQ	EOI	(S,s)
Average total cost [€/day/unit]		0.398 (stock out cost 0%) (cost of holding stock 40%) (order cost 60%)	2.13 (stock out cost 58%) (cost of holding stock 2%) (order cost 40%)
k		1	2.8
TT [day]		2.94	2.32
Operative leverage		EOI = 4 OUTL = 22	ΔT = 4 S = 23 s = 3
	Product 2-MOZZARELLA CHEESE (Shelf life 5 days) (d = 0.20 units/day)		
	EOQ	EOI	(S,s)
Average total cost [€/day/unit]	0.105 (stock out cost 10%) (cost of holding stock 55%) (order cost 35%)	0.167 (stock out cost 0.10%) (cost of holding stock 18%) (order cost 81.90%)	0.48 (stock out cost 0.6%) (cost of holding stock 98%) (order cost 1.4%)
k	1	1	1
TT [day]	15.27	9.54	112.42
Operative leverage	EOQ = 5 OP = 1	EOI = 3 OUTL = 2	ΔT = 5 S = 43 s = 0
	Product 3-YOGHURT (Shelf life 8 days) (d = 0.30 units/day)		
	EOQ	EOI	(S,s)
Average total cost [€/day/unit]	0.151 (stock out cost 8%) (cost of holding stock 64%) (order cost 28%)	0.147 (stock out cost 0%) (cost of holding stock 34%) (order cost 66%)	0.11 (stock out cost 0%) (cost of holding stock 56%) (order cost 44%)
k	1	1	1.2
TT [day]	16.27	10.78	13.77
Operative leverage	EOQ = 7 OP = 2	EOI = 2 OUTL = 3	ΔT = 8 S = 6 s = 1
	Product 4-PLUM PULP (Shelf life 80 days) (d = 0.045 units/day)		
	EOQ	EOI	**(S,s)**
Average total cost [€/day/unit]	0.073 (stock out cost 27%) (cost of holding stock 54%) (order cost 19%)	0.054 (stock out cost 0%) (cost of holding stock 60%) (order cost 40%)	**0.049** **(stock out cost 0%)** **(cost of holding stock 56%)** **(order cost 44%)**
k	1	1	**1**
TT [day]	48.76	49.28	**49.44**
Operative leverage	EOQ = 3 OP = 1	EOI = 11 OUTL = 2	**ΔT = 21** **S = 2 s = 0**
	Product 5-ROYAL JELLY (Shelf life 158 days) (d = 0.06 units/day)		
	EOQ	EOI	**(S,s)**
Average total cost [€/day/unit]	0.072 (stock out cost 0%) (cost of holding stock 75%) (order cost 25%)	0.06 (stock out cost 0%) (cost of holding stock 54%) (order cost 46%)	**0.053** **(stock out cost 0%)** **(cost of holding stock 75%)** **(order cost 25%)**
k	1	1	**1.1**
TT [day]	55.33	42.84	**60.76**
Operative leverage	EOQ = 3 OP = 2	EOI = 1 OUTL = 2	**ΔT = 25** **S = 4 s = 1**

5. CONCLUSIONS

Inventory management is a basic element of competition in order to increase company's efficiency and profitability.

In this paper, we analyzed, through simulation, 3 traditional reorder policies (i.e., EOI, EOQ and S,s) applied to a 2-echelon supply chain, with the purpose of identifying the optimal one with respect to the characteristics of a given product. A particular attention has been paid to fresh products with limited shelf life. As a result, we provided the optimal inventory management policy, its optimal setting, and the resulting total cost, as a function of the product considered and of the supply chain player examined.

The outcomes obtained lead to the following major conclusions. A first consideration is that, for products 4 and 5, which do not fall into the category of fresh products, no significant problems emerge as regards the compatibility of the reorder policies with the product characteristics. Conversely, the correct management of products 1, 2 and 3, which have limited shelf life, is more problematic, and it emerged from our analysis that reorder policies cannot always been adapted to those products. For those products, some practical guidelines can nonetheless be suggested on the basis of the outcomes obtained. As a matter of fact, it seems that the current way such products are managed is inefficient and could be improved. For product 1, the current store facing of 70 items allows satisfying the product demand of less than 2 days (being the daily demand approx. 44 units). No matter the reorder policy applied, this setting always leads to numerous out-of-stock situations. Although stock-out situations could be accepted, because no cost of product disposal arises, they should be properly managed by the retail store. For products 2 and 3, on the contrary, the current store facing is oversized compared to the daily demand of the product at the retail store. For instance, the store facing accounts for 54 items for product 2, while the daily demand is approx. 1.60 items/day. No matter the reorder policy applied, those settings generate a very high TT, which leads to incompatibility with the short product shelf life. Such scenario could be improved by reducing the store facing of those products.

The main contributions of this study can be summarized as follows. First, from the methodological point of view, we compare the different inventory policies to manage the stock level of products, instead of focusing on the optimization of only a specific policy (which is common among the studies available in literature). Second, inventory management policies are examined with respect to a whole supply chain, even if the optimal solution is given separately for each supply chain player. Future research could consider be oriented toward the optimization of the whole supply chain cost. Third, a real supply chain has been chosen for the analysis, so that the outcomes obtained describe are useful to derive practical guidelines for supply chain managers in real scenarios.

Some limitations of the work should be mentioned. One is related to the simulation model developed, which is susceptible to be improved. In fact, under some scenarios (and, in particular, when the real scenario shows some inconsistencies), the simulations performed were ineffective in finding suitable reorder policies. In those cases, the user should change the operating leverages of the reorder policies manually to derive useful results. Moreover, the model omits some specific cost components, such as the disposal cost of expired goods or the cost for checking the stock level, which should be considered for some reorder policies. The cost of disposal was not introduced in the model

because of lack of information in this regard. Specifically, it is not know which supply chain player has to pay such cost. Conversely, the cost for checking the stock level has been voluntary omitted, because typically its amount is very limited compared to the remaining cost components. Nonetheless, on the basis of the considerations above, some improvements will be introduced in the model in future studies.

REFERENCES

Abdel-Malek, L., Montanari, R., 2005a. On the multi–product newsvendor problem with two constraints. *Computer & Operations Research*, 32, 2095-2116.

Abdel-Malek, L., Montanari, R., 2005b. An analysis of the multi-product newsboy problem with a budget constraint. *International Journal of Production Economics*, 97(3), 296-307

Blackburn, J., Scudder, G., 2009. Supply Chain Strategies for Perishable Products: The Case of Fresh Produce. *Production and Operations Management*, 18(2), 129–137

Bottani, E., Montanari, R., 2011. Design and performance evaluation of supply networks: a simulation study. *International Journal Business Performance and Supply Chain Modelling*, 3(3), 226-269

Deniz, B., Scheller-Wolf, A., Karaesman, I., 2009. Managing inventories of perishable goods: the effect of substitution. Working paper of the Carnegie-Mellon University (Pittsburgh). Retrieved June 2012 from http://faculty.fuqua.duke.edu/seminarscalendar/Scheller-WolfSeminar.pdf

Ferguson, M., Jayaraman, V., Souza, G.C., 2007. Note: An Application of the EOQ Model with Nonlinear Holding Cost to Inventory Management of Perishables. *European Journal of Operational Research*, 180(1), 485-490

Goyal, S.K., 1985. Economic Order Quantity under Conditions of Permissible Delay in Payments. *The Journal of the Operational Research Society*, 36(4), 335-338

Liu, L., Lian, Z., 1999. (s, S) continuous review models for products with fixed lifetimes. Operations Research, 47(1), 150-158

Nahmias, S., 1982. Perishable inventory theory: a review. *Operation Research*, 30, 680–708

Raafat, F., 1991. Survey of literature on continuously deteriorating inventory models. *The Journal of the Operational Research Society*, 42(1), 27-37

Ravichandran, N., 1995. Stochastic Analysis of a continuous review perishable inventory system with positive lead time and poison demand, *European Journal of Operations Research*, 84, 444-457

Schmidt, C.P., Nahmias, S., 1985. (S-1,S) policies for perishable inventory. *Management Science*, 31, 719-728

Schneider, H., Ringuest, J.L., (1990). Power approximation for computing (s, S) policies using service level. *Management Science*, 36, 822-834

Silver, E.A., Naseraldin, H., Bischak, D.P., 2009. Determining the reorder point and order-up-to-level in a periodic review system so as to achieve a desired fill rate and a desired average time between replenishments. *Journal of the Operational Research Society*, 60(9), 1244-1253

Waters, D.,2003. *Logistics: an introduction to supply chain management*. New York: Palgrave MacMillan.

Weiss, H.J., 1980. Optimal ordering policies for continuous review perishable inventory models. *Operations Research*, 28, 365-374

AUTHORS BIOGRAPHY

Marta RINALDI is a PhD Student at the University of Parma, and Scholarship Holder in Industrial Engineering at the Interdepartmental Center CIPACK. She got a master degree in Management Engineering with a thesis titled "Analysis and evaluation of energy efficiency of a shrinkwrap-packer". Her main fields of research are discrete event simulation and simulation of industrial plants.

Eleonora BOTTANI is Lecturer (with tenure) in Mechanical industrial plants at the Department of Industrial Engineering of the University of Parma (Italy). She graduated in 2002 in Industrial Engineering and Management at the University of Parma, and got her Ph.D. in Industrial Engineering in 2006. Her research activities concern logistics and supply chain management issues, encompassing intermodal transportation, development of methodologies for supplier selection, analysis and optimization of supply chains, supply chain agility, supply chain modelling and performance analysis, and, recently, the impact of RFID technology on the optimization of logistics processes and supply chain dynamics. She is the author or coauthor of more than 80 scientific papers, referee for more than 40 international scientific journals, editorial board member of 2 scientific journals and Associate Editor for one of those journals.

Gino FERRETTI is full professor of Mechanical Plants at the University of Parma (Italy). He graduated in Mechanical Engineering in 1974 at the University of Bologna, where he served as assistant professor for the courses of “Machines” and “Mechanical Plants”. He worked as associate professor at the University of Padua and as full professor at the University of Trento. In 1988, he moved to the Faculty of Engineering, University of Parma, where at present he is full professor of Mechanical Plants. His research activities focuses on industrial plants, material handling systems, and food processing plants, and have been published in numerous journal and conference papers.

Mattia ARMENZONI is a scholarship holder in Industrial Engineering at the University of Parma (Interdepartmental Center Siteia.Parma). He got a master degree in Mechanical Engineering for the Food Industry, discussing a thesis titled: "Advanced design of a static dryer for pasta with simulation tools". He attended the 10th International Conference on Modeling and Applied Simulation (Rome, 12-14 September 2011) presenting the paper titled "Advanced design of a static dryer for pasta with simulation tools", and the 2011 EFFoST Annual Meeting (Berlin 9-11 November 2011) presenting the paper titled "Advanced design of a UV reactor for water treatment using computational fluid dynamics".

Davide MARCHINI is scholarship holder in Industrial Engineering at the University of Parma (Interdepartmental Center Siteia.Parma). He got a master degree in Mechanical Engineering for the Food Industry, discussing a thesis titled: "Advanced design of a UV reactor for water treatment using computational fluid dynamics". He attended the 10th International Conference on Modeling and Applied Simulation (Rome, 12-14 September 2011), presenting the paper titled "Advanced design of a static dryer for pasta with simulation tools", and the 2011 EFFoST Annual Meeting (Berlin 9-11 November 2011), presenting the paper titled “Advanced design of a UV reactor for water treatment using computational fluid dynamics”.

Federico SOLARI is a PhD Student in Industrial Engineering at the University of Parma; master degree in Mechanical Engineering of the Food Industry, dissertation of the thesis: "Analysis and design of a plant for the extraction of volatile compounds from aqueous matrix"; Attending many international conferences, he’s author or coauthor of 7 international papers. His research activities mainly concern industrial plants and food process modeling and simulation, with a particular focus on the CFD simulation for the advanced design of food plants.

Giuseppe VIGNALI: graduated in 2004 in Mechanical Engineering at the University of Parma. In 2009, he received his PhD in Industrial Engineering at the same university, related to the analysis and optimization of food processes. Since August 2007, he works as a Lecturer at the Department of Industrial Engineering of the University of Parma, and, since the employment at the university, he has been teaching materials, technologies and equipment for food packaging to the food industry engineering class. His research activities concern food processing and packaging issues and safety/security of industrial plant. Results of his studies related to the above topics have been published in more than 20 scientific papers, some of which appear both in national and international journals, as well in national and international conferences. He acts also as a referee for some international journals, such as Prevention Today, Facilities.

Roberto MONTANARI is full professor of Mechanical Plants at the University of Parma (Italy). He graduated (with distinction) in 1999 in Mechanical Engineering at the University of Parma. His research activities mainly concern equipment maintenance, power plants, food plants, logistics, supply chain management, supply chain modelling and simulation, inventory management. He has published his research in approx. 60 papers, which appear in qualified international journals and conferences. He acts as a referee for several scientific journals and is editorial board member of 2 international scientific journals.

MODELLING AND SIMULATION OF A FISH PROCESSING FACTORY SHIP

Nadia Rego Monteil[(a)], Raquel Botana Lodeiros[(b)], Diego Crespo Pereira[(c)], David del Rio Vilas[(d)], Rosa Rios Prado[(e)]

[(a) (b) (c) (d) (e)] Integrated Group for Engineering Research, University of A Coruña (Spain)

[(a)]nadia.rego@udc.es, [(b)] raquel.botana.lodeiros@udc.es, [(c)]dcrespo@udc.es, [(d)]daviddelrio@udc.es, [(e)] rrios@udc.es

ABSTRACT

A fish processing factory ship is a large vessel with on-board facilities for the immediately processing and freezing of caught fish. This is a paradigmatic case of changing environment due to the uncertainty in the quantity and quality of the catch, the importance of the human factor and the frequent adverse operation conditions. Decision making processes are especially difficult to set regarding either resource allocation or the mix of products. This paper presents a case study in a Spanish company with special attention to the sources of variability. An overall exploratory DES model of the plant together with a specific DHM simulation of the workers tasks are developed as a means to gain insight into the process. The process efficiency in different production scenarios, the organizational effects in the packing workstation and the ergonomic and operational assessment of the wrapping operations are studied.

Keywords: Modelling and Simulation, Discrete Event Simulation, Digital Human Modelling, Fish processing, Factory ship

1. INTRODUCTION

A factory ship is a large vessel with on-board facilities for processing and freezing caught fish. There are about 24,000 vessels of more than 100 tons in the world's factory fishing fleet (FAO Archives 2004). According to Eurostat, in terms of tonnage the Spanish fishing fleet is by far the largest (415,000 gross tonnes) of Europe. This fleet produces around 1,000,000t of fish per year (FAO Archives 2007). The Spanish fleet is composed of 11,000 ships, but only 400 trawlers, seiners and liners account for 50% of the global tonnage. These are the ships that fish out of the EU territorial waters (Ministerio de Medio Ambiente, Medio Rural y Marino 2011).

The Spanish company involved in this study has fifteen trawlers that operate in the Southwest and Southeast Atlantic Ocean fishing grounds. Some of these trawlers are multispecies and other are rather specialized in one species. The most complex type of trawler –the multispecies- has two or three parallel lines capable of producing between 25t and 40t of frozen fish per day.

The on-board fish processing lines involve several production workstations with both manual and automatic operations. The analysis of such a system implies to consider the process flow through several parallel lines with different inputs and outputs. In fact, there is a great variety of final products. The uncertainty due to the quantity and quality of the catch makes it difficult to take the optimal decision regarding either resource allocation or the mix of products. The mathematical analysis is then quite complex to conduct. Even more, unless a dynamic approach is adopted, some key factors are very difficult to estimate, such as the time-in-process of the final product. There is a direct relationship between this parameter and the quality of the frozen fish (Trucco et al.1982). As a consequence, a simulation based analysis in order to assess different production alternatives has been considered.

To our knowledge, the fish processing has been seldom analysed under an engineering production perspective. It is remarkable the network-based simulation of a processing facility in land made by Jonatansson and Randhawa (1986). Among the results from their model are statistics on utilization of machines and workers in the process, size of in-process inventory at different locations in the process, and throughput times.

On the other hand, the working conditions of the operator have been widely discussed. Several ergonomic and clinical studies have been carried out. A clear prevalence of shoulder and upper-limb disorders among the workers in eight different factories in the Kaohsiung port (Taiwan) is reported in Chiang et al. (Chiang et al. 1993). A L.E.S.T. analysis (Ergonomic Evaluation Method developed by the Institute of Labour Economics and Industrial Sociology of France) was conducted to characterize the risk in a fish processing plant in Ecuador (Torres and Rodríguez 2007). Regarding the assessment of on board workers, a study of Swedish fishers showed that they experience frequent musculoskeletal disorders (MSDs), according to the type of task, but also a special type of stress due to the natural ship instability (Törner et al. 1988). Another critical factor of these operators is the level of noise, which has been also studied for workers of a fishing trawler (Szcepański and Weclawic 1991).

The process on board is highly dependent on human operators (between 45 and 65 people spend several weeks working on board). Due to hard working conditions, another aim of the study was to characterise

the ergonomic impact of some tasks by means of digital human modelling (DHM). Some studies have already applied this tool to the fishing sector (Zhang 2010, Álvarez-Casado et al. 2011). DHM has been used to jointly consider productivity and ergonomic measures for the workstation design in a very wide range of sectors, like in the food industry (Ben-Gal I and Bukchin 2002), in mining (Rego et al. 2010) and in the automobile industry (Fritzsche 2010).

A combined simulation approach has been adopted for the characterization and improvement of the process: (i) a global analysis of the production system by means of discrete event simulation and (ii) an ergonomic study of the individual tasks. The aim of this paper is to describe the case study and the proposed methodology for its analysis. Although this is an on-going project, some relevant preliminary results are also described.

2. PROCESS DESCRIPTION

In spite the processing process starts and depends on the previous fishing process, only the indoor activities will be explained for the sake of simplicity. The flow diagram of the fish processing is depicted in Figure 1. Initially, the captured fish on deck is introduced into the processing plant by means of a ramp connected to a hopper which feeds a distribution conveyor belt. A manual classification (Figure 2) conveys the fish to the filleting line –Product "Fillet"-. If it is too big or too small to go to this filleting line, it goes to the whole fish line –Product "HG product"-. If it is not able to be processed in time or it does not fit the requirements, it is returned to the sea -Discards-.

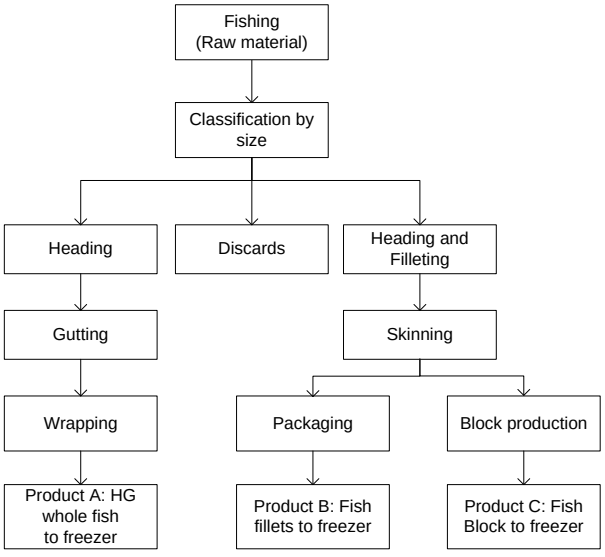


Figure 1. Flow diagram of the factory ship process

On the whole fish line, the fish undergo three sequential operations, named heading, gutting and wrapping. Heading and gutting operations are performed by two machines assisted each one by an operator. The fish without the head and the guts -around a 70 % of the weight of the fish- is washed in the second machine. It waits into an intermediate buffer to be wrapped by an operator. Finally, as showed in Figure 3, it is placed into a box with similar sized HG products (headed and gutted). The full box is ready to go to the freezing stage.

Figure 2. Manual Classification of the Fish.

Figure 3. Wrapped HG fish in a box.

If the fish has an adequate size, it is sent to the fillet lines. Three parallel workstations accomplish the heading, gutting, filleting and skinning of the fish. Each fish yields two fillets that directly go to the skinning, an operation that removes the skin from each fillet. The overall yield is estimated in 40% of the initial weight of the fish. The fillets are then conveyed to a common belt to be manually put into trays. This operation, the packing, consists on selecting similar sized fillets, trimming them if they still have rests on skin or bones, and place them forming several layers into a tray. A plastic sheet is placed between layers to avoid the adherence of fillets. The full box of fillets is then sent to the freezing stage.

Figure 4. Packed Fillets in a Tray

The fillets may be rejected from the packing operations. This eventually happens because they fail reaching the quality requirements or because the packing operators have not enough capacity to process them. In that case, they go to the so-called "fish-block" conformation. This third product is a block of fillets weighing 7.5kg and measuring 485mm x 255mm x 63mm, intended for further processing (breaded sticks, skewered fish, cooked dishes, etc.)

The overall variability exhibited in the process performance can be explained by several causes:

1. Product variability. An important difference between any processing plant and a factory ship is the greater uncertainty about what and when the raw material enters the process. This is due to the heterogeneous distribution of the fish along the sea and the irregular distribution of species and sizes. The different species morphology and size influence the availability to automatic filleting. This obliges the process to be flexible.
2. Process variability. Apart from the above mentioned influence of the species and sizes in the mix HG/fillets, there are other factors that link the product characteristics and the process parameters. First, when the fish waits too long before it is processed the Rigor Mortis makes too rigid to go through certain operations. In that case, a break down in filleting machines may occur. The second factor is the packaging capacity in the filleting line. When the volume of fillets coming from the skinning operation exceeds the manual packaging capacity, the fillets enter to the block production, a less-valued product.
3. Variability due to the resources. The human operations have a natural variability even if they are repetitive, due to factors like skill, mood, tiredness, hour of the day and experience.
4. Variability due to the environment. The ship rocking has a double consequence on the work development. On the one hand, the scales that can be used on board (able to compensate the movement of the ship) are unaffordable. This makes that they cannot be a part in the process as it is in land. As a consequence, weights on trays are estimated by operators and errors are introduced. On the other hand, it is a recognized stressing effect in the operators. Space restrictions often lead the operator to adapt to suboptimal workstation design. Besides, noise, vibration and humidity are factors that increase the risk of accidents.

3. SYSTEM MODELLING AND SIMULATION

Due to important logistic and economic constraints, visiting the actual plant while operating has not been possible. To overcome this disadvantage, process videos and production reports have been extensively analysed. Probability distributions for the sizes on the caught fish (table 1) and the cycle time for activities (operators and machines cycles) have been modelled. The parameters of the size of fish were obtained from the analysis of the production reports that provide data of final product categories packed and frozen during a set of working periods (usually three days). The general operation system was defined from the videos, layout information and interviews. Operators times were obtained from videos observation and machine cycle times were determined with the engineering department help.

Table 1. Fish Weight Categories (in grams)

Number of group	Range of weigh	Description
1	(0, 200]	Fit for whole line
2	(200, 500]	Fit for whole line or fillet line (second category)
3	(500, 800]	Fit for whole line or fillet line (second category)
4	(800,1500]	Fit for whole line
5	(1500, 4500]	Fit for whole line

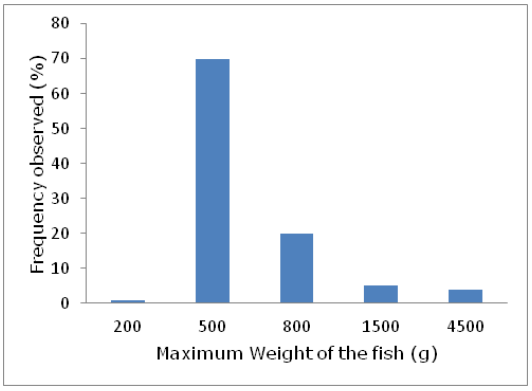


Figure 5. Expected Frequency of Each Fish Weight Category

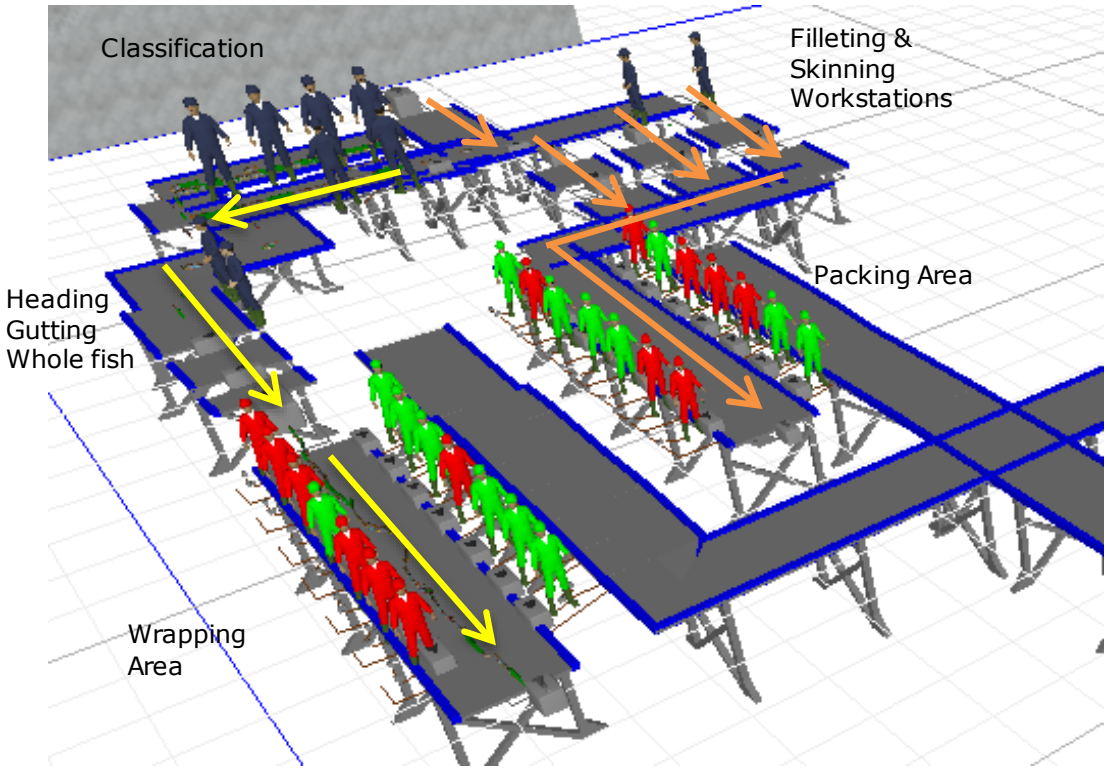


Figure 6. Simulation Model of the Processing Plant in the Factory Ship

The model was developed using SIMIO, an object-based 3D modelling environment. In the simulation model the fish are modelled as entities. The machines and operators are modelled as resources. However, the parameter unit varies depending on the specific operation (fish, lot or fillet).

- Fish. Standard unit from the initial source.
- Lot. During the classification the operator picks up several elements of fish at once –a lot-. The number of fish that compose a lot varies from one to four, according to two empirical distribution functions (to HG process, to fillet process). After the classification, the units are considered individually again.
- Fillet. In the filleting machines, entities modelled as fishes are destroyed after the process time, and two new entities are created as fish fillets.

Wrapping and packing operations imply an individual processing of the product (select, wrap or place) and a common processing as a box or tray (transport). The number of elements that form a box (for HG product) or a tray (for fillets) depends on their size. This has been considered an important factor, because the number of units per container influences the time-in-process of the products, the utilization of the operators and the global time spent in transport to the freezer.

Although a fish size distribution has been defined (Figure 5) this does not mean that all the group 2 and 3 sized fish are sent to the fillets line and the rest of the fish are sent to the whole fish line. A preliminary study of the maximum capacity of the lines regarding the fish supply and the mix of products has been done in order to evaluate this key parameter for the global efficiency of the plant (Figure 7).

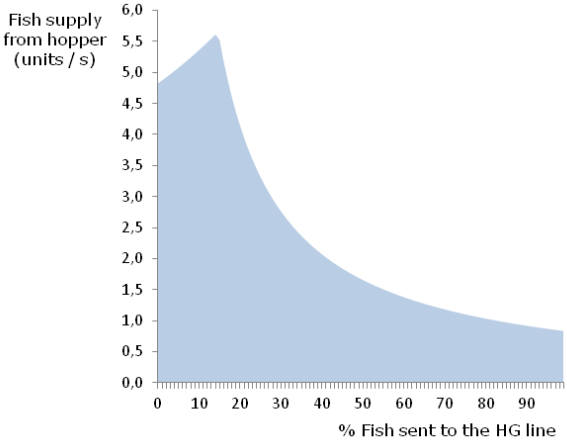


Figure 7. Maximum Supply Rate that can be Processed by the Filleting Machines depending on the Percentage of Product to Whole Fish (HG) Line.

As a result, it is clear that the maximum capacity of the plant occurs when the fish supply is 5.5 units per second and 15% of the supply is sent to the HG line. The normal operation of the factory ship should be close to this optimum working point. Three scenarios will be tested for three different mix of products (10%, 15% and 25% of fish supply sent to the HG line).

In order to define the state of the plant in each operation scenario, a set of performance indicators has to be defined. In this case, we will be accounting for variations in:

- Resource Utilization. Occupation of Operators and Machine compared with the total working time.
- Product Yield. All the fish supply that undergoes the process has four possible outputs: discards, frozen whole fish, frozen fillets and fish block. The production rates of discards and fish block are the variables related to the inefficiency of the system. As a consequence, a better product yield implies reducing them to the minimum.
- Time in process. The time in hours between its exit from the fish hopper and its freezing.
- Production rate. The rates in units per second of the main products of the plant.

Table 2. Summary Statistics of the Three Scenarios

State var.	Element	Scen. 1 (10% to HG)	Scen. 2 (15 % to HG)	Scen. 3 (25% to HG)
Resource utilization	Wrapping operators	48.1 %	58.2 %	78.6 %
	Packing operators	96.3 %	95.3 %	94.3 %
	Average HG machines	39.5 %	52.5 %	90.6 %
	Average fillet machines	48.9 %	47.9 %	44.2 %
Product yield	Fish supply to block	54.3 %	54.2 %	49.9 %
	Fish supply discarded	11.0 %	9.0 %	5.2 %
Time-in-process (h)	Average Time HG	0.12	0.11	0.11
	Average Time of fillets	0.13	0.13	0.13
Produc-tion rate	HG product (units/s)	0.22	0.27	0.37
	Fillets (units/s)	2.12	2.09	2.09

For a fish supply of 3 units/sec, Scenario 3 exhibits better resource utilization for the HG line, a better yield of the products (less discards and less fish supply sent to the lowest valued product) and a similar time in process. With a mix of 25% of fish supply sent to the HG lines and 75% to the filleting lines, the HG machines are close to saturation.

There is not a great difference between the resource utilization in the filleting lines. This can be explained because of the packing bottleneck that reduces the potential capacity of the line. This has been one of the reasons that suggest improving the ergonomics and productivity of the workstation.

4. WORKSTATION MODELLING AND SIMULATION

A supplementary analysis of the wrapping/packing organization task has been done. At present all the operators are placed around a common linear conveyor belt where the products coming from the automatic machines are processed (Figure 8). It has to be remembered that every box/tray has to be filled with similar sized HG fish/fillets (there are up to five categories). We have considered that the first product taken by the operator determines the size of the rest. As a result, in all the scenarios an effect of decrease on the utilization rate of the operators is produced as their distance from the source increases (Figure 9). This can be explained because the amount of products at the end of the line (when all the previous workers have chosen their products) is lower and the operator may eventually be blocked, waiting for a specific size to end a cycle.

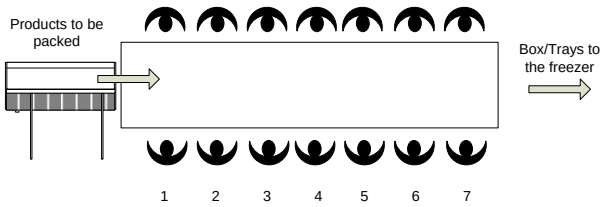


Figure 8. Present Organization of the Wrapping Operators (M1)

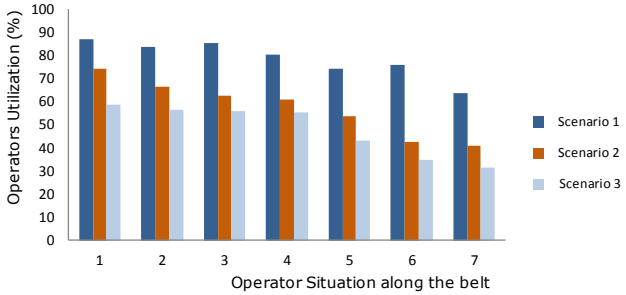


Figure 9. Utilization of the Wrapping Operators related to their Position along the Conveyor Belt

Accordingly, an alternative arrangement of the workers has been modelled. The idea was to divide the

operators in two teams. This idea assumes that neither space restrictions nor technical problems hinder the product flow from being divided.

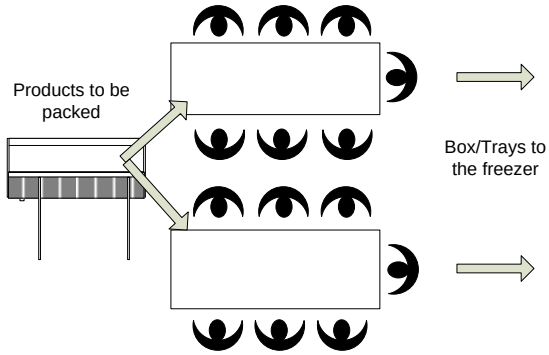


Figure 10. Alternative Distribution of the Wrapping Workstation (M2)

A design of experiments (DOE) approach with two factors –input rate and belt speed- has been considered to compare both models in terms of wrapping performance. The wrapping performance accounts for the proportion of fillets that are effectively processed by the operators and send to the freezing stage.

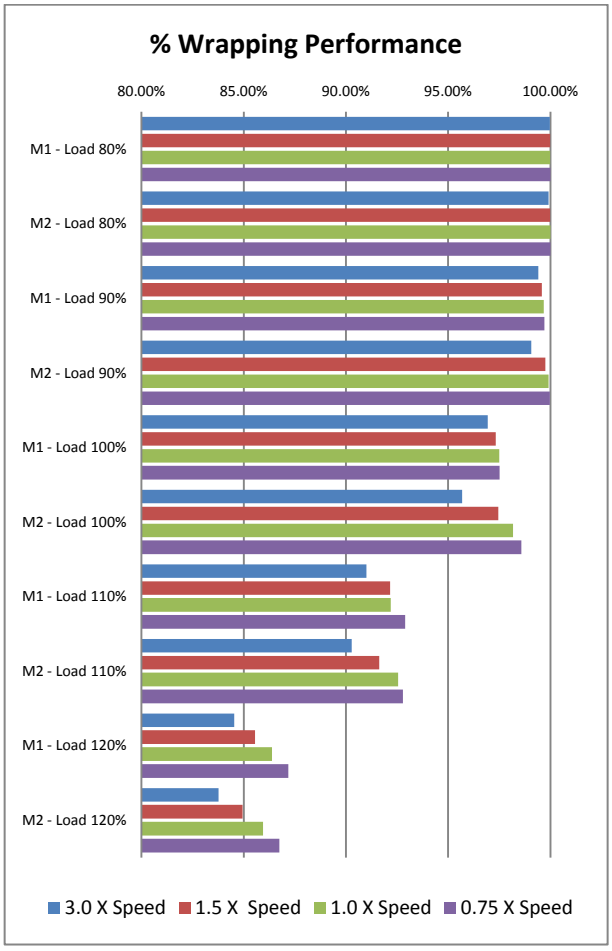


Figure 11. Wrapping Performance depending on the Operators Location along the Belt.

Results are shown in Figure 11. When the input rate is set at 100% and the belt speed is reduced, the M2 distribution achieves better results than the M1 one. For the rest of the cases, M1 behaves better than M2.

For its ergonomic evaluation, the wrapping activity will be decomposed in several components or subtasks. Each subtask is now described as follows:

1. Pick up. The operator bends his back to reach a fillet from the belt (see Figure 12.a.)
   A set of three reach areas have been modelled to cover the entire possible pick up movements.
2. Trimming. If needed, fillets are scissor cut to make them look better and to remove leftover bones (see Figure 12.b).
3. Place on the tray. The fillet is placed on the tray (see Figure 12.c).
4. Plastic sheet between layers. The layer is usually complete after 4, 5 or 6 fillets. A plastic roll is then unwind over the fillets layer. This roll can be seen in Figure 12.c.
5. Tray placed on freezer belt. The tray full of fillets weights around 8 kg. At that point the plastic sheet is cut and the tray is placed over an upper conveyor belt at workers' shoulder level (see Figure 12.d).

Figure 12. Samples of Postures during the Wrapping Operation –(a) Pick up; (b) Trimming; (c) Place Fillet; (d) Place Tray-.

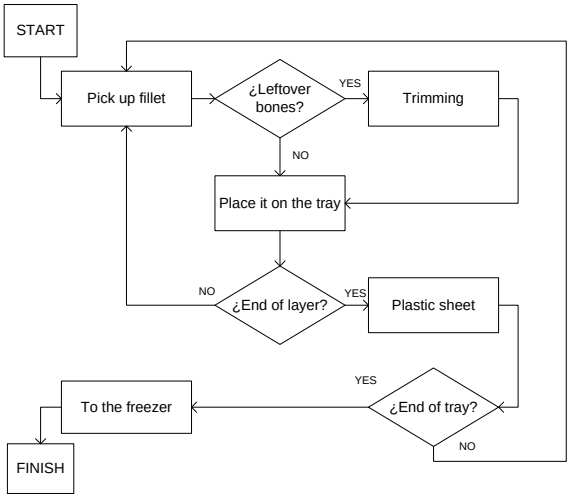


Figure 13. Flow Diagram of the Wrapping Operation

The above mentioned subtasks have been modelled in Delmia V5R20. Their analysis has been done according to the methodology presented in a previous work (Rego et al. 2011). The first stage implies modelling the operators, assuming they fit to the 50th percentile of the French population. The workstation and the tools employed have been modelled by using the geometrical information that the company provided.

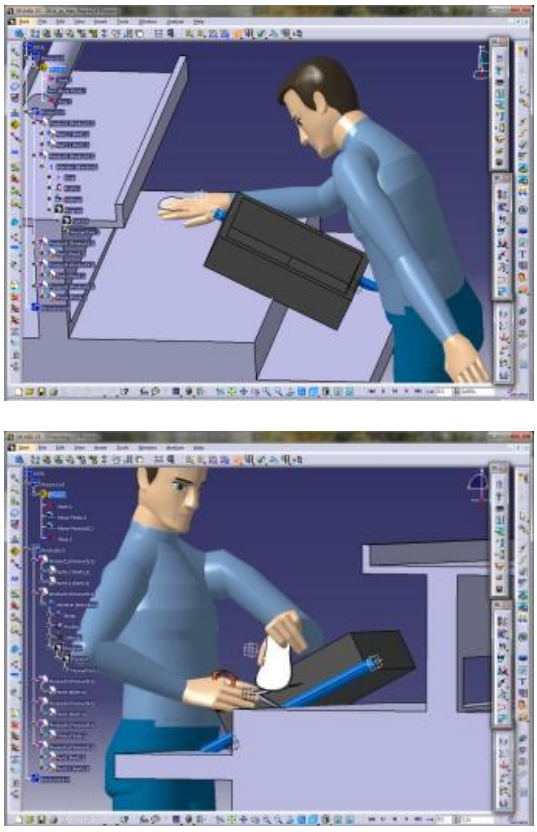


Figure 14. Modelled Postures of the Subtasks -(a) Pick up Max Reach; (b) Trimming; (c) Place fillet; (d) Place Tray.

The RULA score index has been chosen to report the ergonomic evaluation of each subtask. RULA is a well-known and widely used ergonomic assessment method (Cimino et al. 2008) and it is especially thought for the assessment of tasks that mainly imply the upper limbs. The final score is related to the risk of the posture, and goes from 1 (no risk at all) to 7 (urgent need of change to avoid injury). The L4/L5 compression limit has been also considered as an important indicator of the biomechanical risk associated to the adopted postures. The Spine Compression value is a complementary measure of risk of MSDs. According to NIOSH guidelines, compression force on the intervertebral disk above 3.4kN may eventually lead to injuries. Delmia V5R20 provides with both indicators to evaluate each posture of which an activity is made of.

The following charts represent the previous indicators for the different subtasks. As it can be noticed, in Figure 15, the RULA score reaches high levels of risk during the maximum reach pick up task and the place tray task. The L4/L5 compression limits (Figure 16) supports this result with a similar evolution. However, the 3.4 kN limit is never achieved. The rest of the subtasks –place fillet, plastic sheet and trimming operation- remain in relatively “safe” levels.

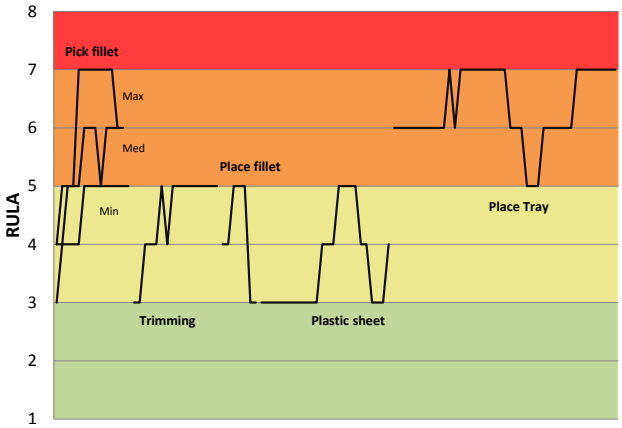


Figure 15. Postural risk of the modelled subtasks

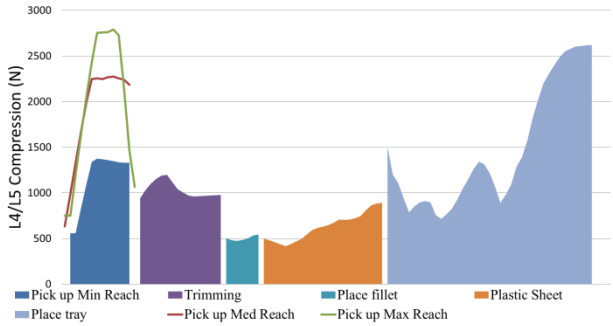


Figure 16. Compression Force on the Intervertebral Disc between L4 and L5 for the modelled Subtasks

Table 3. Summary of results for the wrapping operation

	RULA Score		L4/L5 Comp.	
	Avg	Max	Avg (N)	Max (N)
Pick up Min	4,75	6	1206,18	1375
Pick up Med	5,07	6	1912,71	2275
Pick up Max	6,15	7	2041,38	2789
Trimming	4,50	5	1034,94	1198
Place Fillet	4,14	5	502,43	544
Plastic Sheet	3,58	5	627,65	891
Place Tray	6,37	7	1473,87	2619

A complementary analysis can be done by performing a separately assessment of the different body segments. The RULA method correlates each segment range of movement with the risk of injury. In Figure 17 we present a rate of the average RULA score for each subtask related to the maximum score. In agreement with the literature (Chiang 1993, Törner 1988), there is a clear prevalence of upper limb risks. The forearm and wrists are the most likely parts of developing MSDs. Another remarkable result is that even though the trimming and the place fillet operation were not dangerous in terms of the global analysis, in this analysis they show the highest rates of risk in forearm and wrist.

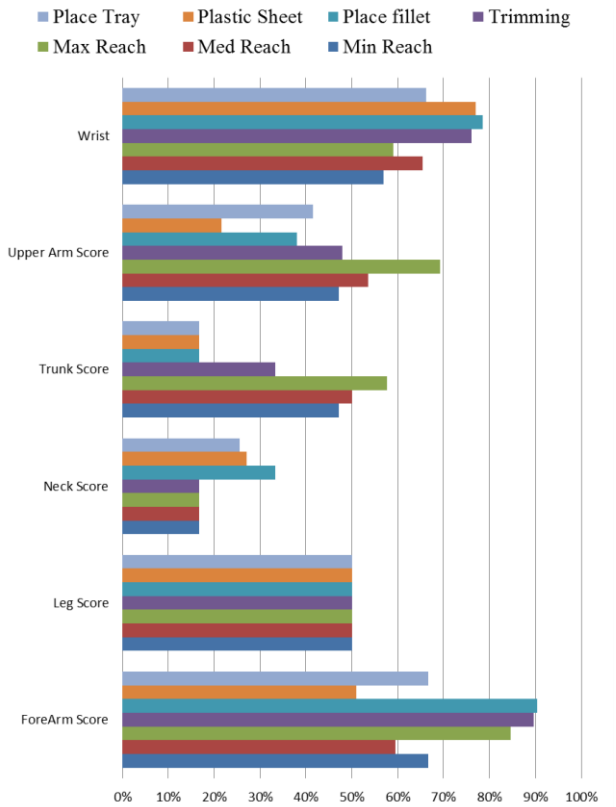


Figure 17. Body Segments Ergonomic Risk for each Subtask

5. VALIDATION AND RESULTS

During the modelling stage a continuous verification effort has been performed by comparing the model with the videos, production reports and the analysis of the real operation times. Validation of preliminary results has been done by a group of experienced workers from the company who found the results to be reasonable according to their experience. For the ergonomic model validation, we also took into account that the literature was in agreement with the main results. Finally, the following findings should be highlighted:

1. The product mix is a key parameter that strongly influences the production rate and the resources utilization. Due to the filleting machines limiting capacity, the point of maximum efficiency of the plant is set to a process input rate of 5.5 units per second and a product mix of 15% to HG lines and 85% fillets.
2. Although the process is oriented to the production of fillets, a more balanced mix between HG product and fillets benefits the product yield and the occupation rates when the input rate is set to 3 units per second.
3. The bottleneck operations are the wrapping operation in the fillet lines and the packing operations in the HG lines.
4. The organization of these workers (wrapping and packing) around a common belt has the effect of decreasing their utilization rate as they are placed

farther from the source. A higher specialisation (two different belts with fewer operators per belt) was tried as an alternative. Their wrapping performance is similar, although in almost all combinations of belt speed and input rate the original organization was slightly better. A 5.02% of fillets could not be processed in the first design, whereas a 5.14% in the second design went to the less valued sub-product (the fish-block).

5. A redesign of the wrapping operation and workplace seems convenient due to their central role in the whole process.
6. The ergonomic analysis of this task showed that placing the full tray of fillets on the upper belt and reaching the fillet from the maximum distance are the hardest tasks in terms of RULA score and L4/L5 compression. The use of smaller trays and an alternative location of the to-freezer-belt should be studied in order to reduce the impact of the "placing full tray task". A redesign of the workplace to reducing the reach distance would indeed decrease the need of back bending.
7. The analysis of the different body segments showed prevalence on wrist and forearm risk. The trimming operation is one of the most demanding in terms of wrist and forearm postures. A better scissor design, amongst other measures, should be proposed in order to reduce the probability of injury.

6. CONCLUSIONS

A simulation analysis of a fish processing plant aboard a common factory ship has been presented. In doing so, a discrete event simulation of the global process and a precise digital human model of the bottleneck operation have been developed. As a result, those parameters affecting the overall process efficiency and the wrapping operation –the actual bottleneck- have been identified. A set of key performance indicators has been defined to evaluate the process efficiency under three different scenarios. Some organizational effects have been found in the last stages of the process, i.e. a decrease in the resources utilization due to the product size variability. An exploratory analysis to assess an alternative organization has been carried out. Finally, an ergonomic and operational analysis of the wrapping operations is presented as a means of improving both the working conditions and productivity.

ACKNOWLEDGMENTS

We wish to express our gratitude to the Xunta de Galicia, which has funded this work through the research project "Ergomar: Proyecto de Investigación y Desarrollo de Métodos de Evaluación Ergonómica en Entornos Dinámicos" (10DPI121E).

REFERENCES

Álvarez-Casado, E., Zhang, B., Sandoval, S. T., Mondelo, P., 2011, Using ergonomic digital human modeling in evaluation of workplace design and prevention of work-related musculoskeletal disorders aboard small fishing vessels, *Human Factors and Ergonomics in Man. & Services Industries*, Vol. 22 (4)

Ben-Gal I, Bukchin J., 2002. The ergonomic design of workstations using virtual manufacturing and response surface methodology. *IEEE Transactions*, 34(4), 375-391.

Chiang H-C, Ko Y-C, Chen S-S, Yu H-S, Wu T-N, Chang P-Y., 1993. Prevalence of shoulder and upper-limb disorders among workers in the fish-processing industry. *Scand J Work Environ Health*, Vol.19(2), pp.126-131.

Cimino, A., Curcio, D., Longo, F., Mirabelli, G., 2008. Workplaces Effective Ergonomic Design: A Literature Review. *Proceedings of the European Modelling and Simulation Symposium*. September 17-19, Campora San Giovani (Amantea, Italy).

FAO Archives, 2004. The State of World Fisheries and Aquaculture (SOFIA). Available from http://www.fao.org/docrep/007/y5600e/y5600e05.htm#P1094_46148 [Accessed April 2012]

FAO Archives, 2007. Food Balance Sheets

Fritzsche L, 2010. Ergonomics Risk Assessment with Digital Human Models in Car Assembly: Simulation versus Real Life. *Human Factors and Ergonomics in Manufacturing*, 20, no.4, pp. 287-299.

Jonatansson E., Randwhawa S.U., 1986. A network simulation model of a fish processing facility, *SIMULATION*, November 1, Vol. 47, pp. 210-211

Ministerio de Medio Ambiente, Medio Rural y Marino, 2011. Estadísticas pesqueras, Spanish Government

Rego N., del Rio D., Crespo D., Rios R., 2010. A Simulaton-based Ergonomic Evaluation for the Operational Improvement of the Slate Splitters Work, *Proceedings of the 22th European Modeling and Simulation Symposium*, pp.191-200

Rego N., del Rio D., Crespo D., Rios R., 2011. An Overall DHM-based Ergonomic And Operational Assessment Of A Manufacturing Task: A Case Study, *Proceedings of the 10th Modeling and Applied Simulation*, pp.375-382.

Szcepański C., Weclawic Z., 1991. Exposure of the crew of a fishing trawler-factory ship to noise, *Bull Inst Marit Trop Med Gdynia*, Vol. 42(1-4), pp. 67-70

Törner M., Blide G., Eriksson H., 1988. Musculo-skeletal symptons as related to working conditions among Swedish professional fishermen. *Applied Ergonomics*; 19(3):191-201.

Torres T., Rodríguez M., 2007. Evaluación Ergonómica de Puestos de Trabajo en la Industria Pesquera del Ecuador, *Revista Tecnológica ESPOL*, 20(1): 139-142.

Trucco, Raúl E., Héctor M. Lupín, Daniel H. Giannini, Marcos Crupkin, Ricarlo L. Boeri, and Carlos A. Barassi, 1982. Study on the Evolution of Rigor Mortis in Batches of Fish. *Lebensm.-Wiss. u.-Technol.*, vol. 15, pp. 77-79.

Zhang B.; Álvarez Casado E., Mondelo P, Tello Sandoval S., 2010. Using digital human modeling in evaluation of working spaces design and prevention of occupational hazards on board fishing vessels. *VIII Congreso Internacional de Prevención de Riesgos Laborales*

AUTHORS BIOGRAPHY

Nadia Rego Monteil obtained her MSc in Industrial Engineering in 2010. She works as a research engineer at the Integrated Group for Engineering Research (GII) of the University of A Coruna (UDC), where she is also studying for a PhD. Her areas of major interest are in the fields of Ergonomics, Process Optimization and Production Planning.

Raquel Botana Lodeiros obtained a MSc in Industrial Engineering in February 2012. She joined the Integrated Group for Engineering Research during her last year degree. Since 2011 she is working in Navantia, the Spanish Public Shipyard devoted to the design and construction of military ships.

Diego Crespo Pereira holds an MSc in Industrial Engineering and he is currently studying for a PhD. He is Assistant Professor of the Department of Economic Analysis and Company Management of the UDC. He also works in the GII of the UDC as a research engineer since 2008. He is mainly involved in the development of R&D projects related to industrial and logistical processes optimization. He also has developed projects in the field of human factors affecting manufacturing processes.

David del Rio Vilas holds an MSc in Industrial Engineering. He is Adjunct Full Professor at the Department of Economic Analysis and Company Management of the UDC. Since 2007, he works in the GII of the UDC in industrial processes improvement projects. He also leads the R&D Department of the Spanish civil engineering company Proyfe S.L.

Rosa Rios Prado works as a research engineer in the GII of the UDC since 2009. She holds an MSc in Industrial Engineering and now she is studying for a PhD. She has previous professional experience as an Industrial Engineer in different engineering companies. Her research areas are mainly devoted to the development of transportation and logistical models for the assessment of multimodal networks and infrastructures.

INTEGRATED SYSTEMS DESIGN IN AN AUTOMOTIVE INDUSTRY - USING CAD AND SIMULATION IN LAYOUT AND PROCESS OPTIMIZATION

Luis **Dias**[(a)], Guilherme **Pereira**[(b)], Pavel **Vik**[(c)] , José **Oliveira**[(d)]

ALGORITMI Research Centre,
Production and Systems Department,
University of Minho, Braga, Portugal
[(a)]lsd@dps.uminho.pt, [(b)]guilherme.pereira@algoritmi.uminho.pt, [(c)]vikpavel@seznam.cz, [(d)]zan@dps.uminho.pt

ABSTRACT
This paper discusses production systems design issues, applied to an internal logistic system in the automotive industry.
In this paper, production systems design software tools – Simulation and Computer Aided Design are integrated, exploring ways of dealing with data diversity and assuring valid and efficient production systems, taking advantage of the mentioned data integration.
This integration is implemented on AutoCAD (layout design) and WITNESS (simulation), using MS Access as the system knowledge repository. The software package developed was called IDS (*Integrated Design of Systems*).
This approach supports global system optimization that considers all important system resources and system performance measures.
Solutions achieved are expected to be better than solutions obtained with non-integrated approaches.
IDS approach is open and accessible, thus enabling different companies to use this advanced production systems design tool, taking advantage of simulation and CAD systems and their integration.
This application intends to validate the concept and functionalities of the proposed tool, on a real industrial case study.

Keywords: Integrated Systems Design; Production Systems Planning and Design; CAD and Simulation; Layout Optimization.

1. INTRODUCTION

1.1. Description of the PSPD problem

Production **S**ystem **P**lanning and **D**esign (PSPD) is a complex set of tasks using knowledge from several fields: scientific, logical, economical, management, statistical, technical and information technology. It consists of planning and evaluating different alternatives of systems aiming the global optimal usage of inputs and all kinds of resources. Alternatives are designed regarding dynamic time changes and stochastic influences (Francis and White 1974)(Heragu 2006).

Nowadays, there is a great pressure on production systems design to be developed or reorganised rapidly and efficiently due to the worldwide competitive market and rapid progress in manufacturing processes. In this dynamic context, flexibility, modularity and robustness are desired production system properties.

This paper deals with production systems design and its improvement. It is focused on the design of systems and layouts based on material flows, on re-layout processes and also on the design of layouts influenced by different types of uncertainties. It discusses production systems design issues, applied to an internal logistic system in the automotive industry.

As far as Production Systems Design is concerned, three basic classes of software tools have been used: **Computer Aided Design**, Process **Simulation** and **Information** Systems. However, these software tools have been used with low levels of integration. The absence of data integration within these three classes of software tools, and also the absence of a systemic approach to Production Systems Design have been causing duplication of work, waste of time, incoherencies, difficulties in project team communication, and errors in the design phase.

The **integration** is implemented through ***I**ntegrated **D**esign of **S**ystems* (IDS) tool, which uses AutoCAD (layout design), WITNESS (simulation) and Microsoft Access (database) (see Figure 1), and makes use of the issues discussed in Vik et al. (2010b)(2010c), concerning the production system design software tool developed and presented in the above papers. Pandey et al. (2000) make an interesting contribution towards this type of tools integration, developing a model that is optimised by simulation and then adapted the results into a layout. Also, Altinkilinc (2004) improved the system with simulation and then used a CRAFT method for layout optimization.

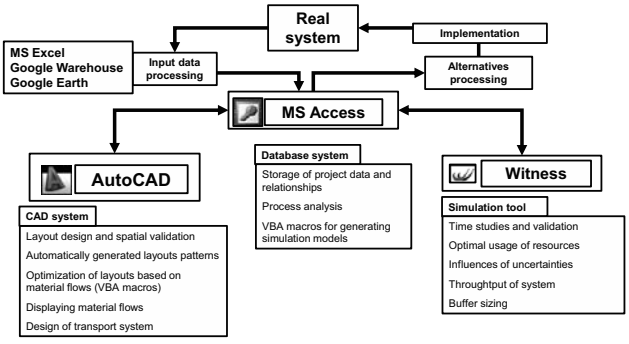


Figure 1 - IDS Overview

MS Access provides an open database structure, allowing integration and data exchange between WITNESS and AutoCAD. Simulation helps on dynamic systems analysis and CAD on static arrangement on a feasible implementation. Iteratively the results from the simulations are used to improve CAD layout design, and CAD layouts are used in new simulation experiments. This approach supports global system optimization that considers all important system resources and system performance measures.

According to Grajo (1995), *layout optimization* and *simulation* are two tasks that are fundamental to facility planning. Burgess et al. (1993) proclaimed that simulation is the only methodology robust enough to the systematic examination of key variables of factory performance. Simulation methodology enables the representation of many attributes of real life problems that are difficult to consider in analytical models for the layout optimization (Tam and Li 1991)(Tang and Abdel 1996)(Pandey et al. 2000)(Castillo and Peters 2002).

The main differences between the traditional (non-integrated) approach and IDS are shown in Figure 2. The traditional approach (upper image) often uses tools separately or with minimum relative integration and data can be stored in several places and in different formats. On the other hand, IDS makes use of a full integration (lower image). A similar idea of integration is described in other works (Chee 2009)(Benjaafar and Sheikhzadeh, 2000)(Sly and Moorthy, 2001).

Figure 2 – Traditional and proposed (IDS) approaches

Integration is managed by one common database. It allows specifying and controlling simulation model from database and read/write the required/received data.

The IDS tool **generates** automatically the **simulation model** and shows several alternatives and provides detailed information on production systems performance measures subject to different designs/configurations, enabling to choose the best solution.

Solutions achieved are expected to be better than solutions obtained with non-integrated approaches.

IDS approach is open and accessible, thus enabling different companies to use this advanced production systems design tool, taking advantage of simulation and CAD systems and their integration.

This application intends to validate the concept and functionalities of the proposed tool.

1.2. Case study description

This case study is focused on the internal logistic system in the automotive industry. The company "Magna Exteriors and Interiors" ("Cadence Innovation" until 2009) is a producer and designer of plastic parts.

It was set up in 1946 and started by producing plastic parts for kitchen and garden. Since 1982, it has been producing plastic parts for the automotive industry, such as: painted bumpers (33%, around 4300 per day), control desks (38%, 3000 per day), door fillers (22%, 6200 per day) and grid of cooler (7%) in 2008.

This paper is focused on the production of bumpers (see illustration in Figure 1) and internal logistics linked with it. A similar topic was discussed on a project in 2007 and 2008 (Jareš, 2008)(Vik and Jareš, 2008)(Vik et al. 2010d).

The factory in Liberec City produces parts for five Škoda car models, four part types for each of them (front and rear bumpers, central strips, front grids). This project involves only "big parts" (bumpers), while "small parts" (strips, grids) are omitted once its production is independent of the bumpers on which this project is focused.

There are 21 different colours available for regular bumpers. The combination of colour and car model (part type) is named by the term **"colour-type"** (**CT**). Every car model is limited to a specific set of colours, as shown in Figure 3. Occasionally, non-standard colours are used for special customers (police, taxis, companies, etc.), yet these colours are omitted here.

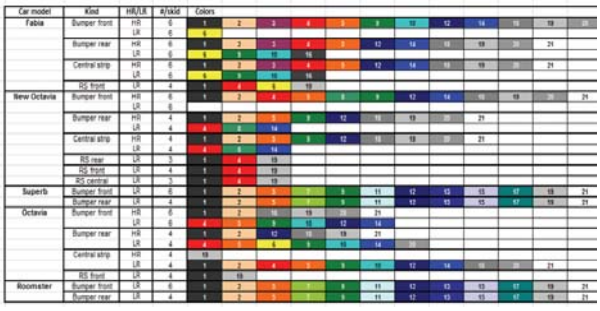


Figure 3 – Colour-Type (CT) table

The complete production processes are in Figure 4, and the factory layout in Figure 13.

Injection moulding machines produce non-coloured bumpers (**batch production**). According to the amounts of items in the "Warehouse of coloured parts", non-coloured bumpers are sent to the "Paint shop". This process is controlled by kanban pull system of orders that cares to hold the established level of products in the Warehouse (so called "safety level").

After the painting operations, painted bumpers (also called colour-type , CT) are transported by a

conveyor to the check quality workstations ("*Check WS*"), where quality is checked and visual defects are brushed and polished. After that, CTs are hung on another conveyor, moved close to the "Warehouse" and stored there in special containers (crates) (see Figure 5).

Based on customer orders, CTs are removed from the Warehouse, assembled and dispatched to the customers. This part of logistics system is based on the JIT principle ("Just-In-Time").

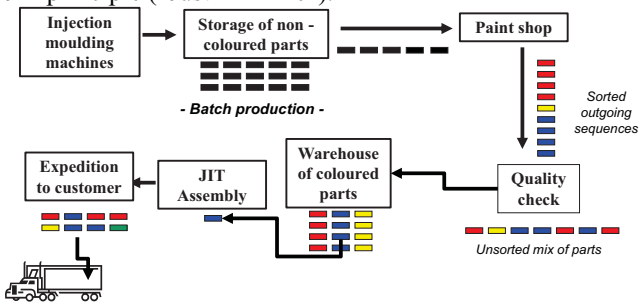


Figure 4 – Schema of production

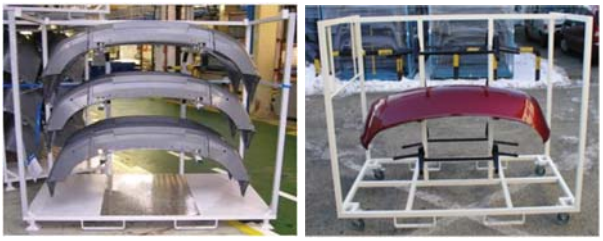


Figure 5 – Transportation units for coloured bumpers

The approach proposed, through IDS tool, should be helpful to acquire the following information:

- Necessary number of brushing operators (Check Quality workstations), number of operators for the take-down operation and forklift drivers
- Utilization rates for system resources
- Size of the buffer and storages (mainly temporary floating buffer and Warehouse)
- Condition of holding a safety level in the storages
- Required number of containers in the system
- Paint shop scheduling approaches/solutions

For this purpose, a simulation model is built (automatic generation through IDS) and a set of experiments is run. Required input data are processed and analysed in the following section. This model can also be used for testing different paint shop schedules.

2. SYSTEM DATA ANALISYS

2.1. Paint shop data processing

Data about product arrivals are taken from production data set, provided by company´s ERP system. That data is recorded in the check workstations where each worker must save information about products in the system (barcodes). These data can be used as schedules for paint shop. Production data consists of time (date), car model, CT name, colour, number of parts, and number of scraps (see Figure 6).

Date	Time	Car model	Part	Color	Count	Scraps	%
2007.03.20	8:09:12	Roomster	Bumper front	Ocean blau	60	1	1
2007.03.20	8:20:26	Fabie	Bumper front	Anthracite	48	2	4
2007.03.20	8:27:52	Fabie	Bumper front	Anthracite	54	0	0
2007.03.20	8:40:52	Roomster	Bumper rear	Capuccino beige	24	2	8
2007.03.20	8:43:55	Superb	Bumper rear	Capuccino beige	20	0	0
2007.03.20	8:51:02	Roomster	Bumper front	Capuccino beige	60	6	10
2007.03.20	9:00:49	New Octavia	RS front	Diamantsilber	24	0	0
2007.03.20	9:08:27	New Octavia	RS rear	Diamantsilber	15	2	13
2007.03.20	9:11:48	Fabie	Bumper rear	Diamantsilber	42	3	7
2007.03.20	9:18:10	Fabie	Central strip	Satin	10	0	0
2007.03.20	9:19:43	Fabie	Central strip	Diamantsilber	6	4	66
2007.03.20	9:22:44	Fabie	Central strip	Diamantsilber	10	0	0
2007.03.20	9:36:53	New Octavia	Bumper rear	Satin	32	0	0
2007.03.20	9:51:46	New Octavia	Bumper front	Satin	36	0	0
2007.03.20	9:52:31	Fabie	Bumper front	Black magic	54	1	1
2007.03.20	9:59:16	Octavia	Central strip	Black magic	3	0	0
2007.03.20	10:01:24	Octavia	Central strip	Diamantsilber	2	0	0
2007.03.20	10:12:30	Fabie	Bumper rear	Black magic	36	0	0
2007.03.20	10:21:06	New Octavia	Bumper rear	Capuccino beige	28	1	3
2007.03.20	10:27:11	New Octavia	Bumper front	Capuccino beige	20	0	0
2007.03.20	10:35:49	Octavia	Bumper front	Black magic	48	0	0
2007.03.20	10:45:18	New Octavia	RS front	Black magic	6	1	16
2007.03.20	10:48:40	Octavia	Bumper rear	Black magic	32	1	3
2007.03.20	10:54:04	New Octavia	Bumper front	Diamantsilber	32	0	0
2007.03.20	11:03:44	Fabie	Bumper front	Candy weiss	12	4	33
2007.03.20	11:08:28	Fabie	Bumper front	Candy weiss	54	3	5
2007.03.20	11:17:03	Roomster	Bumper rear	Corrida	48	0	0
2007.03.20	11:32:52	Octavia	Bumper front	Anthracite	32	1	3
2007.03.20	11:41:30	New Octavia	Bumper front	Candy weiss	24	0	0
2007.03.20	11:44:53	Roomster	Bumper front	Tangarine orange	54	8	14
2007.03.20	12:04:20	Fabie	Bumper rear	Capuccino beige	48	0	0
2007.03.20	12:09:49	Fabie	Bumper front	Capuccino beige	54	2	3
2007.03.20	12:19:25	Roomster	Bumper rear	Diamantsilber	48	0	0
2007.03.20	12:27:18	Roomster	Bumper front	Capuccino beige	54	2	3

Figure 6 - Paint shop data

2.2. P-Q analysis

Production data is summarised and the P-Q analysis is processed. Figure 7 shows production volumes for each CT during a chosen week.

Product portfolio varies – currently 103 CTs are considered. It is not necessary to simulate all of them once they show similar properties and behaviour. For the purpose of the simulation model, 10 CTs (2 HR, 8 LR) were chosen – see arrows in top image of Figure 7.

HR and LR stand, respectively, for "High-runners" (large-size production batch) and "Low-runners" (small-size production batch).

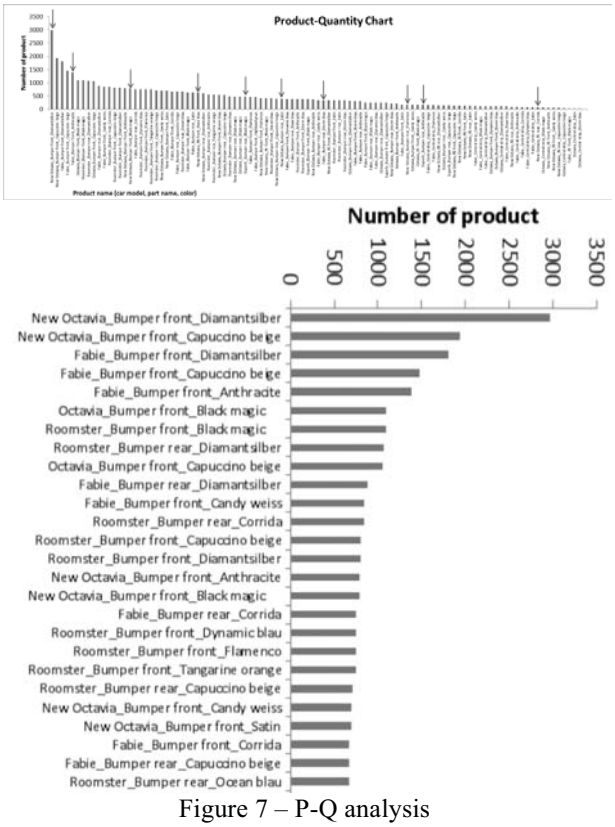


Figure 7 – P-Q analysis

2.3. Output quality

From the data set shown in Figure 6, it was synthesized the percentage of conformity and unconformities. The attribute "vol" was defined as the "quality level" (1..5) and results aggregated in Table 1.

Table 1 – Part quality analysis

Quality level	%	vol
Perfect products	60%	1
Brushing	25%	2
Local repainting	5%	3
Complete repainting	5%	4
Waster	5%	5

2.4. Operation times of Check Workplace

These values were measured and an analysis was performed. All of them have *normal distributions*, as shown in Table 2.

Table 2 – Operation times

Operation	Mean [seconds]	Standard deviation [seconds]
Checking(1)	40	10
Polishing and brushing(2)	160	40
Checking (repaint, waster)(3,4,5)	20	3
Hang on /Take down of bumpers	10	5
Transportation and storing in the Warehouse	120	60
Assembly operation	300	50
Assembly (packing for resellers)	45	7
Painting	10800 (3 hours)	500

2.5. JIT orders, analysis and data preparation

JIT customers' orders come in a very random order. For feeding IDS system, VBA macro in MS Excel generates randomly JIT demands by the following rules (see Figure 8):

- The total percentage of CTs is based on volumes outgoing from the Paint shop,
- 25% of the demands are JIT assembly for direct factory customers, 75% of the demands are for resellers (determinated by *vol* attribute)
- JIT orders are generated at a rate that corresponds to 5500 parts per day

2.6. Safety level of resources in the Warehouse

Customers also demand holding safety storage level of bumpers, usually for two days. The consumption for two days, for the chosen CT, is in Figure 9, as well as the number of full loaded containers.

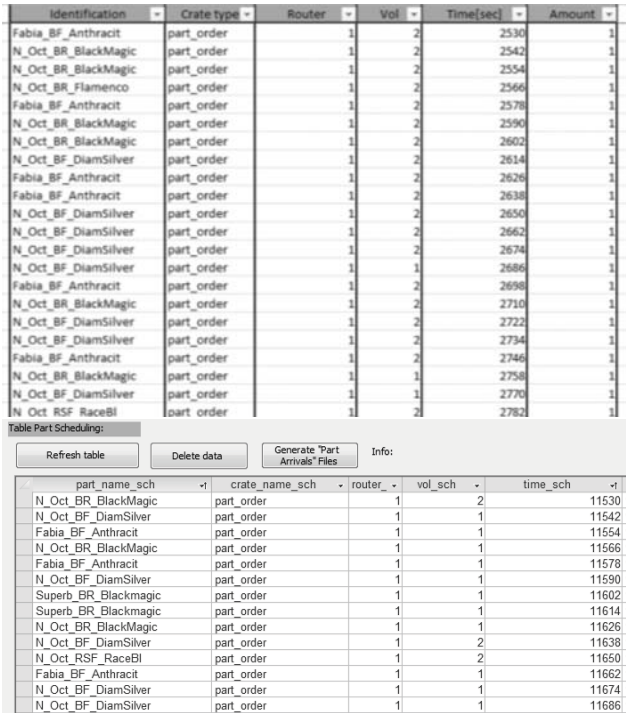


Figure 8 - JIT orders analysis

type	colour	2 days consumption	# bumpers in the skid	#containers
New Octavia, bumper rear	blackmagic	400	8	50
	cappucino	260	8	33
	corrida	40	8	5
	flamengo	50	8	7
	sprint	50	8	7
	island	40	8	5
	highland	20	8	3
	stormblau	100	8	13
	dynamic	60	8	8
	anthracit	250	8	32
	satin	120	8	15
	brilliant	410	8	52
	candy	200	8	25
		2000		***255***

Figure 9 – JIT order analysis

After the complete data are collected, the simulation model can be developed, experiments run and final reports would be available. This type of approach is described in the work of Mecklenburg (2001) and is also based on the author´s experience.

3. PROCESS SPECIFICATION AND SIMULATION MODEL

All the process is specified in the IDS database, then IDS generates de Simulation Model and submits it to WITNESS, to run in batch mode (by default).

A schema of the simulation model is presented in Figure 10. Text boxes with blue background are buffers or conveyors, text boxes with yellow background represent real facilities and grey text boxes are logic control (virtual) processes (that are not in the real system). Flows represent transportation (conveyors, containers or final products), information flows are customers' orders or signals for paint shop.

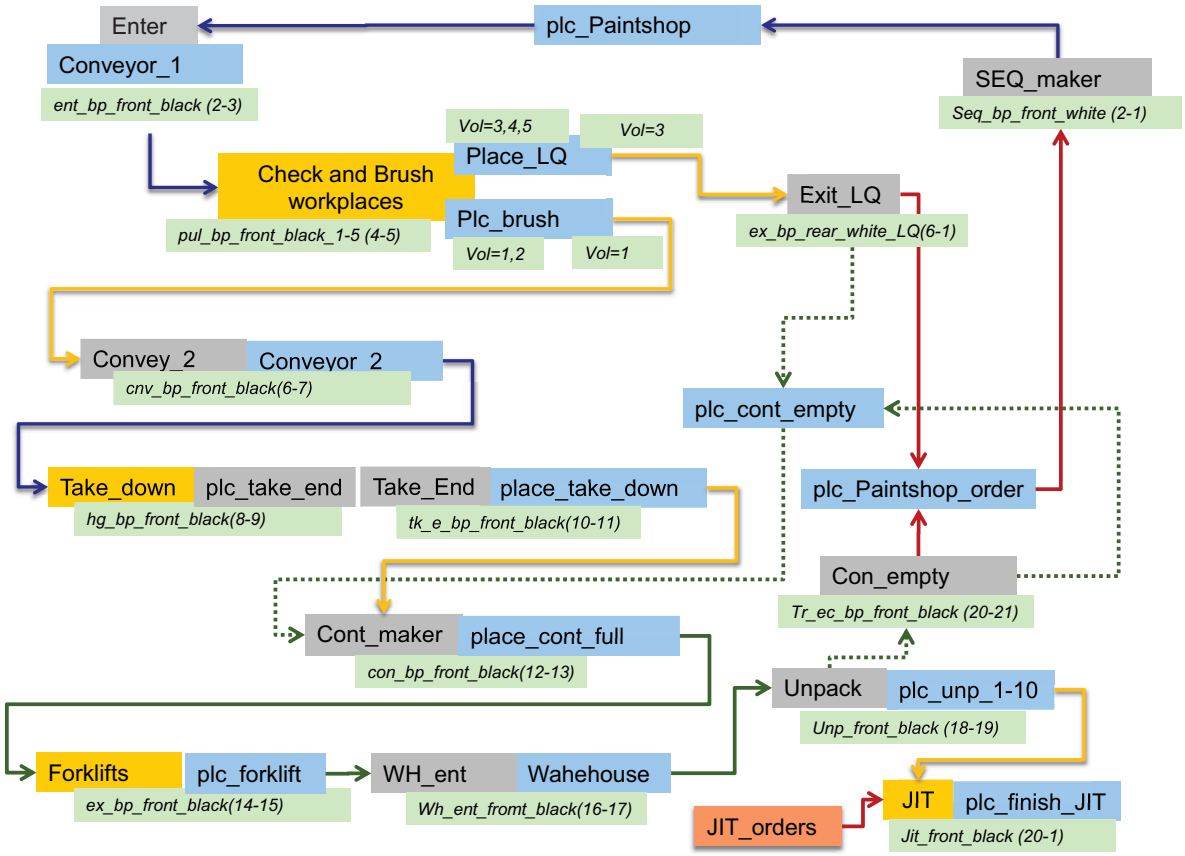
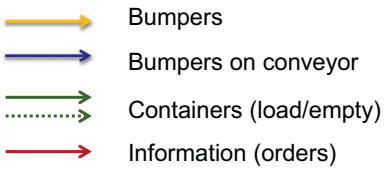


Figure 10 – Schema of the Process (to simulation model)



The *quantity* of resources can be easily changed as parameters in the experiments (e.g. *Checks Quality*, *Forklifts* and *Take down*).

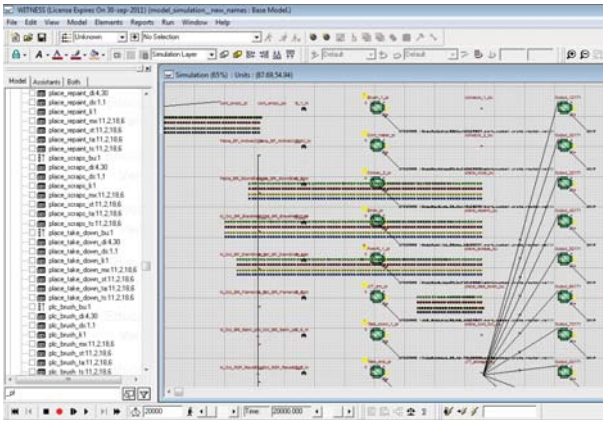


Figure 11 – IDS Generated simulation model

3.1. Brief description of the simulation model

JIT orders are the main inputting entities; all processes are controlled based on them. In a JIT assembly (*"JIT"*), the JIT orders are joined with the corresponding CT. CTs are taken from containers that are unpacked.

Then, empty containers are transported to the storage ("*Con_empty"*) that makes signal for the paint shop (*"plc_Paintshop_order"*) to produce new parts ("*SEQ_maker"*) in sequences of a rate 4:1 (HR:LR). The paint shop (*"plc_Paintshop"*) is represented by a buffer where a part must stay for a defined period, equal to the operation time. Parts going from the paint shop are taken from the conveyor ("*Conveyor_1"*) and checked (*"CheckWP"*). Parts with low quality are sent out (*"place_LQ"*), while the good parts are hung to the conveyor (*"Conveyor_2"*). From there, they are taken down ("*Take_down"*) and stored in the container (*"Cont_maker"*). Full containers are transported by forklifts to Warehouse (*"WH_ent"*). From Warehouse, containers are taken and unpacked (*"Unpack"*).

The simulation model is automatically generated based on data from DB (see Figure 11).

3.2. Results for the current production capacity

The number of containers in the system is estimated by basic calculations. These values must be validated with the simulation results to establish realistic safety levels.

Figure 12 shows a schema of the safety level analysis. Initial simulation model shows the number of full-loaded containers for a given CT (New Octavia Black Magic, HR) in the Warehouse. The simulation model shows between 10-20 full loaded containers in the Warehouse, with a calculated safety level of 41. This value is established by a 2-day production and the recalculated value into full-loaded containers is based on their capacity. Therefore, it must be increased the number of containers to accomplish holding safety level condition. In current state, there are the following number of facilities: *"Check WS"* (10), *"Forklifts"* (2), *"Take down"* (2) and *"JIT WS"* (10).

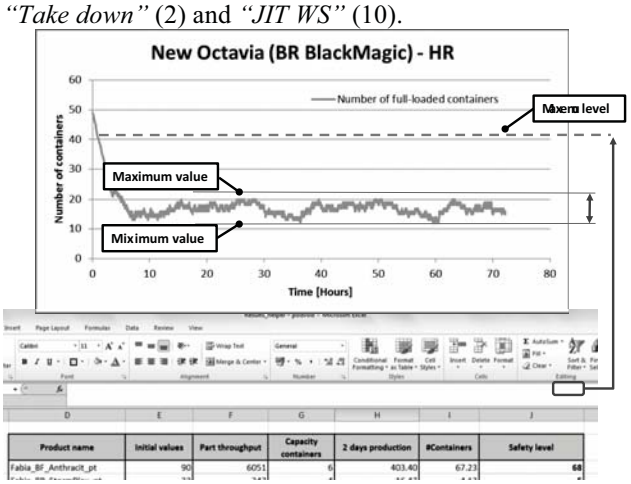


Figure 12 – Safety level determination

3.3. Experiment - increased production

In this experiment, the influences of changes of increased production in the system are studied. There are schedules tested for 10 000 daily production on the current system (paint shop outgoing sequences, operation times, quality).

The simulation model used in the increased production analysis is the same as used in the current system. Several different facilities quantities have been tested in order to guarantee the desired safety level. The main results of this experiment are in Table 3 and define the following number of facilities: *"Check WS"* (10), *"Forklifts"* (**3**), *"Take down"* (2) and *"JIT WS"* (**14**).

The total throughput during three days is 29985 that corresponds to a 10 000 daily production. The utilization of "*CheckWS*" is around 90% and forklifts are busy on 64%. The assembly workstations, work on an average of 90%. The total stable number of full-loaded containers in the Warehouse is around 150, without considering safety storages of each CTs.

Results of the experiment enable the selection of an adequate configuration for the production increase.

Table 3 – Results of increased production experiment

Number of facilities – optimal configuration:

Name	Quantity
Check WS	10
Forklifts	3
Take down	2
JIT WS	14

Throughput:

part_throughput	crate_throughput	value_throughput
N_Oct_BR_BlackMagic	part_order	6758
Superb_BR_Anthracit	part_order	522
N_Oct_BR_Flamenco	part_order	495
N_Oct_RSF_RaceBl	part_order	1305
Fabia_BF_Anthracit	part_order	8172
Fabia_BR_StormBlau	part_order	342
N_Oct_RSR_Corrida	part_order	107
Superb_BR_Blackmagic	part_order	684
N_Oct_BF_DiamSilver	part_order	11177
N_Oct_BR_Satin	part_order	423

Facility usage:

Total time | Percentage

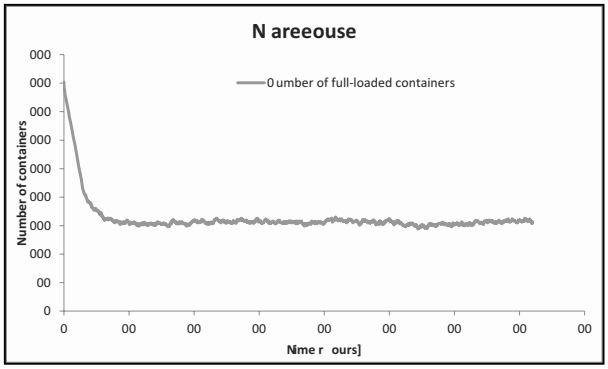
machine_name	kind_of_process	busy_time	blocking_time	waiting_time_part
Brush_1	quantity_function	82.1372809960282	0	17.8627190039718
Take_down_1	quantity_function	27.4269847742791	0	72.5730152257209
Forklift_1	quantity_function	64.7818167627711	0	35.2181832372289

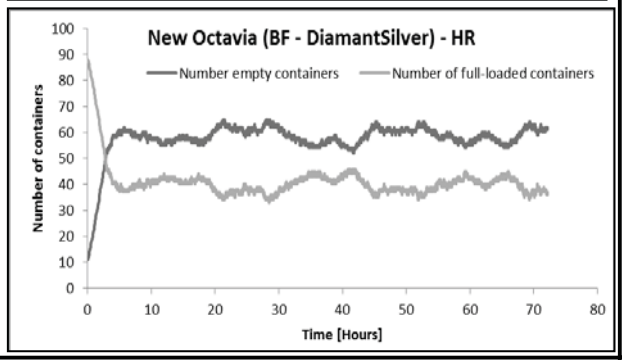
JIT workstations:

Total time | Percentage

machine_name	kind_of_process	busy_time	blocking_time	waiting_time_part
JIT_12	machine	93.7801549047754	0	6.21984509522463
JIT_13	machine	94.6475897333336	0	5.35241026666645
JIT_14	machine	95.7220320434979	0	4.27796795650209
JIT_2	machine	66.9980269493993	0	33.0019730506007
JIT_3	machine	74.8159746992503	0	25.1840253007497
JIT_4	machine	91.5009076732587	0	8.49909232674135
JIT_5	machine	92.2983415484723	0	7.70165845152769
JIT_6	machine	86.2158556645025	0	13.7841443354975

Graphs of containers occupancy in the Warehouse:





4. CAD LAYOUT INTERACTIVE DESIGN WITH IDS

The Design and Improvement (Optimization) of the factory CAD Layout is supported by the IDS system in a very convenient way.

The specification of the material flows and processing's is previously made on the Database by fulfilling proper dialogs and tables, centralizing all the knowledge in the database.

IDS generated automatically a simulation model that runs on WITNESS. Embedded code in the simulation model, interpret simulation results of the system simulation runs and feeds them back into de database, enriching the system knowledge with facts as the intensity of material flows through each path and the needs for real buffer sizes (punctual accumulation of parts, during factory labour).

IDS supports the interactive process of improving successively the CAD layout, evaluating its performance or feasibility through simulation experiments.

4.1. Material flows

Figure 13 identifies the production areas. In Figure 15, there are several layouts with generated material flows. The actual factory layout is on the background (grey colour). The top left image displays schematic layout with direct unconstrained material flows of all CTs and the top right image shows the chosen CT flows. For the design of a transportation aisle, it can be helpful to display flows based on a transportation unit –the bottom left image shows container flows. These flows are generated based on the transportation network of aisles. The bottom right image shows flows based on realistic shapes of aisles, conveyors and transportation roads.

4.2. Buffer size establishment

In order to find the optimal buffer (storages, warehouse) size, it is possible to integrate simulation and CAD results. This approach integrates results from the simulation *("maximum_value_simulation")* and CAD layout *("limitation_CAD")* as shown in Figure 14. Simulation results provide the maximum number of units (e.g. containers) in the given buffer during the simulated period, and in the CAD layout, there is a specific available space for these containers. In Figure 14, data for the Warehouse with available space provided by the layout for each CT is presented. In this case, there is enough space for all CT containers, as also shown in Figure 16.

The appropriate number of blocks representing containers is generated to the layout (see Figure 16, upper image - orange coloured blocks). Those Blocks representing 3D containers are automatically inserted into the layout. These blocks are arranged and placed in the correct position into shelves or stacked (see Figure 16, lower image). This is helpful to the realistic design of a warehouse structure, stocking containers in layers, 3D layout, the required manipulation for space (based on material flows and aisles) and the validation of the required space (the maximal number of parts that can be in the given buffer). IDS doesn't generate 3D animated models, although considers the specification of vertical accumulation. Previous authors experience could be relevant to include 3D animation for illustrative purpose on IDS (Vik et al. 2008a, 2008b, 2010a, 2010b).

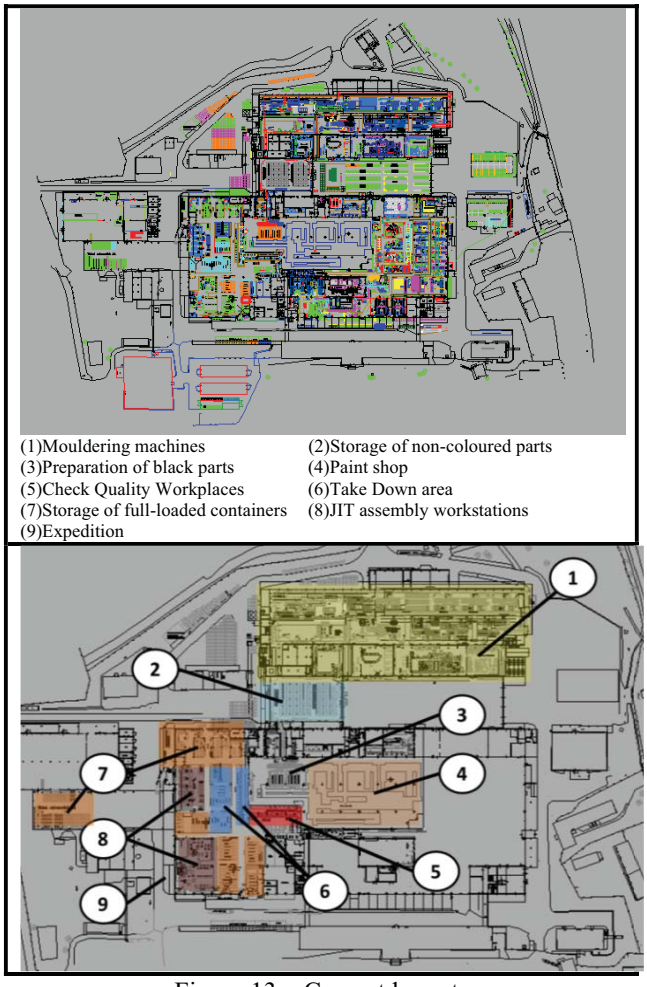


Figure 13 – Current layouts

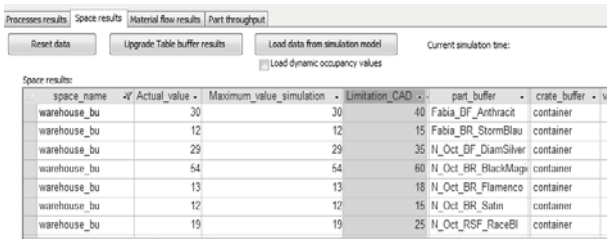


Figure 14 – Warehouse/Buffers size analysis results

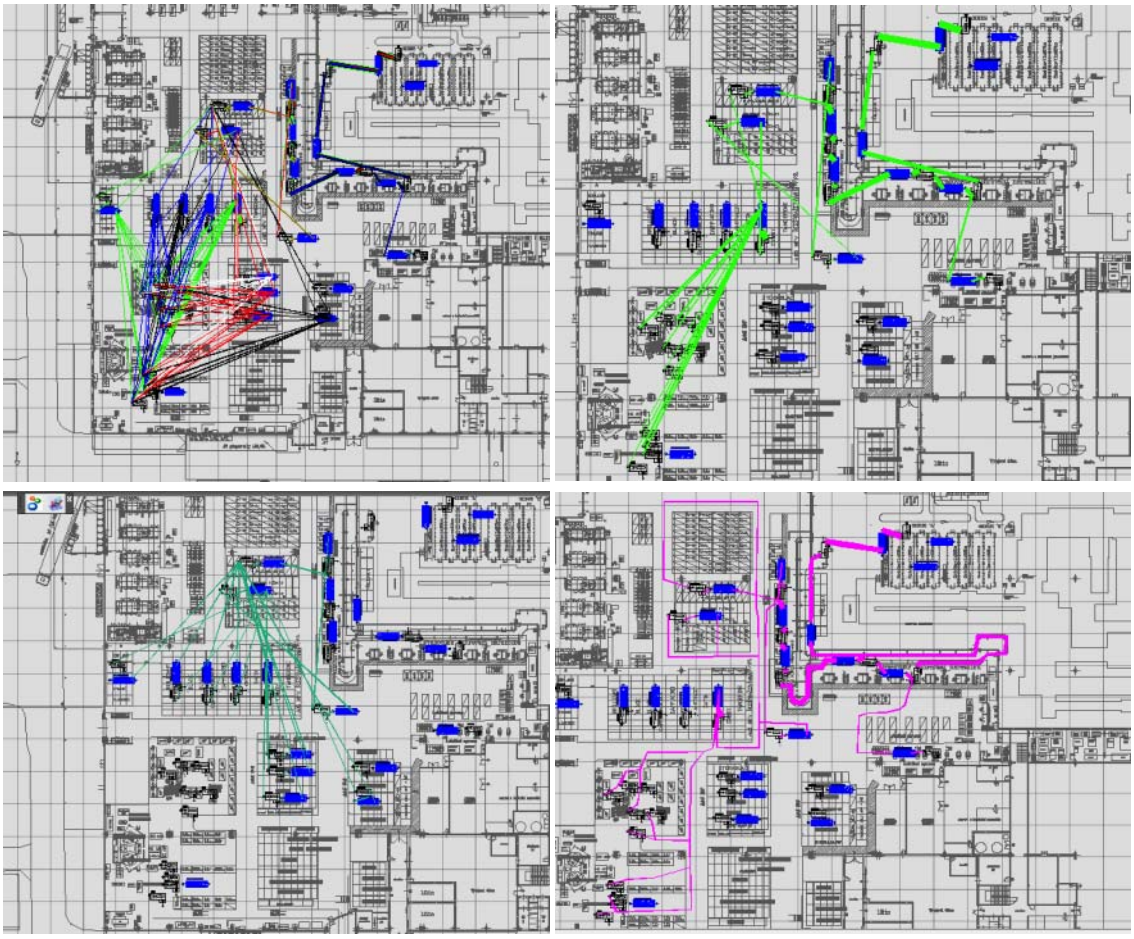


Figure 15 – Layouts with material flows

Figure 16 – Buffers space usage and 3D layout

	Traditional approach	IDS
Project aims establishing	2 days	2 days
Data analysis and input to DB (IDS)	3 days	3 days
Simulation model building	2 weeks	5 minutes
Experiments /alternatives design	3 hours	8 hours
CAD layout 2D	2 days	3 hours
CAD layout 3D	-	1 hour
Results analysis	8 hours	1h
Total time	4 weeks	1.5 weeks

Figure 17 - Time requirement summary

5. CONCLUSIONS

This project solves internal logistics in a factory of an automotive supplier, more precisely the logistics between the paint shop (batch production) and the JIT assembly controlled by customers' orders. To test the planned increased production (doubled), a simulation model was developed (generated) in IDS.

With this model, the optimal configuration was established as well as the safety level of containers in the Warehouse was estimated.

The required time for the complete work using IDS was 1.5 week – traditional approach needed 4 weeks to complete the job (see **Figure 17**).

The use of IDS tool has proved some important advantages of this integrated approach as opposed to the traditional non-integrated approach:

- Bidirectional information flows, i.e. inclusive feedbacks
- One common database
- Easy data transfer and unified data format
- Fast reaction to changes during the project elaboration
- Full use of software tools and their extension by new functions

REFERENCES

Altinkilinc, M. 2004. Simulation-based layout planning of a production plant. In *Proceedings of the 2004 Winter Simulation Conference*, 2:1079–1084. IEEE Computer Society Press.

Benjaafar, Saifallah, and Mehdi Sheikhzadeh. 2000. "Design of flexible plant layouts." IIE Transactions 32 (4) (April): 309-322. doi:10.1080/07408170008963909.

Burgess, AG, I. Morgan, and TE Vollmann. 1993. "Cellular manufacturing: its impact on the total factory." *The International Journal of Production Research* 31 (9): 2059–2077.

Castillo, Ignacio, and Brett Peters. 2002. "Unit load and material-handling considerations in facility layout design." *International Journal of Production Research* 40 (13) (January): 2955-2989.

Chee, Ailing. 2009. Facility layout improvement using systematic layout planning (SLP) and ARENA. Master thesis. Universiti Teknologi Malaysia.

Francis, Richard, and John White. 1974. *Facilities Layout and Location*. New Jersey: Prentice Hall, Englewood Cliffs.

Grajo, E.S. 1995. Strategic layout planning and simulation for lean manufacturing a LayOPT tutorial. In *Proceedings of the 27th Winter simulation conference*, 510–514. IEEE Computer Society.

Heragu, Sunderesh S. 2006. *Facilities Design*. Second edi. Lincoln: iUniverse.

Jareš, David. 2008. Počítačová simulace transportního systému v podniku X (Computer simulation of transport system in the factory). Bachelor thesis. Technical University of Liberec.

Mecklenburg, Karsten. 2001. Seamless Integration of Layout and Simulation. In *Proceedings of the 2001 Winter Simulation Conference*, 1487. Citeseer.

Pandey, P. C., S. Janewithayapun, and M. a. a. Hasin. 2000. "An integrated system for capacity planning and facility layout." *Production Planning & Control* 11 (8) (January): 742-753.

Sly, David, and S. Moorthy. 2001. Simulation data exchange (SDX) implementation and use. In *Proceedings of the Winter Simulation Conference*, 2:1473–1477. IEEE Computer Society Press.

Tam, K.Y., and S.G. Li. 1991. "A hierarchical approach to the facility layout problem." *International Journal of Production Research* 29 (1): 165–184.

Tang, C., and L.L. Abdel-Malek. 1996. "A framework for hierarchical interactive generation of cellular layout." *International journal of production research* 34 (8): 2133–2162.

Vik, Pavel, and David Jareš. 2008. Využití počítačové simulace v interní logistice (Usage of computer simulation in the internal logistics). In *Sborník příspěvků 11. ročníku mezinárodní konferenc*, 81-85. HUMUSOFT s.r.o. & VUT Brno.

Vik, Pavel, Luis Dias, and José Oliveira. 2008a. Using of 3D Simulation in the Industry. P*aper presented at the conference 2nd. International Conference of Technology Knowledge and Information (ICTKI)*. Ústí nad Labem.

Vik, Pavel, Luis Dias, José Oliveira, and Guilherme Pereira. 2008b. Using 3D Simulation in an Internal Logistic Process. In *Proceedings of 15th European Concurrent Engineering Conference (ECEC2008)*, 116-120. Porto: Porto April 2008. EUROSIS-ETI Publication.

Vik, Pavel, Luis Dias, Guilherme Pereira, José Oliveira, and R. Abreu. 2010a. Using Simio for the Specification of an Integrated Automated Weighing Solution in a Cement Plant. In *Proceedings of the 2010 Winter Simulation Conference*, 1534-1546. Baltimore, USA.

Vik, Pavel, Luis Dias, Guilherme Pereira, José Oliveira, and Ricardo Abreu. 2010b. Using SIMIO in the Design of Integrated Automated Solutions For Cement Plants. In *Workshop on Applied Modelling and Simulation*. Rio de Janeiro, Brasil: Universidade Federal do Rio.

Vik, Pavel, Luis Dias, Guilherme Pereira, and José Oliveira. 2010c. Improving Production and Internal Logistics Systems - An Integrated Approach Using CAD and Simulation. In *Proceedings of the 3rd International Conference on Information Systems, Logistics and Supply Chain*. Casablanca (Morocco).

Vik, Pavel, Luís Dias, Guilherme Pereira, and José Oliveira. 2010d. Automatic Generation of Computer Models through the Integration of Production Systems Design Software Tools. P*aper presented at the conference 3th International Conference on Multidisciplinary Design Optimization and Applications*. Paris, France.

BIOGRAPHIES

LUÍS S. DIAS was born Portugal. He graduated in Computer Science and Systems Engineering at the University of Minho (Portugal). He holds an MSc degree in Computers Engineer. He holds a PhD degree in Production and Systems Engineering (Simulation) from the University of Minho. He is Assistant Professor at University of Minho. His main research interests are Modelling, Simulation and Optimization.

GUILHERME PEREIRA was born Portugal. He graduated in Industrial Engineering and Management in the University of Minho. He holds an MSc degree in Operational Research and a PhD degree in Manufacturing and Mechanical Engineering from the University of Birmingham, UK. He is Associate Professor at the University of Minho. His main research interests are Operational Research and Simulation.

PAVEL VIK was born in Czech Republic. He studied at the Technical University of Liberec where he obtained his MSc degree in Manufacturing Systems. He is making doctoral studies at the University of Minho. His main research interests are 3D animation and virtual reality in design of production systems and visualization of computer simulation and automatic generating of simulation models.

JOSÉ A. OLIVEIRA was born in Portugal. He studied Mechanical Engineering at the University of Porto. He holds a Ph.D. in Production and Systems Engineering from the University of Minho. His main research interests are Optimization with Heuristic Methods in Systems Engineering.

MODELING AND THERMO-FLUID DYNAMIC SIMULATION OF A FRESH PASTA PASTEURIZATION PROCESS

Eleonora Bottani[(a)], Gino Ferretti[(b)], Matteo Folezzani[(c)], Michele Manfredi[(d)], Roberto Montanari[(e)], Giuseppe Vignali[(f)]

[(a), (b), (e), (f)] Department of Industrial Engineering, University of Parma, viale G.P. Usberti 181/A, 43124 Parma (Italy)
[(c) (d)] CIPACK Interdepartmental Centre, University of Parma, viale G.P. Usberti 181/A, 43124 Parma (Italy)

[(a)]eleonora.bottani@unipr.it, [(b)]gino.ferretti@unipr.it, [(c)]matteo.folezzani@unipr.it, [(d)]michele.manfredi@unipr.it, [(e)]roberto.montanari@unipr.it, [(f)]giuseppe.vignali@unipr.it

ABSTRACT

The present work aims to analyse a thermal process for pasteurization of fresh filled pasta, by means of Computational Fluid Dynamic (CFD) simulation. The pasta considered is "ravioli" filled of meat. Thanks to many studies on pasta proprieties (Saravacos and Maroulis, 2001 and de Cindio et al., 1992), some product parameters, such as thermal conductivity and heat capacity, have been determined.

All simulations have been performed using Ansys CFX code version 14.0 in a transient state (after 30s, 60s, 120s and 150s), to evaluate the pasteurization temperature at the core of the "ravioli" as a function of the process time. The heat exchange takes place in a pasteurization tunnel by means of water vapour at approx. 96°C.

Finally, experimental tests were performed in order to validate the simulation model of heat exchange. Results show a good agreement with the real pasteurization process and a good level of product quality.

Keywords: heat treatment, CFD simulation, fresh filled pasta, pasteurization process.

1. INTRODUCTION

The pasteurization process is one of the most important steps of the industrial packaging of fresh pasta. This phase should ensure a safe and healthy product. The pasteurization of the fresh filled pasta makes use of a heat treatment, with the aim of reaching a commercial sterility level of the food. The purpose is to reduce all vegetative forms of non-spore-forming pathogenic microorganisms. The timing and temperature of the treatment are established on the basis of the elimination of most heat resistant vegetative forms, such as Salmonella.

In this article, we consider a specific type of pasta, called "ravioli", filled of meat. The storage temperature of this product, throughout the food chain, is kept at 4 ° C ± 2 ° C. However, the presence of microorganisms in the product is inevitable: all technological operations can only reduce the number of microorganisms, but, overall, it remains impossible to completely eliminate them. To provide consumers with an acceptable safety level, the international standard EN ISO 11290-2 and a specific Report ISTISAN89/9 defined the maximum allowable level of contamination as *cfu* (Colony Forming Units) for fresh filled pasta, at the end of production process (Table 1).

MICROORGANISM	LIMIT	REFERENCE
Total microbial	max 10^6 cfu/g	Rep. ISTISAN 89/9
Staphylococcus Aureus	max $5x10^2$ cfu/g	Rep. ISTISAN 89/9
Clostridium Perfringens	max 10^3 cfu /g	Rep. ISTISAN 89/9
Salmonella	absent in 25 g	Rep. ISTISAN 89/9
Listeria Monocytogenes	max 100 cfu/g	EN ISO 11290-2
Bacillus Cerus	max 10^4 cfu/g	Rep. ISTISAN 89/9

Table 1- Maximum level of cfu for the production of fresh egg pasta (International standard EN ISO 11290-2, Report ISTISAN 89/9)

In order to understand the correct value of pasteurization, knowing the quality of raw materials is also paramount. In fact, the microbiological quality of fresh pasta depends not only on the hygiene of workers and environment where it is produced and preserved, but above all on the raw materials used to produce it and on the production technologies. Table 2 reports the maximum contamination level of the flour allowable for the production of fresh pasta (Pagani, 2007).

CBT	max 30-40.000 cfu/g
Yields	max 500 cfu/g
Molds	max 500 cfu/g
Salmonella	absent in 25 g
Staphylococcus	max 100 cfu/g
Enterobacteria	max 100 cfu/g
Fecal coliform	max 100 cfu/g
Bacillus cereus	absent in 1 g

Table 2 – Maximum contamination level of the flour for the production of fresh pasta

The fresh pasta is pasteurized to avoid the presence of pathogenic microorganisms and reduce the

saprophyte micro flora to acceptable limits; as a result, it ensures a longer shelf life. During the heat treatment, the pasta filling contains a large amount of water, and, when the temperature increases, a partial evaporation occurs. The resulting water vapour remains entrapped inside the product. Thus, a particular way of cooking is generated, able to give the final product a typical taste.

In industrial processes, the heat treatment is obtained by putting the final products on a conveyor belt, moving inside a steam tunnel which operates at pressure close to 1 atm. Typical processing times are around a few minutes.

During the heat treatment, the vapour produced inside the product generates a water pressure that overcomes the external pressure. Thus, the dough shell is subject to a pressure gradient that deforms it and, under critical conditions, could break it. Thick shells could be used to prevent this undesired effect, but in practice this solution is never adopted because the dough shell should be very thin so as to enhance as much as possible the filler flavours. It should be noted that thin but resistant shells are necessary to obtain a good quality product (De Cindio et al., 2000).

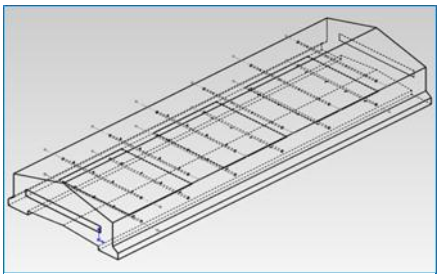


Figure 1. Geometry of pasteurizer

Based on these premises, many researchers try to optimize the pasteurization process of fresh pasta. A traditional approach makes use of experimental tests, changing the main parameters of the process, i.e. pressure, temperature and time (Alamprese et al, 2005; Alamprese et al, 2008; Rizzo et al., 2010).

In the last twenty years, in order to reduce the experimental cost, more and more research activities have been conducted through simulation tools. In particular, Computation Fluid Dynamic (CFD) simulation appears as one of the most used tools to model heat transfer process of food products (Zou et al, 2006, I and II;).

Some research activities about heat transfer in fresh filled pasta concern the simulation of drying phase (e.g., Migliori et al., 2005), using specific formulation for the heat transfer inside the pasta. No scientific works, however, have been found about the application of CFD to the pasteurization process of fresh filled pasta.

In this work, thanks to the support of a numeric solver, simulations were conducted to evaluate the temperatures reached inside the fresh filled pasta during pasteurization, as a function of time. Then the temperature values obtained from experimental tests were reported.

2. MATERIALS AND METHODS

2.1 Materials

The fresh filled pasta considered for the present study has a high moisture content (humidity of at least 24%) and water activity (Aw) between 0.92 and 0.97; for this reason, it can be right considered a perishable product. The main ingredients of the fresh pasta are semolina (or wheat flour), water, eggs and possibly fillings (meat, cheese, spinach, or herbs). A brief description to the recipe and nutritional table are reported respectively in Table 3 and Table 4:

Table 3
Recipe of "ravioli" with meat filling (4 persons)

Fresh pasta	350 g
Roast veal	150 g
Ham	30 g
Mortadella	30 g
Tomato sauce	400 g
Grated parmesan cheese	40 g
1 egg	
Nutmeg	
Salt	

Table 4
Nutritional table of "ravioli" with meat filling

Calories	290 kcal
Carbohydrates	34.5 g
Proteins	14.8 g
Fats	10.4 g

In order to perform CFD simulations, two main elements were considered: water vapour at 96°C and the final product . For the first one, properties are all known and can be easily retrieved from the software library. Conversely, for the food product, it was necessary to find (or compute) the correct properties.

In particular, the following parameters were set for the product : density, thermal conductivity and heat capacity. The density of fresh filled pasta is known, because it is provided by the manufacturing company (ρ = 456.36 kg/m^3); the thermal conductivity and heat capacity, instead, were derived from the available literature.

For the thermal conductivity, the following formulation was used (Saravacos & Maroulis, 2001):

$$k_d(x,T) = \frac{1}{1+x}\lambda_0 \exp\left[-\frac{E_0}{R}\left(\frac{1}{T}-\frac{1}{T_{rif}}\right)\right] + \frac{x}{1+x}\lambda_i \exp\left[-\frac{E_i}{R}\left(\frac{1}{T}-\frac{1}{T_{rif}}\right)\right] \quad (1)$$

where x is the water content of pasta, λ_i, λ_0, E_i, E_0 are used to describe the raw material (Table 5), T [°C] is the pasta temperature and T_{Rif} [°C] is a reference

temperature, which is set at 60°C (Saravacos & Maroulis, 2001).

Table 5
Parameters for pasta thermal conductivity calculation (Saravacos & Maroulis, 2001)

λ_i [W m^{-1}K^{-1}]	0.8
λ_0 [W m^{-1}K^{-1}]	0.273
E_i [kJ mol^{-1}]	2.7
E_0 [kJ mol^{-1}]	0.0

For simplicity, this model was used to obtain an average value of k_d in the range of temperature and water content used in the process. By setting an average pasta temperature of 349 K and an average relative humidity of 24% in Eq. 1, the following value of k_d was obtained:

$$k_d = 0.38 \left[\frac{W}{m \cdot K}\right] \quad (2)$$

Similarly, the heat capacity C_p was computed as weighed average of the heat capacities C_i of the dough main components, i.e. water, starch, gluten and fat (Andrieu, Gonnet & Laurent, 1998):

$$C_p = \frac{\partial h}{\partial T} = \left[\frac{X}{(1+X)} C_{p,water} + \frac{1}{(1+X)} C_{p,solids}\right] \quad (3)$$

where $C_{p,\,solids}$ was derived as:

$$C_{p,solids} = y_{starch} C_{p,starch} + y_{proteins} C_{p,proteins} + y_{fat} C_{p,fat} \quad (4)$$

The following parameters were set for the computation:

Water

$$C_w = 4184 \left[\frac{J}{kgK}\right] \quad (5)$$

Starch

$$C_s = 5.737 \cdot T + 1328 \left[\frac{J}{kgK}\right] \quad (6)$$

Gluten

$$C_g = 6.329 \cdot T + 1465 \left[\frac{J}{kgK}\right] \quad (7)$$

Fat

$$C_f = 2000 \left[\frac{J}{kgK}\right] \quad (8)$$

The pasta composition was obtained from the data proposed in Table 6.

Table 6
Fresh filled pasta composition (Source: data provided by a company manufacturing fresh pasta)

Water	36,09%
Starch	40,70%
Gluten	13,53%
Fat	9,68%

The global C_p is thus calculated as:

$$C_p = 2494.64 \left[\frac{J}{kgK}\right] \quad (9)$$

The above values of d, k_d and C_p for the ravioli have been used in all CFD simulations.

2.2 Mathematical modelling

2.2.1 Simulation setting: geometry and mesh

Given the complex geometry and the large size of the pasteurizer, the simulations were conducted considering 1/3 the length of the pasteurizer (whose total length is 2400 mm).

The width of the pasteurizer, instead, is kept at the original size, because the angle of inclination of the "roof" (cf. Figure 2) was designed specifically to ensure that the condensation of steam does not run directly on the ravioli. Condensation flows along the walls to be deposited in a special section on the bottom of the pasteurizer.

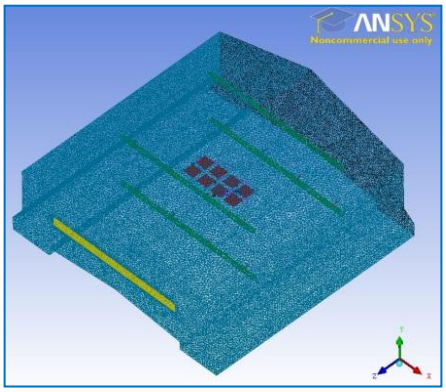


Figure 2. 1/3 of pasteurizer

The thermal analysis was performed on just two adjacent rows of product, each one containing 4 ravioli, although the industrial pasteurizer may contain a number of ravioli equal to 4/5 per 100 cm^2 of the conveyor belt. So, assuming an area of 3375 cm^2 (the size of the conveyor belt is 750x450 mm), the tape can contain up to 126 "ravioli".

In the present study, 8 "ravioli" are included in the model, to consider the impact between two adjacent ravioli and between two adjacent rows. Another reason for this choice is to not overload the mesh, which would compromise the reliability of the results.

The flow rates of steam were defined according to the load of dough in the pasteurizer. The size of the ravioli in the simulations complies with those used in subsequent experiments. The mesh created for this geometry is of tetrahedral type, with a gradient properly set to comply with the Courant number. This latter is of fundamental importance for transient flows. For a one-dimensional grid, it is defined as:

$$Courant = u\frac{\Delta t}{\Delta x} \tag{10}$$

where u is the fluid speed, Δt is the timestep and Δx is the mesh size.

The Courant number calculated in Ansys CFX is a multidimensional generalization of this expression where the velocity and length scale are based on the mass flow into the control volume and the dimension of the control volume (Löhner, 1987).

To allow a correct CFD simulation, the timestep must be chosen so that the Courant number is sufficiently small.

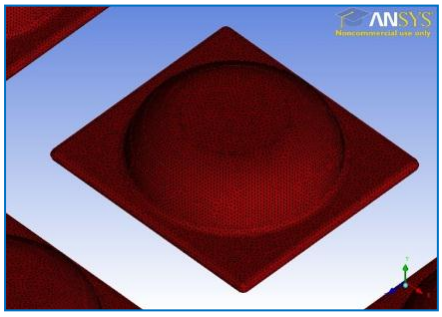


Figure 3. Surface Mesh on the "Raviolo"

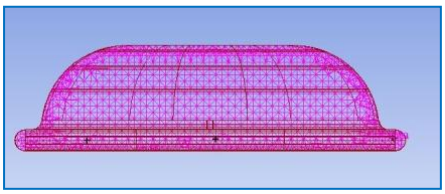


Figure 4. Volume mesh inside the "Raviolo"

2.2.2 Simulation setting: domain equations and boundary conditions

The software was used to solve the governing continuity, momentum and energy equations for the defined geometry and associated boundary conditions. The generalized transport equations solved are:

The continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho V) = 0 \tag{11}$$

The momentum equation:

$$\left(\frac{\partial \rho V}{\partial t} + \nabla \cdot (\rho V \otimes V)\right) = \nabla \cdot (-p\partial + \eta \cdot (\nabla V + (\nabla V)^t)) + S_M \tag{12}$$

In this work, according to the materials used, two domains were created: a fluid (steam) and a solid (filled pasta).

For the energy equation of the fluid domain, the pasteurization process of "filled fresh pasta" was modelled using a "Total Energy" approach, i.e.:

$$\frac{\partial(\rho h_{tot})}{\partial t} - \frac{\partial p}{\partial t} + \nabla \cdot (\rho U h_{tot}) = \nabla \cdot (\lambda \nabla T) + \nabla \cdot (U \cdot \tau) + U \cdot S_M + S_E \tag{13}$$

where h_{tot} is the total enthalpy, which can be expressed as a function of the static enthalpy h (T, p) as follows:

$$h_{tot} = h + \frac{1}{2}U^2 \tag{14}$$

The term $\nabla \cdot (U \cdot \tau)$ in eq. 13 represents the work due to viscous stresses and is known as the viscous work term. The term $U \cdot S_M$ represents the work due to external momentum sources. In this case, as in many applications, this term was neglected.

For the energy equation of the solid domain, CFX enables to create solid regions in which the heat transfer equations are solved, without considering the flow. Within solid domains, the conservation of energy equation can account for heat transport due to solid motion, conduction, and volumetric heat sources:

$$\frac{\partial(\rho h)}{\partial t} + \nabla \cdot (\rho U_S h) = \nabla \cdot (\lambda \nabla T) \tag{15}$$

where h, ρ and λ are the enthalpy, density and thermal conductivity of the solid, respectively; U_s is the solid velocity (set at 0).

The boundary conditions set above are related to 16 inlet holes and 1 output section; in particular, a uniform orthogonal velocity input and a relative pressure for outlet are set. The wall was considered as adiabatic.

3 SIMULATIONS RESULTS

The simulations are carried out in transition state, to evaluate the variation of the pasteurization temperature inside the "ravioli" in function of time, in which the product is in contact with water vapour at 96 °C.

Four series of simulations were performed by setting different values of "Total Time", while keeping the "timestep" unchanged (cf. table 7).

Table 7
Total time and Timestep of simulations

Total Time	*Timestep*
30 s	0.01 s
60 s	0.01 s
120 s	0.01 s
150 s	0.01 s

From all the simulations carried out, it was observed that the product heats up with the increase of the time. The figures below show the pasteurization phase at different time intervals .

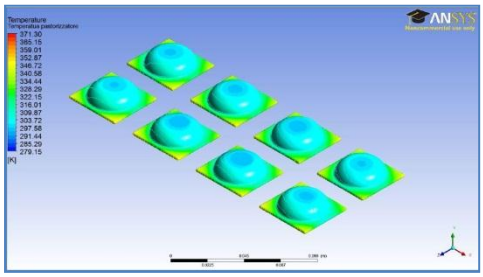


Figure 5. Simulation I - t = 30 s

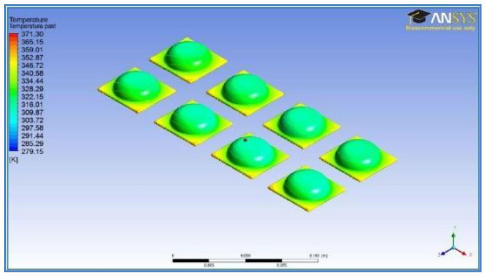


Figure 6. Simulation II - t = 60 s

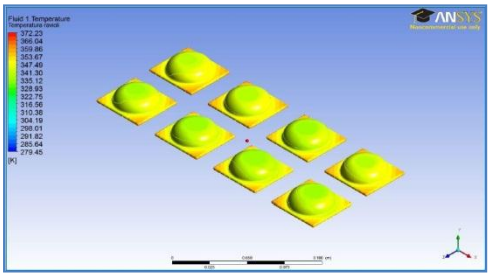


Figure 7. Simulation III - t = 120 s

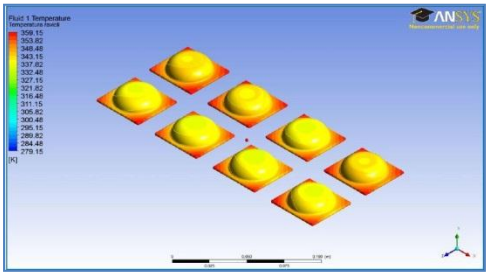


Figure 8. Simulation IV - t = 150 s

For each simulation, 8 points located at the core of each product were created. In these points, a chart to display the "pasteurization temperature ramp" was generated:

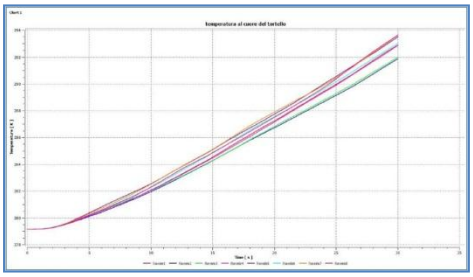


Figure 9. "Pasteurization temperature ramp" between 0 s < t < 30 s

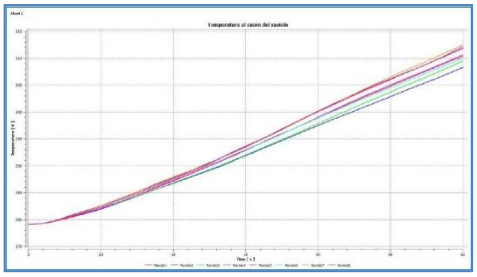


Figure 10. "Pasteurization temperature ramp" between 0 s < t < 60 s

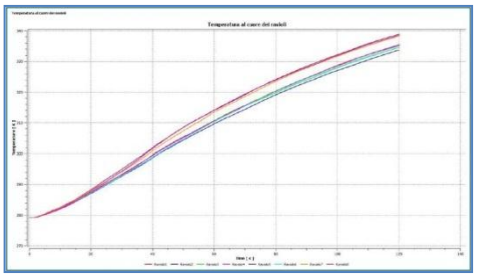


Figure 11. "Pasteurization temperature ramp" between 0 s < t < 120 s

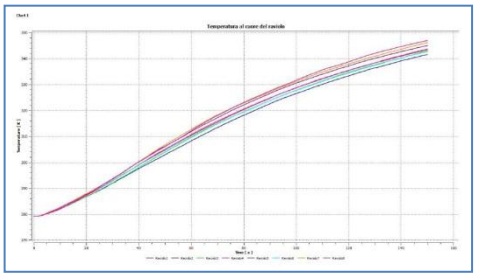


Figure 12. "Pasteurization temperature ramp" between 0 s < t < 150 s

From each simulation, with subsequent Microsoft Excel elaborations, it was possible to obtain all the temperatures calculated in the product volume. With those data, we were able to see the temperature trend in the product volume and not only in a specific point. From table 8 to table 11, the percentage of volume increase was observed. Values equal to 100% mean that the entire volume is located at a temperature higher or equal to the reference temperature set.

Table 8
% of volume for T greater or equal in respect of Tref (t = 30 s)

Tref	°C	K	%Volume
T0	15	288.15	100%
T1	25	298.15	85.06%
T2	35	308.15	42.58%
T3	45	318.15	22.84%
T4	55	328.15	9.19%
T5	65	338.15	1.86%
T6	75	348.15	0.04%

Table 9
% of volume for T greater or equal in respect of Tref (t = 60 s)

Tref	°C	K	%Volume
T0	15	288.15	100%
T1	25	298.15	100%
T2	35	308.15	99.99%
T3	45	318.15	71.38%
T4	55	328.15	33.09%

T5	65	338.15	12.98%
T6	75	348.15	1.75%

Table 10
% of volume for T greater or equal in respect of Tref (t = 120 s)

Tref	°C	K	%Volume
T0	15	288.15	100%
T1	25	298.15	100%
T2	35	308.15	100%
T3	45	318.15	100%
T4	55	328.15	100%
T5	65	338.15	87.03%
T6	75	348.15	22.92%

Table 11
% of volume for T greater or equal in respect of Tref (t = 150 s)

Tref	°C	K	%Volume
T0	15	288.15	100%
T1	25	298.15	100%
T2	35	308.15	100%
T3	45	318.15	100%
T4	55	328.15	100%
T5	65	338.15	100%
T6	75	348.15	56.24%

3.1 Summary of the simulations:

A summary of the tests conducted is reported below. In particular, table 12 shows the initial temperature (always constant), the temperature at the product core, the temperature at the base and the average temperature of the product, at four different time steps.

Table 12
summary of the simulations

	T_{start} [°C]	T_{inside} [°C]	T_{edge} [°C]	$\bar{T}$ [°C]
t = 30 s	6	20	61	41
t = 60 s	6	35	73	54
t = 120 s	6	67	82	75
t = 150 s	6	76	91	84

4 EXPERIMENTAL VALIDATION

4.1 Experimental tests

Experimental tests were performed to evaluate the accuracy and sensitivity of the results provided by the simulations described above. The tests were realized in a way as close as possible to the conditions implemented in the software. During the simulation, as mentioned, only 1/3 of the pasteurizer was considered, with a load of 8 ravioli disposed on two rows; furthermore, as done in the simulation phase, the ravioli are considered not in motion.

During the experimental tests, the same working conditions were set. The entire pasteurizer was used and 24 ravioli for test were loaded on the conveyor belt (i.e., a load three times higher). Moreover, on the basis of the travel time on the tape, it was possible to locate the ravioli in the same position as analysed during the simulation.

A temperature probe was inserted inside a product, with the purpose of recording the time-temperature trend. A series of tests to estimate an average trend of the temperature at the core of the product was performed.

The experimental results are reported in tables 13, 14 and 15:

Table 13
Experimental test 1

T_{start} = 7 °C	T_{inside} [°C]
t = 30 s	21.70
t = 60 s	47.10
t = 120 s	74.10
t = 150 s	82.10

Table 14
Experimental test 2

T_{start} = 6.20 °C	T_{inside} [°C]
t = 30 s	20
t = 60 s	41.80
t = 120 s	72.50
t = 150 s	83.50

Table 15
Experimental test 3

T_{start} = 6,5 °C	T_{inside} [°C]
t = 30 s	20
t = 60 s	38
t = 120 s	67
t = 150 s	77.50

4.2 Summary of the experimental tests:

As done for the simulation phase, a summary of the experimental tests performed is reported. In particular, figure 13 shows the "pasteurization temperature ramp" obtained in the experimental stage.

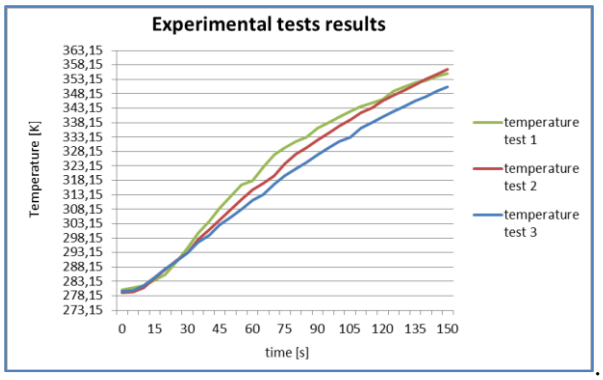


Figure 13. "Pasteurization temperature ramp" from the experimental tests

4.3 Comparison

Three series of experimental tests were performed and the average value of temperature at the beginning of the experiment (t = 0s) and at the four time intervals analysed during the simulation (t = 30, 60, 120, 150s) was calculated.

An average value was derived from the experimental tests , so as to have a single curve for comparison with the simulation trend (cf. figure 14).

Table 16 shows the average values of experimental tests and compares them with the results provided by the software. The following values refer to the ones measured and calculated at the product core.

Table 16
Comparison between software and experimental results

	Average of experiments [°C]	Simulations [°C]	ΔError [°C]
T_{start} [°C]	6.57	6	0.57
t = 30 s	20.57	20	0.57
t = 60 s	39.53	36	3.53
t = 120 s	70.83	67	3.83
t = 150 s	78.92	74.5	4.42

Figure 14 shows the two trends. From that figure, it is immediate to observe a slight deviation between the results calculated by the simulator and those reported during the tests. This can be due to the fact that the pasteurization temperature of the industrial process was not always constant; conversely, some fluctuations were recorded. The temperature set for the simulations, instead, is fixed (96°C).

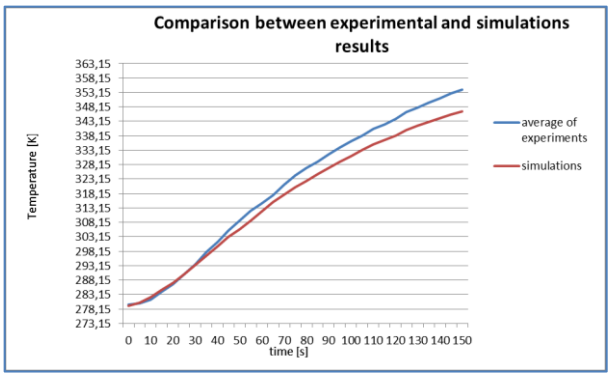


Figure 14. comparison between tests and software results

A further reason for deviation between experimental tests and simulations is that the real industrial scenario can be sensitive to many variables, since because each single product manufactured is different from another. For example, the consistency of the filling, the shape of the ravioli, the empty spaces inside them may vary. In addition, the initial temperature of the ravioli is never uniformly distributed and is not equal respect to another "ravioli".

5 CONCLUSIONS

The present work aims to analyse a thermal process for pasteurization of fresh filled pasta, by means of Computational Fluid Dynamic (CFD) simulation. The pasta considered is "ravioli" filled of meat.

A simulation model was built under Ansys CFX code version 14.0. All simulations have been performedin a transient state (after 30s, 60s, 120s and 150s), to evaluate the trend of the pasteurization temperature at the core of the "ravioli" as a function of the process time. The real industrial process takes place in a pasteurization tunnel by means of water vapour at approx. 96°C.

Experimental tests were performed in order to validate the simulation model of heat exchange. Results show a good agreement with the real pasteurization process and a good level of product quality, with a little underestimation of the inner temperature of the ravioli, due to the higher temperature reached by the vapour during the tests (approx. 99°C).

Analysing and optimizing a pasteurization process by means of experimental tests can be expensive; hence, the availability of a simulation model able to reproduce this process can be helpful in practice for its optimization.

Future researches will be oriented toward obtaining better results, setting an higher temperature in the treatment room and considering the diverse initial temperature of different zone of the product.

REFERENCES

De Cindio, Celot, Migliori, Pollini, 2001. A simple rheological model to predict filled fresh pasta failure during heat treatment. *Journal of Food Engineering, 48,* 7-18.

Löhner, R., 1987. An adaptive finite element scheme for transient problems in CFD. *Computer Methods in Applied Mechanics and Engineering 61 (3), pp. 323-338*

Migliori, Gabriele, de Cindio, Pollini, 2005. Modelling of high quality pasta drying: mathematical model and validation. *Journal Food of Engineering, 69,* 387 – 397.

Pagani, A. 2007. Influenza delle materie prime (sfarinati) sulla qualità della pasta fresca. *Atti Convegno AITA "Il processo di produzione della pasta alimentare fresca". Parma.*

Rizzo, Romagnoli, Vignali, 2010. Process parameter optimisation in the design of a pasteuriser for fresh-filled pasta. *Food Manufacturing Efficiency, volume 3, Issue 1, 25 - 33*

Zou, Opara, McKibbin, 2006. A CFD modelling system for airflow and heat transfer in ventilated packaging for fresh foods: I. Initial analysis and development of mathematical models. *Journal Food of Engineering, 77),* 1037 – 1047.

Zou, Opara, McKibbin, 2006. A CFD modelling system for airflow and heat transfer in ventilated packaging for fresh foods: II. Computational solution, software development, and model testing. *Journal Food of Engineering, 77,* 1048 – 1058.

AUTHORS BIOGRAPHY

Matteo FOLEZZANI is scholarship holder in Industrial Engineering at the Interdepartmental Center CIPACK. In 2011 he has achieved a master degree in Mechanical Engineering for the Food Industry, discussing a thesis titled: "Analysis and simulation of innovation system of sugar dissolution for the food industry". He attended the 2011 EFFoST Annual Meeting (Berlin 9-11 November 2011), with a poster

presentation titled "Analysis and simulation of a powders dissolution system based on the hydrodynamic controlled cavitation.". His main fields of research concern food process modeling and simulation, with a particular focus on the CFD simulation for the advanced design of food plants.

Eleonora BOTTANI is Lecturer (with tenure) in Mechanical industrial plants at the Department of Industrial Engineering of the University of Parma (Italy). She graduated in 2002 in Industrial Engineering and Management at the University of Parma, and got her Ph.D. in Industrial Engineering in 2006. Her research activities concern logistics and supply chain management issues, encompassing intermodal transportation, development of methodologies for supplier selection, analysis and optimization of supply chains, supply chain agility, supply chain modelling and performance analysis, and, recently, the impact of RFID technology on the optimization of logistics processes and supply chain dynamics. She is the author or co-author of more than 80 scientific papers, referee for more than 40 international scientific journals, editorial board member of 2 scientific journals and Associate Editor for one of those journals.

Gino FERRETTI is full professor of Mechanical Plants at the University of Parma (Italy). He graduated in Mechanical Engineering in 1974 at the University of Bologna, where he served as assistant professor for the courses of "Machines" and "Mechanical Plants". He worked as associate professor at the University of Padua and as full professor at the University of Trento. In 1988, he moved to the Faculty of Engineering, University of Parma, where at present he is full professor of Mechanical Plants. His research activities focuses on industrial plants, material handling systems, and food processing plants, and have been published in numerous journal and conference papers.

Michele MANFREDI is a PhD Student at the University of Parma, and Scholarship Holder in Industrial Engineering at the Interdepartmental Center CIPACK. He has achieved a master degree in mechanical engineering of food industry, discussing a thesis titled: "Analysis of the sterilization process of pouch packaging in an aseptic line". His main fields of research concern analysis LCA (Life Cycle Assessment) of food process and food products.

Giuseppe VIGNALI: graduated in 2004 in Mechanical Engineering at the University of Parma. In 2009, he received his PhD in Industrial Engineering at the same university, related to the analysis and optimization of food processes. Since August 2007, he works as a Lecturer at the Department of Industrial Engineering of the University of Parma, and, since the employment at the university, he has been teaching materials, technologies and equipment for food packaging to the food industry engineering class. His research activities concern food processing and packaging issues and safety/security of industrial plant. Results of his studies related to the above topics have been published in more than 20 scientific papers, some of which appear both in national and international journals, as well in national and international conferences. He acts also as a referee for some international journals, such as Prevention Today, Facilities.

Roberto MONTANARI is full professor of Mechanical Plants at the University of Parma (Italy). He graduated (with distinction) in 1999 in Mechanical Engineering at the University of Parma. His research activities mainly concern equipment maintenance, power plants, food plants, logistics, supply chain management, supply chain modelling and simulation, inventory management. He has published his research in approx. 60 papers, which appear in qualified international journals and conferences. He acts as a referee for several scientific journals and is editorial board member of 2 international scientific journals.

A LOT-SIZE SIMULATION MODEL WITH BATCH DEMAND WITH SPECIAL ATTENTION TOWARDS THE HOLDING COSTS

Gerrit K. Janssens[(a)], Roongrat Pisuchpen[(b)], Patrick Beullens[(c)]

[(a)]Operations Management and Logistics, Hasselt University
Diepenbeek, Belgium
[(b)]Department of Industrial Engineering, Kasetsart University
Chatuchak, Bangkok, Thailand
[(c)]School of Mathematics/School of Management, University of Southampton
Southampton, United Kingdom

[(a)]gerrit.janssens@uhasselt.be, [(b)]fengros@ku.ac.th, [(c)]P.Beullens@soton.ac.uk

ABSTRACT
The problem deals with the optimization of a multi-echelon supply chain with, at the downstream end, the final customer with random demand but with a pre-determined service level. In such a chain with several levels including production and distribution, safety levels appear for various types of products. Decisions on safety stock are made based on various costs, including the holding cost. It is shown how this holding cost could be calculated and whether it should be based on purchase prices or selling price. In a stochastic scenario it is not so clear what the consequences are of using the wrong type of price. A simulation in Arena has been constructed to show an example of such a supply chain, under various levels of uncertainty and various types of demand distributions.

Keywords: buyer-vendor system, multi-echelon supply chain, uncertain demand

1. BACKGROUND AND LITERATURE

The problem under study deals with the optimization of a multi-echelon supply chain. Such a supply chain consists of at least two levels, of which at the downstream end we have the final customer for which a service level has to be fulfilled. As demand from this final customer is random, safety stock needs to be provided, but as there are several levels with production and distribution, these safety levels appear for raw materials, intermediate products and finished products. The decision about service levels in between and the levels of safety stock is a decision matter of the supply chain, and is not seen by the final customer. He just wants goods delivered according to a pre-specified service level.

While the cost of holding inventory includes the opportunity cost of the money invested, expenses for running the warehouse, handling, insurance, losses for deterioration and damage, it is generally accepted that the largest portion of the holding cost is made up of the opportunity cost of capital (Silver et al. 1998). Thus in many traditional models the following convention is adopted for the holding cost per year:

$$\alpha v E(I) \quad (1)$$

where v is the unit variable cost to be invested for every unit placed in inventory, $E(I)$ is the average inventory in unit, and α is defined as the return on investment that could be earned on the next best alternative for the company.

For practical purposes, the question arises which cost elements and how to calculate from these the correct value for v. Furthermore it has to be looked at how to handle other variable out-of-pocket costs, in case they are considered important, like the cost of insurance or the rent of warehouse space. Starting from the important contribution by Grubbström (1980), and further development by Van der Laan and Teuntner (2002), this research develops some further analysis.

In Net Present Value (NPV) analysis, all cash flows, which are related to an activity, are valued by their time of occurrence using one common discount factor α. When applied to our practice, the NPV framework provides annuity stream (AS) profit functions for an inventory system. The NPV approach is powerful in deriving optimal inventory decisions in cases where the moments in time that cash flows occur are not based on the movements of product in the chain.

For certain classes of production and inventory problems, the difference between the classical approach and the NPV framework seems to be large, as shown in Grubbström and Thorstenson (1986) and in Teuntner and Van der Laan (2002). Why this difference appears is still a major issue to explain and to understand. Starting from Grubbström (1980 and further work by Haneveld and Teuntner (1998), it is shown that linearisations of the AS functions can be directly compared with the functions derived in the classical approach.

The linearization of the AS functions constructs a link between the NPV analysis and classical

frameworks but sometimes produces counterintuitive results. One example is the batch sales economic order quantity model, which this paper studies in further detail. The optimal lot size in this model is the basic EOQ result but the inventory is to be valued at sales price rather than at invested costs, as the classical

This contradiction is well-known, but it is shown in Beullens and Janssens (2011) that this outcome is not the only valid outcome. Their model introduces the concept of the anchor point, which allows to construct NPV models under either push and pull conditions. When the first activity has to start at some fixed, but arbitrary point of time in the future, the anchor point coincides with this fixed time in the future. In many classical production-inventory models optimal current decisions are restricted by the past.

The case which is studied in this paper is the elementary lot-size model with batch demand from Grubbström (1980) and Kim et al. (1984). A producer fulfills deterministic demand that occurs at a constant rate of y product units per year in batches of size Q. Stock-outs are not allowed and the producer has matched his production rate to the demand rate. The following additional information is required: the sales price (w), the discount rate (α), and the variable cost per product for producing at annual volume y ($c(y)$). The cash flows involved are: (1) a set-up cost s at the start of every cycle $T=Q/y$; (2) production cost equal to $c(y)y$ which is a continuous stream during the year; (3) income equal to wyT arises upon delivery of Q units at the end of every cycle. The annuity stream s can be calculated, from which the optimal order quantity and the holding cost can be obtained. .

The question seems to be whether the opportunity cost of capital should be made or at the rate what can be generated through sales of this inventory and that, by this, classical inventory theory is wrong. In Beullens and Janssens (2011) it is shown that the difference between results by the NPV approach and classical inventory theory depends on the choice of the location of the anchor point. Literature always has assumed that the start of the most upstream process in the supply chain is fixed. Larger production volumes delay the downstream activities and the final sales.

2. BASIS OF THE SIMULATION MODEL

In order to study the effect of using the wrong type of holding cost a simulation model is built making use of the Arena simulation software. A two-level supply chain is simulated in which customers have a demand to a retailer. This type of chain is called a buyer-vendor system (Goyal and Gupta 1989). It is a type of vertical integration in which the buyer and the vendor co-operate by synchronizing production with demand. The objective of this type of co-operation is joint profit maximisation. The dynamic of the inventory levels in this system is shown in Figure 1. The buyer puts orders in batches of size Q and faces a constant demand (y). The vendor works at a finite production rate $R \geq y$. In Figure 1, it is assumed that delivery time equals zero and no shortages exist. However, the buyer-vendor system in our simulation makes use of a stochastic demand and delivery time, in which shortages may appear.

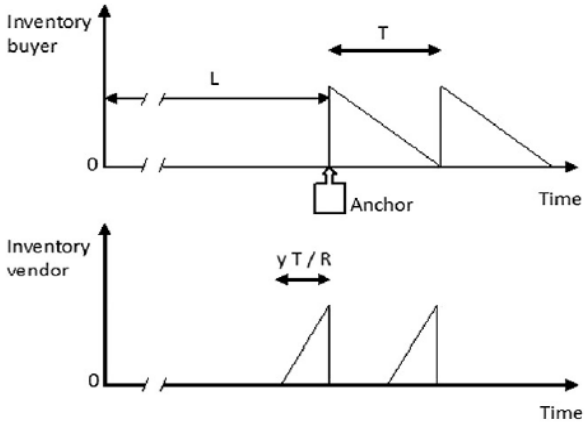


Figure 1: Buyer-vendor system with finite production rhythm

The simulation model makes use of the following data for the buyer-vendor system. The yearly demand y equals 3960 units, which means that demand during a simulation run of one year of 360 days equal 11 units per day. The vendor faces a fixed set-up cost per production (s_p) equal to 600 and a fixed shipment cost per shipment (s_s) equal to 25. The variable cost per product (c) equals 10. The buyer (retailer) faces a fixed cost per order (s_b) equal to 300 and pays a fixed price per unit (w) equal to 20. The sales price at the retailer's site (p) equals 40. The fixed ordering cost includes both the transportation cost and the administrative cost. The holding cost equals α = 20% of the purchase value of the product. The backorder cost is independent of time and is equal to $B_2 = 10$.

The retailer makes use of an (s, Q) inventory policy, in which s represents the re-order point and Q the fixed order quantity for placing his orders to the production site. The vendor aims to synchronise his production with the buyer's demand, i.e. R = 11 units per day. If delivery time is strictly positive, production should start before the start of the demand as the vendor wants to have the required order size ready for delivery.

In most buyer-vendor systems the vertical co-operation is realised by determining a fixed order quantity by means of the following formula:

$$Q^* = \sqrt{\frac{2*(s_b+s_p+s_s)*y}{\alpha(2p-c)\frac{y}{R}+\alpha c}} \qquad (2)$$

With the data mentioned before, formula (2) leads to an order size of 676.66 units. In the simulation we will use the closest multiple of daily demand which is 61 * 11 = 671 units. The total relevant cost (TRC) to be used in the simulation model is the cost function as defined by Banerjee (1986) and by Goyal (1988):

$$TRC = (s_p + s_b + s_s) * O + \alpha c * avg(invr) + \alpha w * avg(invm) + B_2 * ES \quad (3)$$

where

- *O:* number of orders/production runs per year
- *avg(invr)*: yearly average inventory at retailer's site
- *avg(invm)*: yearly average inventory at manufacturer's site.

3. EXPERIMENTS AND RESULTS

Table 1: Sensitivity analysis: scenario 1 – Normal distribution

Normal Distribution		**Variation Coefficient (σ_L/μ_L)**		
μ_L ↓	Element	**0,2**	**0,6**	**1**
2	σ_L	0.4	1.2	2
	Inventory Manuf.	334	333	333
	Inventory Retail	333	333	332
	# Shortages	18	32	49
	Orders Not satisfy	3	3	3,02
	Orders satisfy	2,52	2,56	2,54
	TRC	**€ 7.726**	**€ 7.866**	**€ 8.030**
	P_2	99,55%	99,19%	98,76%
6	σ_L	1,2	3,6	6
	Inventory Manuf.	337	337	336
	Inventory Retail	334	334	339
	# Shortages	33	88	115
	Orders Not satisfy	3,12	2,98	3,2
	Orders satisfy	2,6	2,7	2,5
	TRC	**€ 7.890**	**€ 8.440**	**€ 8.728**
	P_2	99,15%	97,79%	97,09%
10	σ_L	2	6	10
	Inventory Manuf.	339	339	339
	Inventory Retail	334	335	333
	# Shortages	48	139	223
	Order Not Satisfy	3,1	3,18	3,18
	Orders Satisfy	2,76	2,66	2,66
	TRC	**€ 8.044**	**€ 8.958**	**€ 9.790**
	P_2	98,77%	96,47%	94,35%

The simulation model is run for a scenario without safety stock and a scenario with safety stock.

The case of no safety stock is simulated under the following conditions: the demand follows a Poisson distribution and the lead time L follows a Normal distribution, with various parameter values of the mean value μ_L and its standard deviation σ_L. Table 1 show some simulation results for various parameter values. The table shows the average yearly inventory at the manufacturer's site (Inventory Manuf.), the average yearly inventory at the retailer's site (Inventory Retail), the number of shortages on a yearly basis (# Shortages), the number of times per year that a manufacturer cannot deliver the full order quantity (Orders Not satisfy), the number of times per year that a manufacturer can deliver the full order quantity (Orders satisfy), total relevant cost (TRC), the P_2-service level (also called fill rate) (P_2).

Different combinations of the parameter values on the lead time distribution do not lead to big changes into average inventory levels both at the retailer's and at the manufacturer's site.

It can be expected that the number of units short (# shortages) increases with an increasing level of lead time variability (as there is no safety stock). This increase in the numer of units short leads to a lower service level (P_2) and to an increase of the shortage cost part in the total relevant cost (TRC). The decrease in service level is more explicit when the lead time becomes bigger.

In a second scenario it is assumed that th retailer hold a level of safety stock. The level depends on the parameters of the distribution of demand during lead time. Via the *Input Analyzer* (Arena software) the distribution during lead time is determined in an empirical way. The best fitting distribution is used to determine the re-order corresponding to a pre-specified service level. Figure 2 shows such an empirical distribution and also the best fitting distribution (based on the Mean Square Error criterion), which in this case is the Normal distribution (for μ_L = 22.2 and σ_T = 7.99*)*.

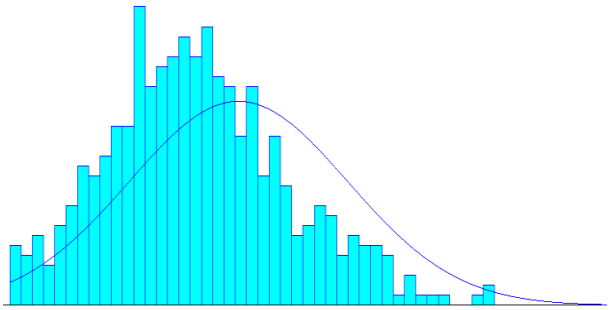


Figure 2: Distribution of demand during lead time (μ_L = 22.2 and σ_T = 7.99*)*.

Table 2: Sensitivity analysis: scenario 2 – Normal distribution

Normal Distribution		Variation Coefficient (σ_L/μ_L)		
μL ↓	Element	**0,2**	**0,6**	**1**
2	σ_L	0.4	1.2	2
	Inventory Manuf.	332	335	334
	Inventory Retail	344	351	359
	# Shortages Orders	2	3	6
	Not satisfy Orders	3,38	3,32	3,4
	Satisfy	2,26	2,26	2,22
	TRC	**€ 7.610**	**€ 7.654**	**€ 7.714**
	P_2	99,95%	90,92%	99,85%
6	σ_L	1,2	3,6	6
	Inventory Manuf.	337	337	339
	Inventory Retail	353	401	415
	# Shortages Orders	4	8	12
	Not satisfy Orders	3,44	3,48	3,4
	Satisfy	2,32	2,5	2,56
	TRC	**€ 7.676**	**€ 7.908**	**€ 8.008**
	P_2	99,90%	99,82%	99,65%
10	σ_L	2	6	10
	Inventory Manuf.	337	337	339
	Inventory Retail	365	423	440
	# Shortages Orders	5	20	36
	Not satisfy Orders	3,32	3,42	3,44
	satisfy	2,64	2,62	2,58
	TRC	**€ 7.734**	**€ 8.116**	**€ 8.348**
	P_2	99,87%	99,54%	99,10%

From Table 2 it can be learned that the total relevant cost (TRC) show a smaller increase in case a safety stock is used. The service level is always above the 99% level. As could be expected, the level of safety stock is higher with a higher level of lead time variability. The increase in the total relevant cost (TRC) is mainly due to the increase in inventory cost. When the lead time is small (μ_L= 2), the increase in inventory level is 4.36%, but when it comes to an intermediate level of lead time (μ_L= 6), the increase is more than 17%. For a high level of lead time it is even more than 20%.

4. CONCLUSIONS

A simulation model has proven to be of high value in order to investigate the effect of lead time variability in a buyer-vendor co-operation. While both partners are seeking to maximise their joint profit, they should also worry about variability in delivery as it lay destroy their joint positive ideas. They should use the arguments brought forward by the simulation to negotiate with their third party logistics providers.

REFERENCES

Banerjee, A., 1986. On 'a quantity discount pricing model to increase vendor profits'. *Management Science*, 32(11), 1513-1517.

Beullens, P. and Janssens, G.K., 2011. Holding costs under push and pull conditions – the impact of the anchor point. *European Journal of Operational Research*, 215, 115-125.

Goyal, S.K., 1988, A joint economic lot-size model for purchaser and vendor: a comment. *Decision Sciences*, 19, 236-241.

Goyal, S.K. and Gupta Y.P., 1989. Integrated inventory models: the buyer-vendor coordination. *European Journal of Operational Research*, 41(3), 261-269.

Grubbström, R.W., 1980. A principle for determining the correct capital costs of work-in-progress and inventory. *International Journal of Production Research*, 18(2), 259-271.

Grubbström, R.W. and Thorstenson, A., 1986. Evaluation of capital costs in a multi-level inventory system by means of the annuity stream principle. *European Journal of Operational Research*, 24, 136-145.

Haneveld, W.K.K. and Teuntner, R.H., 1998. Effect of discounting and demand rate variability on the EOQ. *International Journal of Production Economics*, 54, 173-192.

Kim, Y.H., Chung, K.H. and Wood, W.R., 1984. A net present value framework for inventory analysis. *International Journal of Physical Distribution and Logistics Management*, 14(6), 68-76.

Silver, E.A., Pyke, D.F. and Peterson, R., 1998, *Inventory Management and Production Planning and Scheduling*, 3rd ed., J. Wiley & Sons, New York.

Teuntner, R. and van der Laan, E., 2002. On the non-optimality of the average cost approach for inventory models with remanufacturing. *International Journal of Production Economics*, 79, 67-73.

Van der Laan, E., and Teuntner R., 2002. Average cost versus net present value: a comparison for Multi-source inventory models. In: Klose, A., Speranza, M.G., Van Wassenhove, L.N. (eds.) *Quantitative Approaches to Distribution Logistics and Supply Chain Management*, Springer-Verlag, 359-378.

AUTHORS BIOGRAPHY

Gerrit K. JANSSENS received his Ph.D. in Computer Science from the Free University of Brussels (VUB), Belgium. Currently he is Professor of Operations Management and Logistics at Hasselt University (UHasselt) within the Faculty of Business Administration. He has been president of the Belgian Operations Research Society (ORBEL) in 2006-2007. He is president of the board of Eurosis (the European Multidisciplinary Society for Modelling and Simulation Technology). His main research interests include the development and application of operations research models in production and distribution logistics.

Roongrat PISUCHPEN received her Bachelor of Engineering in Industrial Engineering from Kasetsart University (Thailand), Master of Engineering in Industrial and Systems Engineering and PhD in Industrial and Systems Engineering from Asian Institute of Technology (Thailand). Currently, she is an Associate Professor in Industrial Engineering at Engineering Faculty, Kasetsart University in Bangkok. She is author of *The Handbook of Simulation Modelling with ARENA* published in Thai. Her main research interests include Petri Net models, simulation modelling and analysis, and applications of operations research.

Patrick BEULLENS obtained his PhD in Industrial Management at the Catholic University Leuven (Belgium) in 2001 on location, process selection and vehicle routing models for reverse logistics. He was a visiting research associate at INSEAD (Fontainebleau, France) and a post-doctoral researcher at the Erasmus University Rotterdam (the Netherlands, 2002). From 2004 till 2011 he was at the Department of Mathematics of the University of Portsmouth (UK). He was a guest professor at the Faculty of Economics of the University of Hasselt (Belgium, 2009-2011). Currently he is a reader/senior lecturer at the School of Mathematics and School of Management at the University of Southampton, where he lectures in the area of OR/MS.

GRID GENERATION FROM VIDEO CAPTURE FOR MESHLESS METHOD THERMAL SIMULATIONS

Khaoula Lassoued [(a)], Tonino Sophy[(a)], Luis Le Moyne[(a)], Nesrine Zoghlami[(b)]

[(a)] DRIVE – ISAT EA 1859, 49, rue Mademoiselle Bourgeois, 58000 NEVERS, FRANCE.
[(b)] CONPRI – 99/UR/11-29, Rue Omar Ibn-Elkhattab, 6029 GABES, TUNISIA

[(a)] khaoulalassoued@gmail.com, [(b)] nesrine.zoghlami@gmail.com

ABSTRACT
A dynamic grid generation tool based on contour reconstruction and a Meshless Diffuse Approximation Method (DAM) is developed. The purpose is extracting the points cloud involved in the Meshless simulation from a digitalized picture or a video capture. Each frame taken from an ultra-speed camera video is transformed into a set of points.
After several image treatment investigations, a threshold-Hough association method is adopted and tested with circular and more complex shaped object.
The obtained grids are then used in numerical simulations with the Diffuse Approximation Meshless method. Transient heat diffusion and steady state convection are simulated. Results are presented as isotherms and streamlines.
The presented method seems to create DAM compatible grids as the isotherms are in respect with the conduction phenomenon and the streamlines fits to the references corresponding to similar cases.

Keywords: heat transfer, fluid flow, images and video processing, circular and complex shape object.

1. INTRODUCTION

Numerical simulation is widely used in advanced technology studies, especially in engineering field. Indeed, it is necessary to use adapted numerical method to produce convincing simulations of physical event such as fluid flow (smoke or water), thermal diffusion and natural or forced convection.
The choice of numerical methods to solve and represent the describing equations depends ideally on application fields.
In the past, to solve systems governed by Partial Differential Equations (PDEs), Finite Difference Method (FDM) Boutayeb (1991) and the Finite Element Method (FEM) Long (1995) which reduce the computational problem of complex geometries, were widely used. In spite of the great success of the so-called method as effective numerical tools for the solution of boundary values problems on complex domains, there has been, over the past decades, a growing interest in numerical methods which not requires finite element meshes. This class of method is known as “meshfree” or “particle” method. Indeed, the main advantage of the meshless methods is the needlessness of any predefined mesh or elements between the nodes. It just requires a grid of points for the discretization. So far, a number of methods have been proposed, the first one beeing the Smooth Particle Hydrodynamics (SPH) Monagahn (1992). For example, Ahmadi *et al*. (2010) used a Moving Least Square approximation method (MLS) to simulate steady-state heat conduction in heterogeneous materials.
Even if particle methods do not need meshing of the domain, the spatial discretization requires a set of points. This can be done by mathematical functions, grid extracting from mesh generation software or more recently by image processing.
Furthermore, in image processing scientific field, many works have been done concerning object detection or reconstruction. Main applications concern medical field, automotive comfort or security, face detection (Viola 2001, Li 2002 and Sochman 2004) or national defense department.
So, the focus of this paper is the association of image processing advance for the automation of static or dynamic grid construction from an image or video capture and simulation by a meshless method.
In this work, a Diffuse Approximation Method (DAM) based on a moving weighted least square approximation Sophy (2002) is developed to simulate heat diffusion around free falling ball and to simulate a thermal fluid flow around different objects. This method was first introduced by Nayroles *et al*. in the beginning of the 90s Nayroles (1991).
In this contribution, we start by providing a general description of DAM. We then present several methods of image processing used to generate the points’ grid from the image or video capture. Then, heat conduction numerical results are presented in terms of isotherms or temperature field and some results of natural or forced convection around a circular or a more complex shaped object are respectively shown.

2. METHOD DESCRIPTION

2.1. Meshless Diffuse Approximation Method

Let φ: $R^n \rightarrow R$ be a scalar field whose values φ_j are known at the points x_j of a given set of n nodes in the

studied domain D ∈ R^n. The diffuse approximation gives estimates of φ and its derivates up to an order k at any point M(x,y) ∈ D. The order 2 Taylor expansion of φ at M gives:

$$\varphi_j = \varphi + \frac{\partial \varphi}{\partial x}(x_j - x) + \frac{\partial \varphi}{\partial y}(y_j - y) + \frac{\partial^2 \varphi}{2!\partial x^2}(x_j - x)^2 + \frac{\partial^2 \varphi}{\partial x \partial y}(x_j - x)(y_j - y) + \frac{\partial^2 \varphi}{2!\partial y^2}(y_j - y)^2 \quad (1)$$

For reason of simplification, let assume the following notation:

$$(x_j - x) \leftrightarrow x_j \quad (y_j - y) \leftrightarrow y_j$$

With the minimization of the quadratic error one can obtain expressions of φ and its derivatives at any desirable nodal point:

$$\begin{Bmatrix} \varphi \\ \frac{\partial \varphi}{\partial x} \\ \frac{\partial \varphi}{\partial y} \\ \frac{1}{2!}\frac{\partial^2 \varphi}{\partial x^2} \\ \frac{\partial^2 \varphi}{\partial x \partial y} \\ \frac{1}{2!}\frac{\partial^2 \varphi}{\partial y^2} \end{Bmatrix} = [A^M]^{-1} \cdot \sum_{M_j \in v^M} \omega(M_j, M) \begin{Bmatrix} 1 \\ x_j \\ y_j \\ x_j^2 \\ x_j y_j \\ y_j^2 \end{Bmatrix} \cdot \varphi_j \quad (2)$$

where the matrix $[A^M]$ is:

$$[A^M] = \sum_{M_j \in v^M} \omega(M_j, M) \cdot \begin{bmatrix} 1 & x_j & y_j & x_j^2 & x_j \cdot y_j & y_j^2 \\ x_j & x_j^2 & x_j \cdot y_j & x_j^3 & x_j^2 \cdot y_j & x_j \cdot y_j^2 \\ y_j & x_j \cdot y_j & y_j^2 & x_j^2 \cdot y_j & x_j \cdot y_j^2 & y_j^3 \\ x_j^2 & x_j^3 & x_j^2 \cdot y_j & x_j^4 & x_j^3 \cdot y_j & x_j^2 \cdot y_j^2 \\ x_j \cdot y_j & x_j^2 \cdot y_j & x_j \cdot y_j^2 & x_j^3 \cdot y_j & x_j^2 \cdot y_j^2 & x_j \cdot y_j^3 \\ y_j^2 & x_j \cdot y_j^2 & y_j^3 & x_j^2 \cdot y_j^2 & x_j \cdot y_j^3 & y_j^4 \end{bmatrix} \quad (3)$$

and ω is a weight-function of compact support, equal to unity at this nodal point, decreasing when the distance to the node increases and equal to zero outside a surrounding zone (mentioned as v^M) near the calculation node. In our study, we chose the following Gaussian window:

$$\begin{cases} \omega(X, X_j - X) = \exp\left[-3\ln(10).\left(\frac{|X_j - X|}{\sigma}\right)^2\right] \\ \omega(X, X_j - X) = 0 \quad if\ ((X_j - X)^2 \geq \sigma^2) \end{cases} \quad (4)$$

where σ is the radius of the weight function support.
Thus, any Partial Derivative Equation can be written in terms of different φ_j. This requires information on the position of each point of the grid which describes the domain.

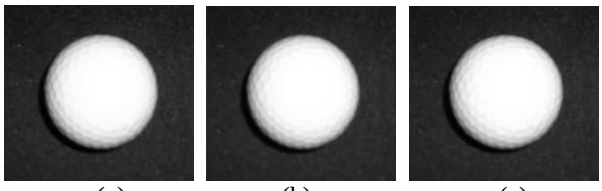
As the ultimate objective of this contribution is the association of DAM with an automatic procedure to obtain the set of points, a description of different stage of our image processing is given in next section.

2.2. Image processing

In this section we show different used methods for image treatment. A brief description will be given for these methods. Our code development is based on a C++ language using an Open source Computer Vision library (OpenCV). To achieve the grid generation, the captured image has to be smoothed by filtering operations. This is necessary for edge detection then pixel classification.

2.2.1. Image preprocessing

When an image is acquired by a camera or other imaging system, the vision system for which it is intended is often unable to use it directly. Good image smoothing should be able to deal with different types of noise. In this paper, two image smoothing filters are used. Figure 1 shows an example of original image and the results obtained with a Gaussian linear filter and a Median non-linear filter which is very effective in removing salt and pepper and impulse noise while retaining image details.



(a) (b) (c)

Figure 1: Smoothing filters, Original image (a), Gaussian filter (b) and Median filter (c).

2.2.2. Edge detection

As seen before DAM also needs the type of any point (if it is a point belonging to an edge, an object or a background point). It is then necessary to detect contours in the image. Edge detection is a fundamental tool used in most image processing applications to obtain information from the frames as a precursor step to feature extraction and object segmentation. This process not only detects boundaries between objects and the background in the image, but also the outlines within the object. To detect edges, many operators such as Sobel operator or Canny detector can be applied. The OpenCV library gives an important function that can detect contours. This function is called cvFindContours Bradski (2008). During our work many detection operators like Sobel or Canny, FindContour function and the Hough operator has been used (Figure 2 and 3).

- Sobel Operator:

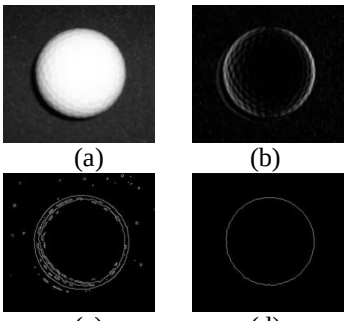
It is a discrete differentiation operator, computing an approximation of the opposite of the gradient of the image intensity function.

- Canny Operator:

Canny's aim is to discover the optimal edge detection algorithm to satisfy good detection, good localization and minimal response

- cvFindContours:

The function cvFindContours retrieves contours from the binary image and returns the number of retrieved contours. The pointer "firstContour" is filled by the function. It will contain pointer to the first most outer contour or NULL if no contours is detected (if the image is completely black). Other contours may be reached from firstContour.

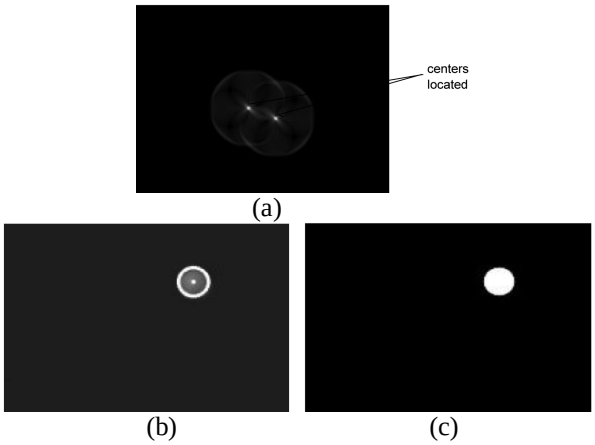


(a) (b) (c) (d)

Figure 2: Edge detecting filters, Original image (a), Sobel detector (b), Canny detector (c), Find Contour (d)

- Hough transform:

The Hough transform is a feature extraction technique used in image analysis, computer vision and digital image processing Shapiro (2001). In our work, we used Hough to reconstruct the circles contours Kimme (1975). The Hough circle reconstruction technique highlights in the image the potential centers of r radius circles (figure 3-a). The center being detected, one can reconstruct the circle's contour (figure 3-b) or the entire object (figure 3-c).



(a) (b) (c)

Figure 3 Hough circle transform, Center detection (a), (Ic) Contour image (b), (Io) Object image (c).

2.2.3. Acquisition

Two types of acquisition are made. Stand images are captured with a 640x480 resolution digital camera. No filters are used during the acquisition as the image is treated after. The video captures are made with a Photron FASTCAM Ultima APX-RS (Figure 4) high-speed video camera that can reach 250 000 frames per second. During our work we capture 500 frames per second. This implies a dynamic system. Using Hough Transform (section 2.2.2) we treat this video frame per frame. Sometimes happen that the Hough detection is not suitable. This leads to the temporary disappearance of the object. A particular treatment is then applied. Our capture concerns a free falling ping pong ball. A black background is used to avoid additional treatment of the vicinity. High power spotlights are set at both sides of the scene to avoid a privileged light exposition side that can drive to spurious shadow (which can be interpreted as a contour). All this wariness is applied to reduce the future image processing time. Indeed, stand images are captured without black background and spotlights so the spurious contours are treated by the process.

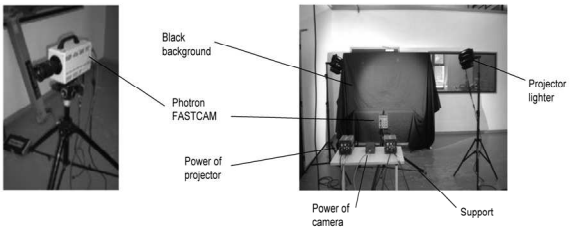


Figure 4: Photron FASTACAM Ultima APX-RS and material used.

An example of frame and the corresponding grid are shown in Figure 5. For the stand image, the picture is treated as one frame of the video.

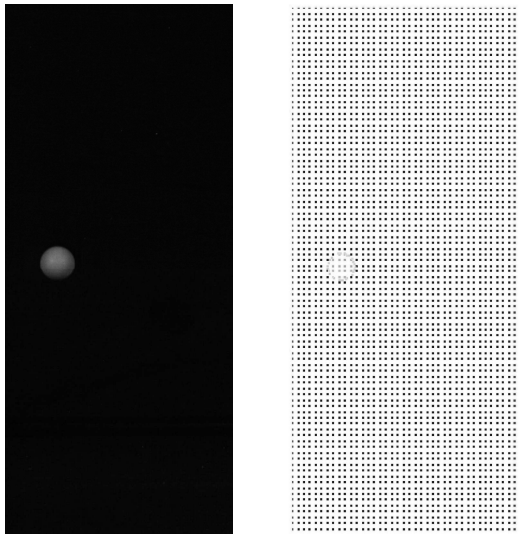


Figure 5: Grid point generated from frame

3. RESULTS

In this section some results of simulations obtained with DAM associated with the presented grid generation technique are shown. The first case concerns transient heat diffusion of a free falling circular object. The isotherms are presented for two different times and two different thermal conductivities. The second and the third cases, involve convective exchanges of a circular and a more complex shaped object. Streamlines and isotherms are calculated.

3.1. Heat diffusion

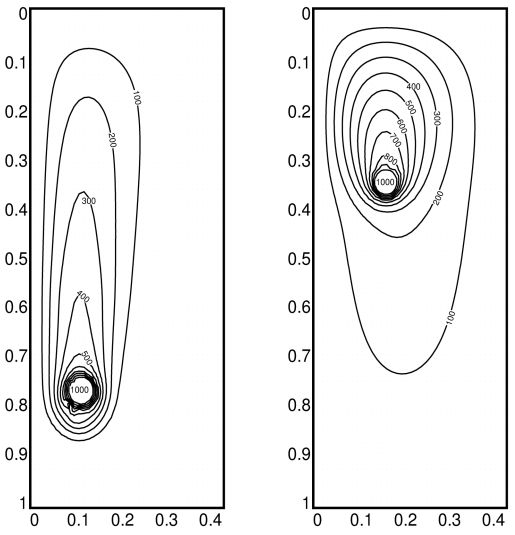
The dynamic grid is obtained with a video capture of a free falling rebounding ping-pong ball (Figure 5). The diffusion term of the transient heat equation (Eq 4) is discretized with the DAM and the transient term is approximated with a forward first order Finite Difference scheme:

$$-\frac{\rho C}{\lambda}\frac{\partial T}{\partial t}+\Delta T=0 \qquad (4)$$

where, ρ is the density, λ is the thermal conductivity, C is the specific heat capacity, T is the temperature and t is the time.

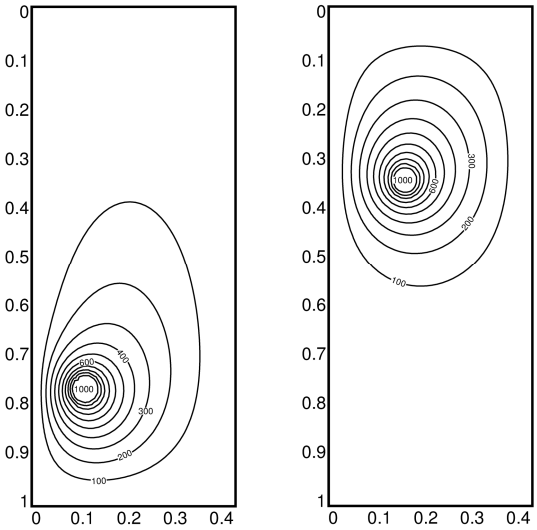
The boundary conditions are fixed temperatures on the limits of the domain (T_c=0) and the object (T_h=1000). All temperatures are initially set to $T_{init}=T_c$. The thermal conductivity values are λ = 1 W/mK and $\lambda = 8.10^{-3}$ W/mK while other physic properties are fixed to the value 1 SI. The time step Δt is fixed to 10^{-3}s.

The presented isotherms for times corresponding to 184 Δt (after the first rebound) and 404 Δt (during the second descent) (Figure 7) are in total accordance with the physical phenomenon. The increment of the thermal conductivity (Figure 8) intensifies the diffusive character of the problem and reduces the thermal inertia.



(a) (b)

Figure 7: Isotherms form for λ =8 10^{-3}, Ts=184Δt (a) and Ts=404Δt (b)



(a) (b)

Figure 8: Isotherms form for λ =1, Ts=184Δt (a) and Ts=404Δt (b)

3.2. Flow over a circular shape object

In this case the steady state buoyancy flow around a circular shape object is simulated.

The governing equations are the Navier-Stokes equations in their dimensionless secondary variable form and the energy equation (Eq 5-7). As a pseudo-instationnary algorithm is used, the equations are in their transient form:

$$\frac{\partial^2\Psi}{\partial x^2}+\frac{\partial^2\Psi}{\partial y^2}=-\Omega \qquad (5)$$

$$\frac{\partial\Omega}{\partial t}+u\frac{\partial\Omega}{\partial x}+v\frac{\partial\Omega}{\partial y}=Pr\left(\frac{\partial^2\Omega}{\partial x^2}+\frac{\partial^2\Omega}{\partial y^2}\right)+Ra\,Pr\frac{\partial T}{\partial x} \qquad (6)$$

$$\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}=\nabla^2 T \qquad (7)$$

where Ψ represents the dimensionless stream function, Ω the vorticity, T the temperature, Ra the Rayleigh number and Pr the Prandtl number which are defined as:

$$Ra=\frac{g\beta\Delta TL_{ref}^3}{\nu\alpha}. \qquad (8)$$

$$Pr=\frac{\nu}{\alpha} \qquad (9)$$

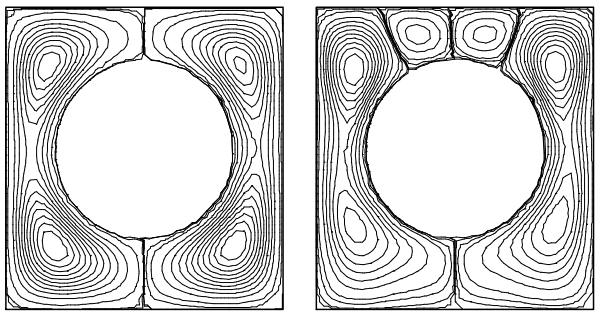
whith g, β, ΔT, L_{ref}, ν and α being respectively the gravity, the thermal expansion coefficient, a characteristic temperature difference (T_h-T_c), a characteristic length (diameter of the circular object), cinematic viscosity and the thermal diffusivity.

The velocity component (u , v) are obtained with the stream function according to the equations:

$$u=\frac{\partial\Psi}{\partial y}\quad v=-\frac{\partial\Psi}{\partial x}. \qquad (10)$$

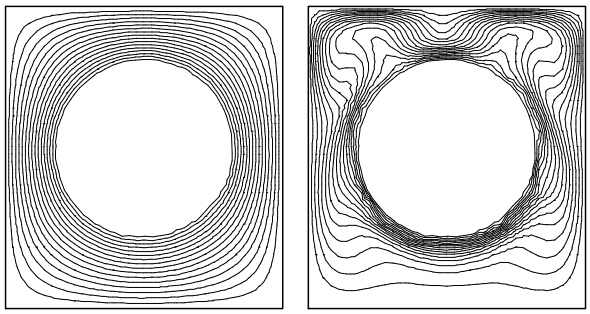
The fluid is assumed to be air (P=0.7). All the physical properties are constant except the density where the Boussinesq approximation is applied. The boundary conditions are fixed temperatures and stream functions

on the object (T_h, Ψ_{object}) and the walls (T_c, Ψ_{wall}). Non slip boundary conditions are applied (null velocities on the walls and the object).
Calculations are made for several Rayleigh number ranking from 1000 to 10^5. Figure 9 shows examples of streamlines obtained for Ra=1000 and Ra=10^5. Figure 10 shows temperature field for the same Ra numbers.



(a) (b)

Figure 9: Streamlines for natural convection of Circular shaped object, Ra = 1000 (a) and Ra= 10^5 (b).



(a) (b)

Figure 10: Isotherms for natural convection of Circular shaped object, Ra = 1000 (a) and Ra= 10^5 (b).

The results are in accordance with the references for similar problems as the structures of the flows respect the natural convection flows.

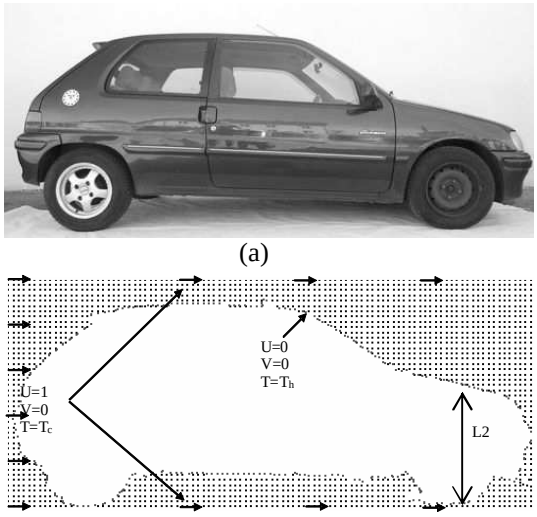
3.3. Flow over a car shape object

As a main advantage of meshless methods is the flexibility of the grid to deal with complex shapes, a car image is taken as original picture to built the grid. Nevertheless, the hardiness of the point cloud generation increases with the complexity of the involved shapes. The proposed method allows to generate complex grid from a picture (Figure 11). Then the DAM can be used to simulate convection exchanges on the obtained discretized domain.
For the natural convection the equations still the same as those used in section 3.2, with L_{ref} being the length (along the direction x) of the calculation domain.
For the forced convection the (Eq 6) and (Eq 7) respectively become:

$$\frac{\partial \Omega}{\partial t} + u\frac{\partial \Omega}{\partial x} + v\frac{\partial \Omega}{\partial y} = \frac{1}{Re}\left(\frac{\partial^2 \Omega}{\partial x^2} + \frac{\partial^2 \Omega}{\partial y^2}\right) \qquad (11)$$

$$\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} = \frac{1}{Pr\,Re}\nabla^2 T \qquad (12)$$

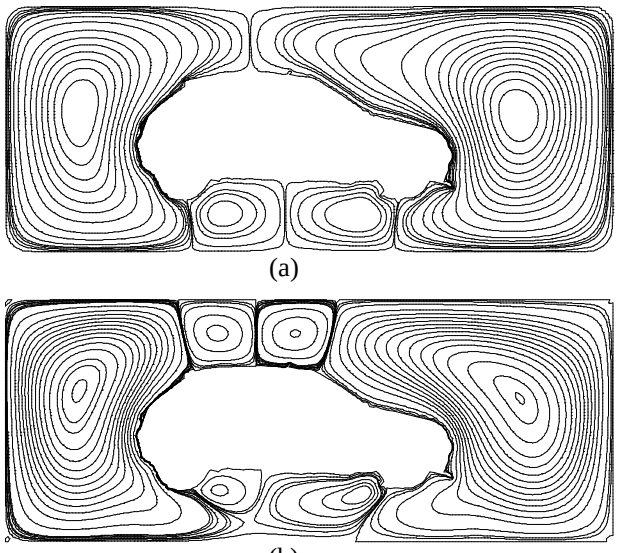
and L_{ref} is fixed as the length L_2 shown in Figure 11b.



(a)

(b)

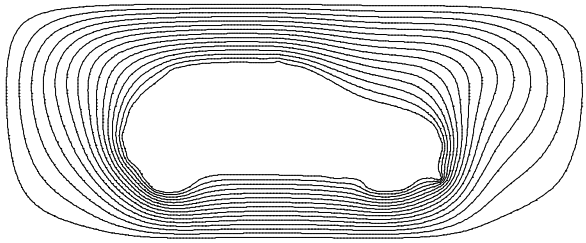
Figure 11: Complex shaped object (a), Automatic generated grid and Boundary Conditions (b)



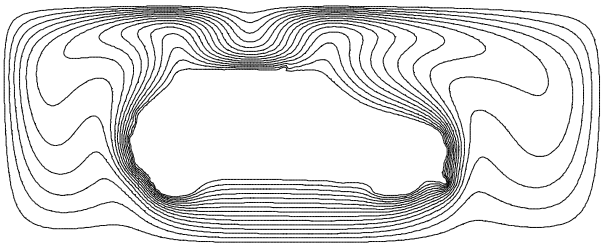
(a)

(b)

Figure 12: Streamlines for natural convection of car shape object, Ra = 1000 (a) and Ra= 2.10^4 (b).

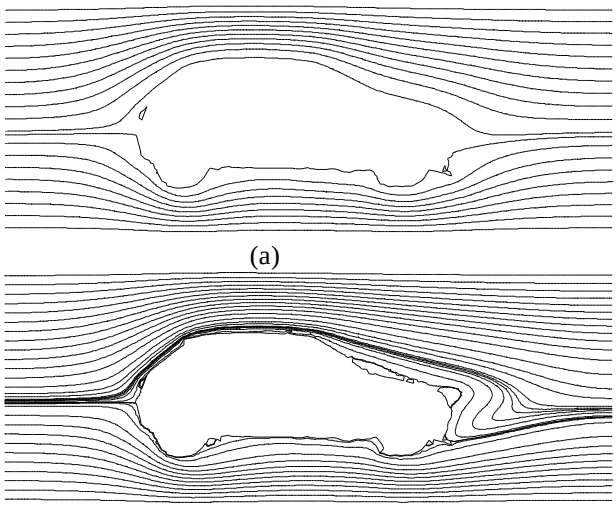


(a)



(b)

Figure 13: Isotherms for natural convection of car shape object, Ra = 1000 (a) and Ra= 2.10^4 (b).

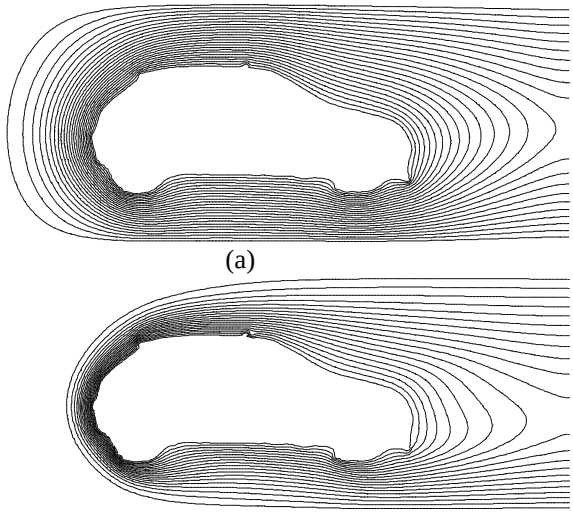


(a)

(b)

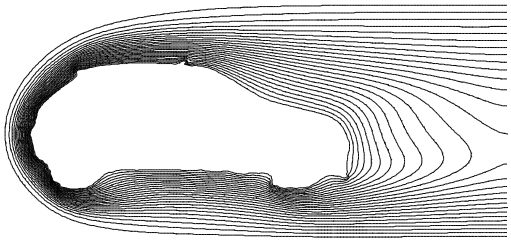
(c)

Figure 14: Streamlines for forced convection of car shape object, Re = 30 (a) Re=180 (b) and Re=210.



(a)

(b)



(c)

Figure 15: Isotherms for forced convection of car shape object, Re = 30 (a), Re= 180 (b) and Re=210 (c).

Streamlines and temperature fields are presented for natural and forced convection respectively on (Figure (12,13)) and (Figure (14,15)). Concerning natural convection, one can see that for low Rayleigh numbers (Figure (12a,13a) isotherms correspond to a diffusion regime and when Ra increases a plume flow appears. When simulating forced convection, a similar observation is made in respect of the Reynolds number. Simulations are made with Re=30, 180 and 270. Boundary conditions are in accordance with Figure 11b. The flow inlet is on the rear side of the car. This corresponds to a reverse car travelling simulation.

The ground is not taken into account so the flow can get around the tires. Indeed a 2D simulation can not simulate the flow within the car track. The simulation of the ground would lead to fluid accumulation near the tires.

For low Re numbers, streamlines are horizontal and incurve near the object. The increment of Re involves appearance of localized vortexes. These recirculations are representative of a flow detachment. They take place at the junction between the windshield and the hood, and between the hood and the headlights. The presence of these recirculations is probably relevant to the care brought to these locations in car conception. Figure 14c shows that another increment of Re leads to a merge of the eddies.

4. CONCLUSION

A meshless method grid generation procedure has been developed. It is the first image or video automatic grid generator associated to the Diffuse Approximation meshless Method. It seems to construct suitable grids either it concerns circular or more complex shaped object. The obtained point clouds have been tested with a Diffuse Approximation Method on transient heat diffusion or natural convection problems. Results are in agreement with the physical phenomena.

The obtained tool brings some facilities to the engineer or the searcher in terms of domain spatial discretization.

REFERENCES

Ahmadi, I., Sheikhy, N., Aghdam, M.M. and Nourazar, S.S., 2010. A new local meshless method for steady-state heat conduction in heterogeneous

materials. *Engeneering Analysis with Boundary Element*, 34, 1105_1112.

Boutayeb, A., Twizell, E.H., 1991. Finite-difference methods for twelfth-order boundary-value problems. *Journal of Computational and Applied Mathematics*, 35, 133_138.

Bradski, G.R., Kaebler, A., 2008. Learning OpenCV,Computer Vision with *OpenCV Computer Vision with the OpenCV Libarary*, O'Reilly Media, 234-244.

Kimme, C., Ballard, D. H., Sklansky, J., 1975. Finding circles by an array of accumulators. *Communications of the Association for Computing Machinery*, 18, 120_122.

Li, S., Zhu L., Zhang, Z.Q., Blake A., Zhang, H.J. and Shum, H., 2002. Statistical learning of multi-view face detection. *Proceedings of the 7th European Conference on Computer Vision*, May 2002 Copenhagen, Denmark.

Long, P., Jinliang, W. and Qiding, Z., 1995. Methods with high accuracy for finite element probability computing. *Journal of Computational and Applied Mathematics*, 59, 181_189.

Monaghan, J.J., 1992. Smooth particle hydrodynamics. *Annual. Reviews Astronomy and Astrophysics*, 30, 543-74.

Nayroles, B., Touzot, G., Villon, P., 1991. L'approximation diffuse, *Comptes Rendus de l'Académie des Sciences Paris*, 313, 133_138 Série II.

Shapiro, L. and Stockman, G., 2001. *Computer Vision. 1st ed*. New Jersey (USA):Prentice-Hall.

Sochman, J. and Matas, J., 2004. AdaBoost with totally corrective updates for fast face detection, *Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition*, pp.445-450.

Sophy, T., Sadat, H., Prax, C., 2002. A meshless formulation for three dimensional laminar natural convection. *Numerical. Heat Transfer part B: Fundamental*, 41, 433_445.

Viola, P. and Jones, M., 2001. Rapid object detection using a boosted cascade of simple features, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 01)*, pp.511-518.

CUSTOMER / SUPPLIER REQUIREMENTS AND BEHAVIOUR MODELING & SIMULATION IN SERVICE DELIVERY

Thècle Alix[(a)], Gregory Zacharewicz[(b)], Bruno Vallespir[(c)]

[(a)(b)(c)] University of Bordeaux IMS UMR CNRS 5218

[(a)]thecle.alix@ims-bordeaux.fr, [(b)]gregory.zacharewicz@ims-bordeaux.fr, [(c)]bruno.vallespir@ims-bordeaux.fr

ABSTRACT

Service has become over the years a very popular word discussed in the whole world. The economy is dominated by the tertiary sector of activity only one able to create new jobs because the service demand is not yet covered by the private or by the public sector. Citizens see services as a way to have access to basic or complex amenities, authorities as a way to address environmental problems and manufacturers as a way to differentiate them from the competition, to be closer to their customers and to improve the shopper experience. Services are of a huge importance in the national and international economy and are discussed in numerous domains: human service, business service, IT service, manufacturing area, etc. Several concepts related to service have merged as well as new scientific disciplines. Problematic linked to service design, service implementation, service operation management, service quality, service system simulation, product-service system design, service modeling are still under consideration and the multiplicity of the domain concerned failed to come up with unanimous answer. This paper proposes a contribution on service delivery process modeling and simulation that can potentially be used in any area. The proposed model is based on the most relevant concepts coming from a specialized literature review on services. A G-DEVS model of the service delivery process is then proposed.

Keywords: Service study, PSS, G-DEVS M&S

1. INTRODUCTION

Pragmatically, customers are currently looking for individual solutions to meet the challenge of their everyday life. They are torn between their willingness to have the ownership of physical products in a consumer economy and their new enthusiasm for a virtual economy based on use or on functionality. Products are perceived as containers of service corresponding to the product functionality. Two business models can be envisaged to address customers concerns: the first one consists in selling the product together with extra-services. The second one proposes to lease, to rent, to share or to pool the product (Tukker, 2004) and to sell the associated service.

Although the business model is different, one can conclude that from the customer point of view product and service are the two facets of a same object. Economically there exist products oriented service and use oriented services (Manzini, 2001) and besides some pure services. To summarize, everything can be considered as a service and the ability to shift from a product dominant logic to a service dominant logic is of paramount importance in the capital good industry. Service systems are a hallmark of the industrial economy.

Consequently, the management science end economics inputs have been endowed with other concepts proposed by the SSME, the IT industry, the mechanical engineering science together with academics from environmental and social science, etc. All the contributions have in common the willingness to develop models, methods and theories to support the shift. However till now, approaches are service domain centered and it is quite difficult to transpose a contribution from one domain to another one.

Our works rests on the provision of a generic model of Service Delivery allowing simulation. After a first recall on what a service is, we present the most relevant concepts from the disciplines abovementioned regarding our problematic. Then the paper initiates a possible use of modeling & simulation to characterize study and measure the capacity of entity called artifacts to receive or to deliver services. Each artifact can be alternately customer / supplier and it is associated to a service potential definition. M&S can support the understanding of this new paradigm by studying the artifacts behavioral aspects. The potential to supply a service by an artifacts and the need to receive it by another is also discussed to be quantified in order to be used in formal simulation models here G-DEVS models.

2. SERVICE LITERATURE REVIEW REGARDING SERVICE DELIVERY

The presentation of the service literature review is subjected to the limits of our work and mainly concerns the concepts related to service delivery: stakeholders, relation, activities, system dynamic, etc.

2.1. Service basic concepts

Many definition of the term exist having in common the three following elements: i.e the service provider, the service client and the service target (Spohrer, 2007).
The three main characteristics of service are:

- The co-creation of value: idea of the customer as co-producer of the value extracted from the service system and input to the service process
- Relationships: the relationship with the customer is of paramount importance and is a source of innovation and differentiation. Long-term relationships facilitate the ability to tailor the service offerings to the customers' needs
- Service provisioning: there is a provision service capacity to meet fluctuations in demands while retaining quality of service

2.2. Inputs from the SSME

The new scientific approach that merged around 2004 to study, design and implement service is the SSME (IBM, 2004). Defined as the application of science, management, and engineering disciplines, SSME proposes to build service knowledge and basic theories, to manage and optimize the process of creating value with service and to apply theory to solve practical service problems.

Among the highlighted concepts that we recall are the service dominant logic concept, the service system role-holders concept, the service mindset concept and the service system.

- Service-dominant logic (S-DL): The service-dominant logic world view, upon which service science is based, advocates that service is value co-creation interactions undertaken when service systems create, propose and realize value propositions, which may include things, actions, information and other resources. Value propositions are built on the notion of asset sharing, information sharing, work sharing (actions), risk sharing, as well as other types of sharing and exchange that can co-create value in customer- provider interactions.
- Stakeholders: also known as role-holders in service systems. Role-holders are people, or other service systems, that fill named roles in service systems. The two main roles in any service system are provider and customer.
- Service mindset: a focus on innovating customer-provider value co-creation interactions (service systems and value propositions, SSME qualified) that is combined with the interactional expertise capabilities of an adaptive innovator to enable team work across academic disciplines and business functional silos.
- Service system: is a dynamic value co-creation configuration of resources. Service systems are a type of system of systems; in which value proposition connect internal and external service systems. The smallest service system is a single person and the largest service system is the global economy.
- Service sourcing: agreed commodity definitions, identifying expected outcomes of customer needs and outcomes, determining cost drivers, defining and communicating requirements, defining supplier evaluation criteria

2.3. Inputs from the PSS community

At the same time other school of thought (mainly academics from environmental and social sciences (Baines, 2007) as well as more recently from engineering technologists) have focused on Product Service System in an environmental awareness perspectives. "Product(s) and service(s) combined in a system to deliver required user functionality in a way that reduces the impact on the environment" (Tukker, 2004). The definition extended through the years acknowledges that the concept of PSS also embraces value in use and sustainability (Goedkoop, 1999). Main contributions concern the principles, strategy, and development in PSS, service design methods and service engineering (Tomiyama, 2001). Concerning these two last point, the main value added is to consider each step of the system life cycle to ensure value delivery. Recent works concerning specifically service simulation are discussed hereafter.

2.4. Inputs from the IT industry

The service industry considers service system under the following definition: a configuration of technology and organizational networks designed to deliver services that satisfy the needs, wants or aspirations of customers. Service system includes: service provider, service customer, service environment and technical support. In the IT industry, software as a service is a widespread PSS (Bohmann, 2008).

The computer science has proposed a service-oriented architecture that rests on the combination of a process innovation with an effective-governance, a technological strategy centred on the definition and on the re-use of services. Here, a functionality is decomposed in a set of functions or of services supplied by components. A business service is a company functionality that seems to be atomic from the service consumer point of view. In this frame, a service is a connection to

3. DECISION SUPPORT TOOL BASED ON SERVICE MODELING AND SIMULATION

Based on the previous literature review, the first assumption we defend in the frame of service delivery is that each system designed and delivered can be seen as a unique complex coupled set composed of products subsets or components and services subsets or components. The second one is that everything is service which means that each product can be described through its functionalities designed to fit service needs. Based on the previous assumptions, we propose to model a service system as a set of services components.

There already exist works on service modeling and simulation (Alix, 2012). The conclusions of the bibliographical analysis are the following:

Regarding service modeling and PSS modeling: Service modeling is a recent domain, which has not yet adopted a unique common standard for developing

frameworks to manage services processes. There is a lack of one specific modeling language for PSS. The specification of Service Modeling can involve different process, application and actor components, which are essential to the service execution, but heterogeneous. The specification standards are numerous. Some authors transpose to service the administrative or production workflow process sequence description. Others use the graphical definition of a Service-Oriented Modeling Framework (SOMF). An essential breach concerns the model correctness checking. The W3C proposed an XML representation of Service Modeling Language (SML) accepted as a standard in the Service modeling community but the proposed description is more Computer Science Service Modeling oriented than industrial PSS oriented.

Regarding service simulation and PSS simulation: Most of the projects performed over the previous decade discuss the dynamic behavior of the system and were driven by the goal to provide information to the designers on how to handle the system and to verify desired properties. Others studies focus on Discrete-Event System modeling of PSS. The different researches identify the variables to be followed during simulation including the price, process costs lifetime, sales frequency, lifetime, etc. None research already specifically focus on customer quantifiable level of demand and supplier capacity to answer the need and synergies between products and services or services according to our previous assumptions.

4. SERVICE MODELING METHODOLOGY

4.1. Service definition statement

A service is an interaction between a provider that has a function and a consumer who has a need. A potential difference between both allows an exchange. This potential difference is the trigger of the service.

Touzi in his thesis has considered that the supplier has a potential level higher than the consumer (Touzi, 2011). Based on these service play rules, we will define and study the trigger conditions of service and types of service. The consumer is a service user, in the definition of a service, an object that has a potential negative service is regarded as a consumer and it is called potential need. The provider is the object that renders service to another. Its service potential is positive and will be called a function.

4.2. Service potential

The service is implemented through the existence of a potential difference expressed as follows:

Pot.service(A)>Pot.service(B) // Service A/B

The particular case of a hybrid object can be distinguished as it can be both provider (in French *Fournisseur*) and consumer (in French: *Bénéficière*) (see figure 1). This leads to:

Pot.service (A)>Pot.service (B)>Pot.service(C)

Figure 1: The hybrid object, both supplier and consumer

The nature of supplier or provider is an absolute value that is not relative to other objects. This allows telling from the above equation that A is considered as a supplier because its potential is the greater one; C is lower potential which defines it as a consumer while B is a consumer or supplier under the sign of its potential.

4.3. Notion of artifact

An artifact is an object which has undergone a transformation, however simple, by human and which is distinguished from another only created by natural phenomenon. On the other hand in management and process management, the artifact is any document (rule, graphic, procedure, etc.) identified within a process.

Thus it can be defined that an artifact has a capacity to provide a service. Moreover, this ability has an only interest regarding the need to consume of another object. From this postulate, we can understand that the concept of service can be considered only as part of a couple service provider/consumer.

According to the definitions so far, a subject is able to serve or consume. Yet an artifact capable of making several services of different nature, an object can also use several different types of services. In general we consider that an object can be a consumer or supplier of a range of services.

From the analysis of the complexity of the relationship between provider and consumer, it results multiple possible relationships between objects. The diagram figure 2 below can illustrates this concept.

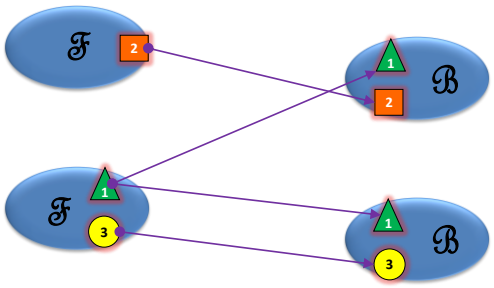


Figure 2: Multi relationships between objects

Given the complexity mentioned above we will define the functions and needs of an object from the following notations. *Fn (A)* (*provider*) means the ability of the object A to supply a service *n* when the notation *Bn (A)* (*Consumer*) will describe a need *n*.

4.4. Service delivery process

Service Delivery (SD) process is a service production process. For realizing the SD, a coupling between a consumer and a provider is required. The steps of the coupling are illustrated in figure 3.

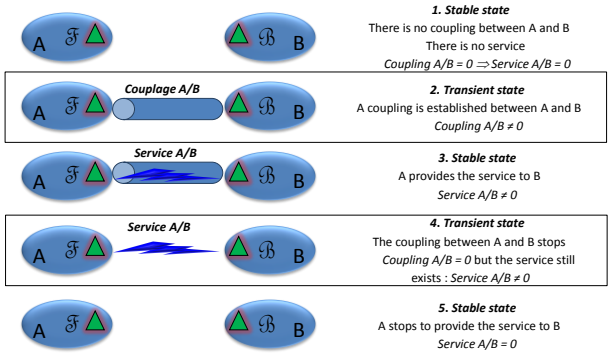


Figure 3: Service consumer / provider coupling steps

However, additional situations pre and post-coupling are required for each object. These situations are initialization phases; contextualization, de-contextualization and disconnection they are detailed below.

4.4.1. Trigger conditions of a service

A service can be triggered according to the figure 5 three cases:

- Case 1: The consumer requests the service, it is a pull service.
- Case 2: The vendor initiates and provides a service to consumers; it is a push service.
- Case 3: It is a third actor action that sends information to both supplier and consumer for a SD. This is called service-driven.

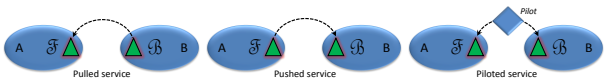


Figure 5: Triggering the Services

4.4.2. Notion of capacity and load

An object has a function that is a capacity to serve another object that has a need load. Hence the notion of capacity and load requires to be quantified. $CF_n(A)$ is the capacity function of the object A to make n service(s) belonging to a size interval [0, MaxCF]. The capacity can be Boolean, expressed on a continuous or discrete scale. Similarly we can define a load $CB_n(A)$.

4.4.3. Parameters of a SD process

Assuming two objects A and B, respectively provider and consumer, the possible situations that can initiate a SD procress are:

- $CF_n(A) = 0$: the service can't be made, regardless of the evolutionary stage of the SD process,
- $CB_n(B) = 0$: the service can't be made, whatever the stage evolutionary SD process,
- $CF_n(A) = CB_n(B)$: the function capacity perfectly matches the intensity of need of B, a is fully occupied and the need of b is filled,
- $CF_n(A) < CB_n(B)$: the service needed by b can't be offered by a,
- $CF_n(a) > CB_n(b)$: b can receive entirely the service needed, a is partially occupied to deliver.

5. G-DEVS MODELLING AND SIMULATION OF SERVICES

Indeed, the service modeling requires also the modeling of the interactions between multiple services; this process can lead very quickly in a significant level of complexity. We therefore focus in this paper on the establishing operations for a single service coupling in the G-DEVS formalism (Giambiasi, 2000). This discrete formalism is selected for its formal property and its time management. We propose to model each service component through a G-DEVS model based on attributes. The model attributes are described from a qualitative and quantitative point of view and all elements (actors and material) that interact within its environment are required. Once the description is complete, the described component can integrate a G-DEVS based library of service components: service repository. The prospect of a break in service into four subsets then seemed obvious. These four subsets are:

- The object requesting the service (here object B) is the consumer, the one who feels a need.
- The service provider (here object A) is the supplier, the one who has the ability to satisfy a need.
- The coupling is the association between two objects that will achieve the SD process.
- The SD process is the service producing.

5.1. The service requester: Object B

The atomic model corresponding to the object B is an applicant for a service. The G-DEVS model (Figure 6) and its operations are detailed. This model describes its behavior during the process of SD coupling.

The G-DEVS model (figure 8) follows the coupling steps described in the § 5.3. It communicates with the service delivery model. To assume the simulation dynamical execution, the states, event and temporal information have been added. These data are not related to any information coming from a real system.

1. Every 5 time units (arbitrary chosen), the model B launches the comparison of its P/S potential intensity over a threshold value. This comparison is expressed as a condition on the internal transition.
2. The model function sends a request (DFb1) to the object service provider (the object A).
3. Positive response is received when the object A is able to achieve this service.
4. Negative response is received when A is not available or not competent for this service achievement, a request will be send to another object.
5. Sending an acknowledgment to the object A to tell them that the service can be achieved and to lock between them a SD process.
6. Sending "ok_couplage" to external produce model.
7. Pending the external event "fin_SD" meaning that the production of the service is completed.
8. Back to the waiting phase, the service was rendered, the intensity of the object B to decrease the capacitance value function of the object A demand. We note that some strategy can keep the value of the object B since some service potential is infinite.

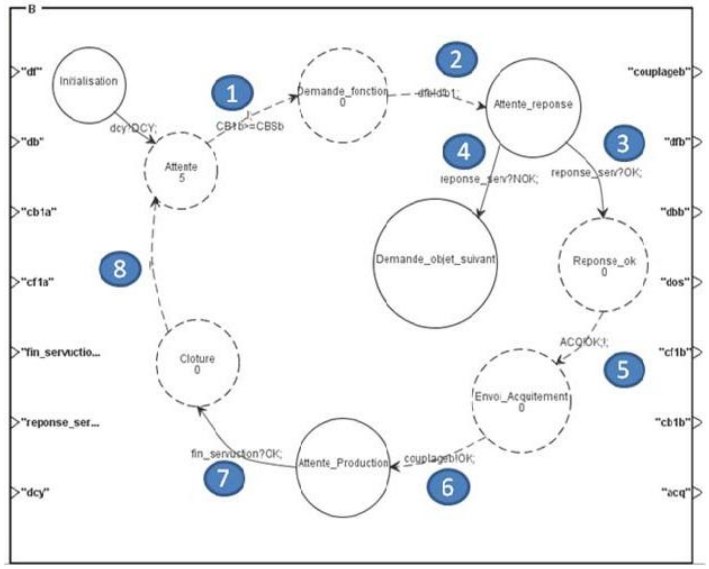


Figure 6: Service Requester Object B

5.2. The service provider: Object A

This section is introducing the G-DEVS atomic model of the service provider object A. An explanation of the model and its operations is detailed here.

1. The model A is expecting a request from object B.
2. A tests the ability to get the function (CF1a) and the load required by the object B (CB1b). If the ability and capacity of A is greater than or equal to the need and load of B then the service is feasible.
3. Same as step 2; if the capacity is lower than the load of B then the service will not be feasible.
4. The service is not feasible, the supplier is informed the consumer via a message object "NOK".
5. The service can be done, A sends to B "OK".
6. The model is waiting for the acquittal of the object B to produce the service.
7. The model is awaiting a response from an external "process" models and indicating that the service is in progress.
8. The model is pending an external event "fin_SD_OK" from the "process" model.
9. The process is over and the ability to provide the function of the particular service A can be reduced (consumption) or increased (experience).

The service is completed. The ability of A can stay decreased or can recover its initial value with a gain of experience that increases its ability (the hypothesis can be to gain 10% capacity acquired for each SD).

5.3. Coupled model

The coupled model (figure 7) is presenting the G-DEVS component required for the global simulation of the process. The component 1 & 2 are the A and B models. The component 3 is a coupling model used to connect the models that will be paired for the service. The component 4 is used to orchestrate the process steps defined in § 5.3.

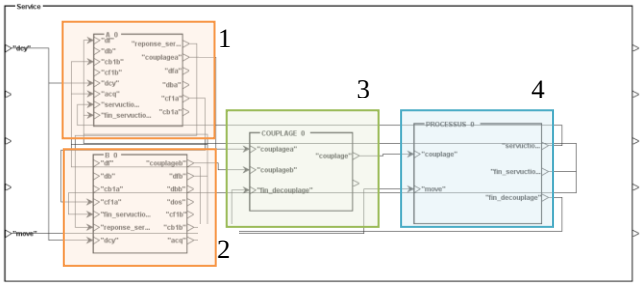


Figure 7: Coupled Model of SD

5.3.1. Coupling (and decoupling)

This atomic model is permitting the coupling orchestration of two objects before the SD process; managing the notion of decoupling, which is at the end of the service. This model is labeled 3 in figure 6. Its main operations are the followings.

It starts by waiting an event from the object A informing that is ready and looking for a SD coupling. Then it is waiting for an event from the object B answering that it is available and capable regarding the load and competence required by A. Then the coupling can be realized. Gathering the information, the model is informing both participants. Then it is waiting for an event informing of the end of the SD to return to the standby state. This is producing the end of the coupling.

5.3.2. The SD process

A last G-DEVS atomic model is required to defines SD the characteristics and simulate the SD process behavior (SD model number 4 in figure 6).

This model starts by waiting an event to be informed that a coupling is ok. At this time, it computes the characteristics of this SD including coupling, duration, quantity of load and experience acquired at the end. It informs by output sending the "SD" settings to the model participants and set them in progress. When the service is ended, it informs the participants by sending to output "fin_SD".

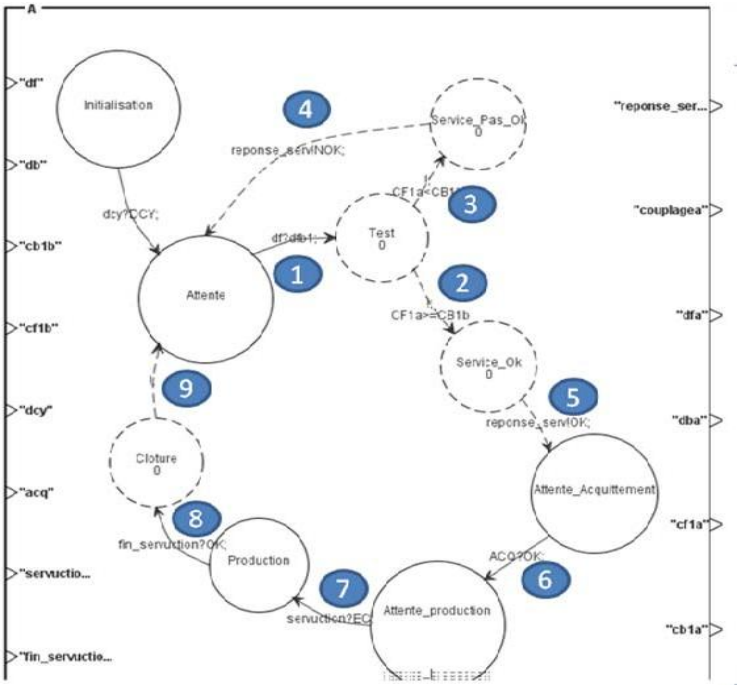


Figure 8: Service Provider Object A

6. CONCLUSION

The paper has provided a preliminary work to define the base of what can be modeled and simulated about the concept of PSS. It has focused on decomposing the SD steps used to couple a service supplier and a receiver. It opens the research in this domain where few works considering dynamic are done. The simulation is considered under the idea of defining a space of objects moving autonomously. When an object needs a service, it will try to connect to a service supplier in its neighborhood. The construction of this space is still under the consideration of the authors at the moment.

ACKNOWLEDGMENTS

This research work has involved Moustapha Gueye and Nicolas Compere during their master student's internship period at IMS lab.

REFERENCES

Alix, T., Zacharewicz, G. 2012. Product-service systems scenarios simulation based on G-DEVS/HLA: Generalized discrete event specification/high level architecture. *Computers in Industry* 63(4), 370-378

Baines, T. S. et al. 2007. State-of-the-art in product service-systems. DOI: 10.1243/09544054JEM858

Compere, N., Guye, M. 2012. Analyse, Modelisation et Simulation d'un Service, *Master intermediary report*, University of Bordeaux.

Goedkoop, M. et al. 1999. Product Service-Systems, ecological and economic basics. Report for Dutch Ministries of Environment (VROM) and Economic Affairs (EZ).

Berkowich, M. et al. 2011. Requirements Engineering for Product Service Systems: A State of the Art Analysis. *Business and Information Systems Engineering*. DOI 10.1007/s12599-011-0192-2

Manzini, E. et al. 2001. Product service systems: using an existing concept as a new approach to sustainability. *Journal of Decision Research* 1(2).

Spohrer. J, Maglio. P, Bailey. J et Gruhl. D. 2007. Steps Toward a Science of Service Systems. IEEE Computer 40(1), 71-77.

Tomiyama, T. 2001. Service engineering to intensify service contents in product life cycles. In Proceedings of the Second International Symposium on Environmentally Conscious Design and Inverse Manufacturing, 613–18.

Touzi, W., Alix, T., Vallespir, B. 2012. Contribution To The Development Of A Conceptuel Model Of Service And Services Delivery. *Proceedings of the APMS conference*, Rhode, Greece

Touzi, W., 2011. *Conceptualisation et modélisation de la production de service: application aux domaines de la santé et de l'enseignement.* Thesis (PhD). University of Bordeaux.

Tukker, A. 2004. Eight types of product-service system: eight ways to sustainability experience from suspronet. *Bus. Strat. Env.* 13, 246-260

Zacharewicz G., Alix T., Vallespir B., 2009. Services Modeling and Distributed Simulation DEVS / HLA Supported, *IEEE Proceedings of the Winter Simulation Conference*, pp. 3023-3035, Austin, Texas, USA

MODELING SELECTIVITY BANKS FOR MIXED MODEL ASSEMBLY LINES

Alex Blatchford[(a)(1)] , Yakov Fradkin[(a)(2)] , Oleg Gusikhin[(a)(3)],
Ravi Lote[(b)(4)], Marco Pucciano[(c)(5)], Onur Ulgen[(b,d)(6)]

[(a)] Ford Motor Company, USA
[(b)] PMC, USA
[(c)] University of Genova, Italy
[(d)] University of Michigan-Dearborn, USA

[(1)]ablatchf@ford.com, [(2)]yfradkin@ford.com, [(3)]ogusikhi@ford.com,
[(4)]rlote@pmcorp.com, [(5)]marco.pucciano@gmail.com, [(6)]ulgen@umich.edu

ABSTRACT
This paper presents a simulation study on the influence of different designs and operational policies of selectivity banks on the efficiency of automotive assembly sequencing. Specifically, we analyse whether or not a bypass lane could improve the performance of selectivity banks. A special purpose simulation framework has been developed and implemented using Witness simulation integrated with an Excel file containing the dataset and the decision making algorithms.

1. INTRODUCTION

A mixed-model automotive assembly plant is typically comprised of several different production areas including the body shop, paint shop and final assembly as shown in Figure 1. The plant's efficiency largely depends on the sequence of different types of vehicles moving down the line. The quality of the vehicle sequence can be defined by how well it satisfies the preferences or constraints and supply chain considerations posed by different production areas: see for example Fradkin (2006).

In general, different production areas have different preferences in terms of ideal sequences (Blatchford 2008). For example, paint operations prefer batching of similar colours to minimize changeovers (Spieckermann *et al.* 2004), while body shop and final assembly prefer smoothing of similar vehicle types to balance the labour and supply chain consumption rates.

To accommodate these conflicting requirements, many automakers have faced the problem of resequencing the vehicle flow by adding storage/ resequencing areas between the various assembly areas. Figure 1 shows these areas within the assembly plant, denoting them with ellipses. In particular:

- the White Body Storage (WBS) is located between the body shop and the paint shop;
- the Colour Rescheduling Storage (CRS) is located within the paint shop;
- the Painted Body Storage (PBS) is located between the paint shop and the trim shop.

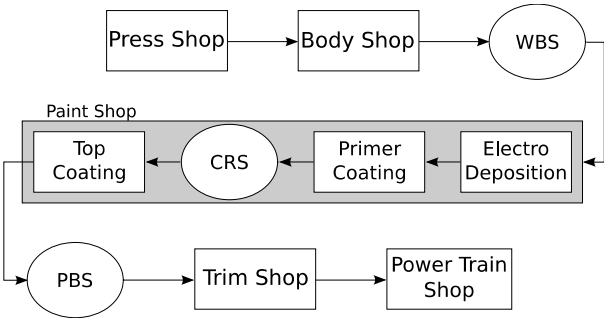


Figure 1: Storage Areas Scheme.

Those not only allow absorption of fluctuations in output from the previous phase and in input for the next one, but they can be also used as an opportunity to resequence the stream of items in order to better satisfy the constraints of the downstream process.

Under the term "constraints," we mean rules that are used in order to prevent certain sequences of models on the assembly line – as explained in Nagane (2002) – and therefore even prevent situations in which the lane could be imbalanced.

The resequencing activities could be useful not only between the production areas having different sequencing requirements, but also between the areas having the same constraints. For example, a defect could be discovered on a certain vehicle in the upstream area, requiring that the vehicle be put aside for the repair. This disrupts the initial, possibly perfect, sequence. To contain the effects of the disruption, some form of resequencing may be needed before feeding the vehicles to the downstream area.

Several authors such as Inman (2003), Inman *et al.* (1997), and Gusikhin *et al.* (2007) analysed the resequencing structures and algorithms. The commonly used resequencing structures are:

- automated storage and retrieval system (ASRS or AS/RS);
- pull of table;
- repair holding area;
- selectivity bank (SB).

Each of the above has advantages and disadvantages. For example, ASRS has the best resequencing capabilities (it allows random access) but is typically very expensive (both in initial investment and in work in progress costs) and requires a lot of physical space (which means that ASRS cannot be easily added to an existing plant). In comparison with ASRS, SB has marginal resequencing capabilities (described in Section 3) but is simpler and less expensive, which allows its use when ASRS is not a viable option.

The aim of this work is to develop methods and tools that could improve the selectivity bank's performance. Particularly we study whether or not the introduction of a bypass lane could improve SB's resequencing capabilities. We built a tool that could test whether adopting a bypass lane could be useful, given the characteristics of the plant and the vehicles that have to be assembled.

We used Modelling & Simulation (M&S) to achieve these goals. The benefits of M&S have been widely discussed in the literature. Karakal (1998) and Banks (1998) point out that the application of M&S is particularly useful in order to perform what-if analyses and analyse the behaviour of complex systems. The usefulness of the simulation approach to problems of this type has been also pointed out in Han *et al.* (2003) (in which the authors used a similar approach in order to reduce the number of colour changes within the paint shop), in Ulgen (1994), and in Park *et al.* (1998).

We used a simulator to represent the plant and study the behaviour of the system with and without (respectively WIB and WOB) a bypass lane. In this paper, we describe the theoretical advantage related to the usage of a bypass lane and discuss the results.

This paper is structured as follows: in Section 2 we introduce the selectivity banks and their characteristics; in Section 3 we describe the simulation model that we used to perform the analyses. The results and conclusions are given in Sections 4 and 5.

2. SELECTIVITY BANKS AND BYPASS LANE

Selectivity bank (SB) is a kind of a buffer that is commonly used for both storage and resequencing purposes, as shown in Narayanaswamy *et al.* (1997). As highlighted in Figure 2, it usually is a multilane structure in which each lane can carry any type of model and that is characterized by a certain length – i.e. the maximum number of vehicles that it can contain. In some applications, particular lanes may be dedicated to a certain single model (in order to be able to easily access it) or to a subset of models.

SB could be a traditional structure with a certain number of lanes (each one characterized by a certain length) – like the one in Figure 2 – but it could also include a bypass lane as in Figure 3 and/or a return lane like the one in Figure 4 (just the first two configurations have been considered in this work).

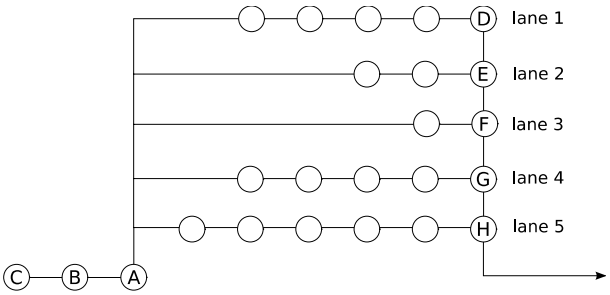


Figure 2: Example of Selectivity Bank Without Bypass Lane.

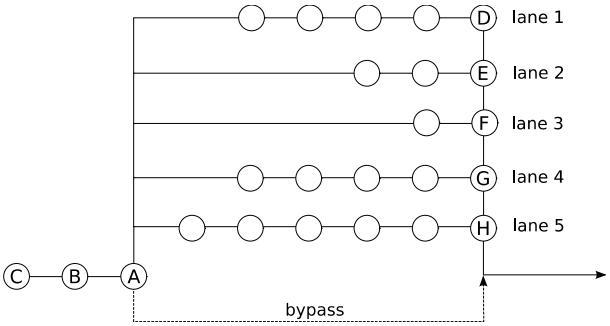


Figure 3: Example of Selectivity Bank With Bypass Lane.

The two major processes that have to be managed in SB are the input and the output ones. The input process determines in which lane a certain item has to be sent when it enters the SB, while the output process determines which vehicle should be sent downstream to the next phase. These two processes could be managed in different ways, typically using heuristics. Two important characteristics of SB are the maximum number of selection and retrieval points that determine the resequencing capabilities of this kind of buffer; in both cases they're equal to the number of lanes (bypass lane included).

3. THE SIMULATION MODEL

To test how different SB designs could affect system performance, we created a simulation model. Model building methodology included defining the problem, designing the study, designing the conceptual model, formulating inputs and assumptions, building and verifying the model, conducting experiments and documenting the results. We developed the following three components of the model:

- a simulator (realized using Witness);
- input data (in an Excel spreadsheet);
- decision making algorithms (realized in VBA within the Excel file).

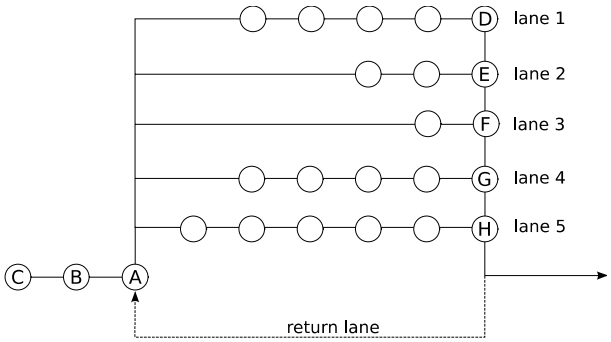


Figure 4: Example of selectivity bank with Return Lane.

One of the modelling decisions was how to represent the assembly line constraints. We reduced them to "no more than x out of y" constraints specifying the maximum allowed number (x) of vehicles that have a certain option out of (y) of consecutive items. Figure 5 shows how these constraints are stored in the Excel file.

We defined a metric to measure the performance of the system: with each constraint violation we associated a penalty (time lost, measured in seconds). According to this framework the performance of the system is measured as the total penalty that has to be sustained after having resequenced the whole input sequence. Note that these penalties could be different from plant to plant and therefore just a comparison between different designs of the same plant is meaningful. In other words, this definition doesn't allow us to compare performance of different plants.

The model's three components are linked together as shown in Figure 6: Witness drives the simulation and passes input data to the Excel spreadsheet when a decision making process is needed; within the Excel file the VBA algorithms use this information to decide to which lane a vehicle that is currently entering the SB should be sent and from which lane a vehicle should be retrieved for sending downstream of the SB. The decision is communicated to the simulator via the Excel spreadsheet.

In summary, the developed framework is based on the following assumptions:

- The possibility of implementing a bypass lane is considered (as previously mentioned).
- The lanes (but not the bypass lane) are shared between the various vehicle models.
- The constraints are considered as explained above.

Model verification was completed to make sure that the algorithm and programming is error free. The modelling analysts used various techniques to verify the model. These techniques included:

	Constraint	1	2	3	4
	x	2	3	2	2
	in y	10	4	6	6
(secs)	Penalty	60	30	280	60

Figure 5: How the Constraints are Stored in the Excel file.

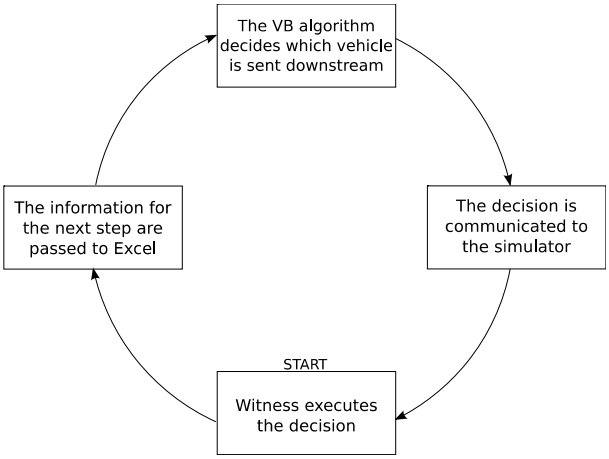


Figure 6: How the Three Components are Linked Together.

- Using modular programming concepts;
- Sending only one vehicle into the model and ensure that the flow and cycle time is correct;
- Undertaking "directional testing" (e.g., if a cycle time increases, throughput should decrease or remain the same);
- Executing deterministic runs: removing all randomness and making sure that the results match static analysis;
- Error-trapping the events that "can't" happen;
- Making sure that the time units and distance units are consistent throughout the model;
- Using simulation traces extensively.

The implementation was corrected based on the errors found in the underlying model, often resulting in retesting to ensure integrity of the programming done. However, since the model is built at the abstract level, validation was not possible.

While building and verifying the model, we used the help of subject matter experts from assembly plants, utilized input data coming from the real plants and performed sensitivity analyses.

Table 1: The considered selectivity buffer's design for the 1st Test Case.

#	Lanes	Length	Bypass Lane
1	5	8	No
2	10	8	No
3	5	8	Yes
4	10	8	Yes

Table 2: The considered selectivity buffers' design for the 2nd Test Case.

#	Lanes	Length	Bypass Lane
1	3	17	No
2	2	17	Yes
3	5	7	Yes
4	6	7	No

Table 3: Some Details of the 1st Test Case.

3.1.	% of vehicles	x	y
Option 1	6.9	2	10
Option 2	72.1	3	4
Option 3	20.9	2	6
Option 4	14.4	2	6
Option 5	1.6	1	8
Option 6	13.7	5	7
Option 7	3.1	2	10
Option 8	4.9	2	8
Option 9	39.1	3	4
Option 10	2.1	1	3
Option 11	21.9	2	5
Option 12	4.6	1	8
Option 13	14.4	2	6
Option 14	1.1	2	5

Table 4: Some Details of the 2nd Test Case.

3.2.	% of vehicles	x	y
Option 1	61.9	2	3
Option 2	7.8	1	2
Option 3	14.4	1	3
Option 4	7.6	1	3

4. RESULTS

We used two dataset from different plants, as shown in Tables 1 and 2. With the first dataset we've used a perfect (having no constraint violations) input sequence that has been disrupted considering four different levels of repair (5%, 10%, 15%, 20%). For the second dataset we've just considered a random input sequence (and therefore we didn't use any level of disruption/repair as in the previous case).

These different choices have been made in order to create a more realistic model of the considered plants.

The first test case has 14 constraints and an input sequence composed of 14,814 vehicles, while the second one has only 4 constraints and 8,809 vehicles. The characteristics of the input sequences of the considered test cases are shown in Tables 3 and 4; the second column in those tables shows the percentage of vehicles in the input sequence that had a particular option.

The first dataset is more complex and presents more difficult constraints than the second one. The second dataset is simpler: it only has four constraints, of which only the first one is tight, and the remaining three should be easily satisfied.

The obtained results for the 1st test case are shown in Figures 7, 8, and 9. These tests have been performed using two different heuristics in order to manage the SB, called policy 1 and 2. Policy 1 – that is a greedy algorithm – has been designed in order to consider just the explicit constraints of the downstream phase of the production plant and perform a local optimisation at every step; this means that at every iteration the algorithm evaluates which vehicle determines the minimum selection penalty.

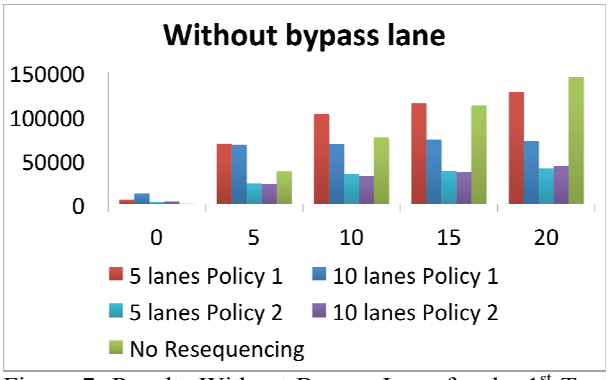


Figure 7: Results Without Bypass Lane for the 1st Test Case.

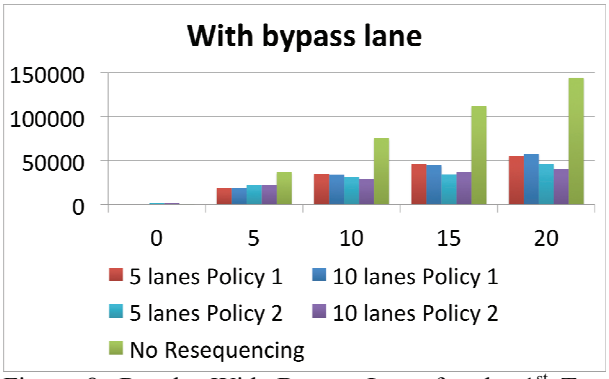


Figure 8: Results With Bypass Lane for the 1st Test Case.

Policy 2 is more complex than the previous one; it considers in particular not just the explicit constraints of the assembly area but spread the vehicles with certain characteristics within the production sequence.

This is realized considering an objective function (used in order to compare the various vehicles) that is composed of two parts:

- the total penalty related to the constraint violations;
- an additional term that is used in order to give priority to vehicles that present options that should be selected more frequently;

These two terms are then adequately weighted in order to give more or less importance to the first or to the second one using another parameter.

This more complex logic allows to better adjust the behaviour of the heuristic to the characteristics of the input sequence than Policy 1; therefore, it should be – in most cases – more efficient than this last one, but also substantially more difficult to implement in practice.

As expected with a higher number of lanes, the results are better; this happens both with and without a bypass lane. This is related to the characteristics of SBs and in particular to their resequencing capabilities (described in Section 2). The interesting thing we noticed is a big difference in terms of performance when implementing a bypass lane while using the first policy.

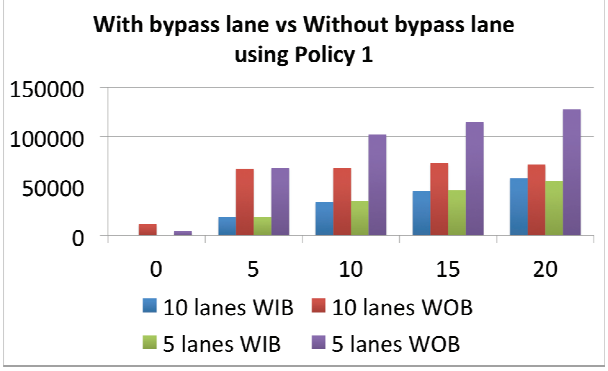


Figure 9: Comparison of the Results With and Without Bypass Lane for the 1st Test Case using the 1st Policy.

As can be seen in Figure 9, when using Policy 1, the SB with a bypass lane outperforms the SB without a bypass lane at any considered disruption level. However, this observation does not hold true when using a more efficient Policy 2, as can be seen in Figure 10. Based on previous results, we performed some other tests, in particular in order to compare whether using a bypass lane is more – or less – efficient than adding another traditional resequencing lane.

Using the first test case, we compared an 11 by 8 SB configuration without a bypass lane *vs.* a 10 by 8 SB configuration with a bypass lane. The results are reported in Figure 11.

The obtained performance measures indicate that adding a bypass lane results in performance improvement that is almost equal to the one that could be obtained with an additional resequencing lane.

We performed similar comparison using the second test case, as shown in Figure 12, under both Policies 1 and 2. This test case had a random input sequence, thus we didn't need to consider various disruption levels.

As seen in Figure 12, implementing a bypass lane doesn't seem to improve the performance of the buffer.

5. CONCLUSIONS

We developed a simulation network consisting of a Witness driver and an Excel file containing the dataset and the VBA decision making algorithms.

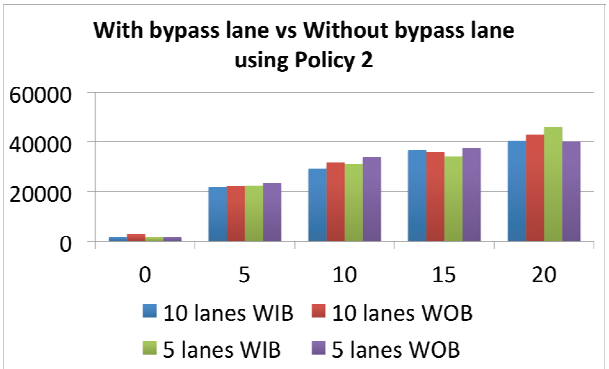


Figure 10: Comparison of the results with and without bypass lane for the first test case using the second policy.

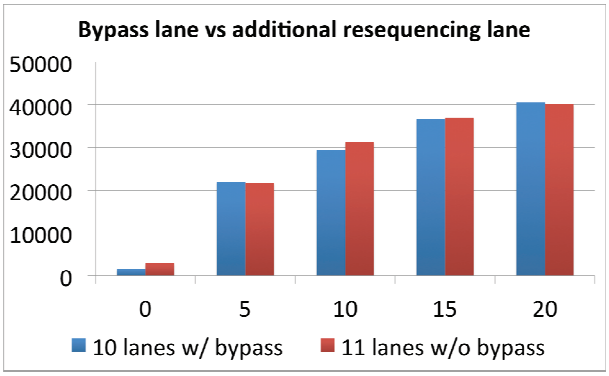


Figure 11: Comparison of the results with a bypass lane *vs.* the ones with an additional resequencing lane.

This framework allows us to simulate what happens within a selectivity bank, typically used in assembly lines both for storage and resequencing activities. It could be used for two main purposes:

- As a decision support tool that could be used during the design phase of a new plant;
- To perform what-if analyses both regarding the usage of different managing policies for the bank and for its layout.

We evaluated two input sequences, and we've tested them with different designs (in particular with or without using a bypass lane) and different sequencing policies. We investigated in particular whether using a bypass lane could improve SB's performance both in comparison with layouts with the same number of resequencing lanes and with others characterized by the same total number of lanes.

The obtained results seem to indicate that different policies could significantly affect system performances. In particular the more complex Policy 2 (described in Section 4) yields better results than its counterpart (Policy 1).

Further considerations regarding the obtained results are that:

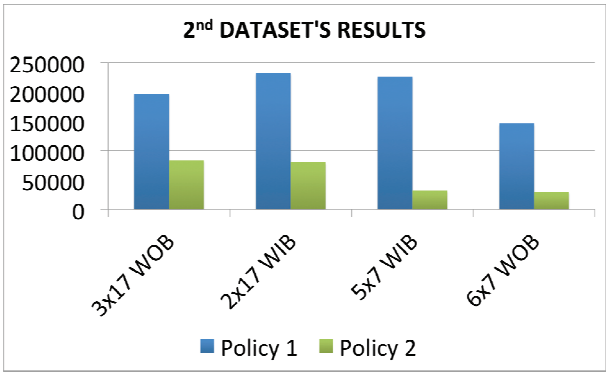


Figure 12: Comparison of the results with and without bypass lane for the second test case.

- a higher number of lanes tends to reduce the performance advantage of the more complex policies;
- the usage of a bypass lane tends to reduce the difference in terms of performance between policies 1 and 2 in certain conditions (in particular, using the more complex dataset).

Regarding the comparison between different SB designs, the results indicate that the usage of a bypass lane does not significantly affect system performance as compared with traditional designs with the same total number of lanes. Nevertheless, this could be considered an interesting result because it indicates that even if the current plant layout doesn't allow to increase the number of traditional resequencing lanes it is sufficient to implement a bypass lane (that could be for example extremely short) in order to obtain the same performance improvement.

Future work could allow to use this framework to evaluate buffer performance depending on the input mix complexity as well as to consider more than just one buffer of the plant. It could allow plant managers to use this tool to formulate recommendations regarding:

- possible improvements of the buffer's design;
- possible new policies used in order to manage resequencing activities;
- best combination of the previous elements considering a certain – fixed – budget dedicated to plant modifications.

REFERENCES

Banks, J., 1998. *Handbook of Simulation*. John Wiley & Sons, Inc.

Blatchford, A., 2008. *Quantifying Product Mix on Flow Lines with High Product Variety*, MSc Thesis, University of Cardiff.

Fradkin, Y., 2006. *Automotive Assembly Line Sequencing Application*. Ford Technical Report SRR-2006-0142, Ford Motor Company, Dearborn, USA.

Gusikhin, O., Caprihan, R., and Stecke, K., 2007. Least in-sequence probability heuristic for mixed-volume production lines. *International Journal of Production Research*, 46(3):647–673.

Han, Y., Zhou, C., Bras, B., McGinnis, L., Carmichael, C., and Newcomb, P., 2003. Paint Line Color Change Reduction in Automobile Assembly Through Simulation. *In Proceedings of the 2003 Winter Simulation Conference.*

Inman, R., 2003. ASRS sizing for recreating automotive assembly sequences. *International Journal of Production Research*, 41:847–863.

Inman, R., Bhaskaran, S., and Blumenfeld, D., 1997. In-plant material buffer sizes for pull system and level-material-shipping environments in the automotive industry. *International Journal of Production Research*, 35:1213–1228.

Karakal, S.C., 1998. A novel approach to simulation modeling. *Computers & Industrial Engineering*.

Nagane, S., 2002. *Simulation Study of Multilane Selectivity Bank in Automotive Industry*, MSc Thesis, University of Kentucky.

Narayanaswamy, R., Jayaraman, A., and Gunal, A.K., 1997. *A Sortation System Model*.

Park, Y.H., Matson, J.E., and Miller, D.M., 2008. Simulation and Analysis of the Mercedes-Benz All Activity Vehicle (AAV) Production Facility. *In Proceedings of the 1998 Winter Simulation Conference*.

Spieckermann, S., Gutenschwager, K., and Voß, S., 2004. A sequential ordering problem in automotive paint shops. *International Journal of Production Research*.

Ulgen, O., 1994. The Role of Simulation In Design And Operation Of Body And Paint Shops In Vehicle Assembly Plants. *In Proceedings of The European Simulation Symposium*.

INTERMITTENT DEMAND FORECASTING AND STOCK CONTROL: AN EMPIRICAL STUDY

Adriano O. Solis[(a)], Letizia Nicoletti[(b)], Somnath Mukhopadhyay[(c)], Laura Agosteo[(d)], Antonio Delfino[(e)], Mirko Sartiano[(f)]

[(a)] School of Administrative Studies, York University, Toronto, Canada
[(b) (d) (e) (f)] MSC-LES, Mechanical DepartmentUniversity of Calabria, Rende, Italy
[(c)] Information & Decision Sciences Department, The University of Texas at El Paso, El Paso, Texas, U.S.A.

[(a)]asolis@yorku.ca, [(b)]letizia@msc-les.org, [(c)]smukhopadhyay@utep.edu, [(d)]agosteo.laura@libero.it, [(e)]delanto@hotmail.it, [(f)]mirko.sartiano@yahoo.it

ABSTRACT

Statistical accuracy measures are generally used to assess the effectiveness of demand forecasting methods. In the final analysis, however, these methods should be judged according to whether they actually lead to better inventory control performance. We empirically evaluate four methods (simple moving average, single exponential smoothing, Croston's method, and the Syntetos-Boylan approximation) in terms of statistical forecast accuracy and, more importantly, inventory system efficiency. We apply four forecasting methods to an industrial dataset involving more than 1000 stock-keeping units of a firm in the professional electronics sector. Demand is often intermittent, erratic, or both (i.e., lumpy). We devise and use a two-stage distribution involving uniform and negative binomial distributions to model the actual demand distribution, where possible. We then simulate the stock control performance of a (*T*,*S*) inventory system with respect to target customer service levels.

Keywords: intermittent/lumpy demand forecasting, forecast accuracy, inventory control, order-up-to periodic review system, simulation

1. INTRODUCTION

Demand for a stock-keeping unit (SKU) is said to be *intermittent* if there are periods in which demand is zero. When demand is intermittent and there are large variations in demand sizes, demand is said to be *lumpy*. Syntetos, Boylan, and Croston (2005) proposed a theoretically coherent scheme for categorizing demand into smooth, erratic, intermittent, or lumpy. In this categorization scheme, the average inter-demand interval (*ADI*) and the squared coefficient of variation (CV^2) of demand are compared with cutoffs of 1.32 for *ADI* and 0.49 for CV^2, as follows:

- *smooth* demand: $ADI < 1.32$, $CV^2 < 0.49$;
- *erratic* demand: $ADI < 1.32$, $CV^2 > 0.49$;
- *intermittent* demand: $ADI > 1.32$, $CV^2 < 0.49$;
- *lumpy* demand: $ADI > 1.32$, $CV^2 > 0.49$.

This categorization scheme has been cited and applied by various researchers (e.g., Ferrari, Pareschi, Regattieri, and Persona 2006; Gutierrez, Solis, and Mukhopadhyay 2008; Altay, Rudisill, and Litteral 2008; Mukhopadhyay, Solis, and Gutierrez 2011).

In the intermittent demand forecasting literature, many papers have been published on the relative performance with respect to statistical measures of accuracy of various forecasting methods, most notably simple exponential smoothing (SES), Croston's method (Croston 1972), and an estimator proposed by Syntetos and Boylan (2005). Schultz (1987) suggested that separate smoothing constants, α_i and α_s, be used for updating the inter-demand intervals and the nonzero demand sizes, respectively, in place of Croston's single smoothing constant α. We note, however, that Mukhopadhyay, Solis, and Gutierrez (2011) investigated separate smoothing constants, α_i and α_s, in forecasting lumpy demand and did not observe any substantial improvement in forecast accuracy.

Syntetos and Boylan (2001) pointed out a positive bias in Croston's method arising from an error in his mathematical derivation of expected demand. They proposed (Syntetos and Boylan 2005) a correction factor of $\left(1-\alpha_i/2\right)$ – where α_i is the smoothing constant used in updating the inter-demand interval estimate – to be applied to Croston's estimator of mean demand. The revised estimator is now often referred to (e.g., Gutierrez, Solis, and Mukhopadhyay 2008; Boylan, Syntetos, and Karakostas 2008; Babai, Syntetos, and Teunter 2010; Mukhopadhyay, Solis, and Gutierrez 2011) in the intermittent demand forecasting literature as the Syntetos-Boylan approximation (SBA).

We also evaluate the 13-month simple moving average (SMA13) method, which is based upon dividing the 52 weeks in a year into 13 four-week "months". SMA13 has been applied in a number of recent intermittent demand forecasting studies (e.g., Syntetos and Boylan 2005, 2006; Boylan, Syntetos, and Karakostas 2008) in view of its being built into some commercially available forecasting software.

This paper is organized as follows. In section 2, we discuss the forecasting methods under evaluation, the statistical measures of forecast accuracy that we use, and the nature of the industrial dataset and how data partitioning is performed. In the next section, we propose a two-stage approach to the modeling of demand distribution. We proceed to report on our empirical investigation of forecasting performance, based upon statistical accuracy measures, on the performance block of the actual data and on the simulated demand distribution. The performance of the forecasting methods in terms of inventory systems efficiency is reported in section 4. We present our conclusions in the final section.

2. FORECASTING METHODS AND DEMAND DATA

2.1. Forecasting Methods and Accuracy Measures

We evaluate four methods that are well-referenced in the intermittent demand forecasting literature: SMA13, SES, Croston's, and SBA. For the latter three methods which involve an exponential smoothing constant α, low values of α up to 0.20 have generally been suggested for lumpy demand (e.g., Croston 1972; Johnston and Boylan 1996). We evaluate four α values of 0.05, 0.10, 0.15, and 0.20 as used in a number of recent studies (e.g., Syntetos and Boylan 2005; Gutierrez, Solis, and Mukhopadhyay 2008; Mukhopadhyay, Solis, and Gutierrez 2011).

In the following sections where our results are presented, we only report those pertaining to SBA and not those for Croston's method, as we have found the former to consistently outperform the latter.

Gutierrez, Solis, and Mukhopadhyay (2008) and Mukhopadhyay, Solis, and Gutierrez (2011) found a neural network (NN) model, when applied to an industrial dataset exhibiting lumpy demand, to perform better overall than the SES, Croston's and SBA methods across different scale-free error measures. However, NN modeling requires a substantial number of time periods to 'train' or calibrate the model, which is not the case in the current study.

In addition to applying the more commonly used root mean squared error (RMSE) and mean absolute deviation (MAD) as forecast accuracy measures, we have also used mean absolute percentage error (MAPE), which is the most widely used accuracy measure for ratio-scaled data. The traditional MAPE definition, which involves terms of the form $|E_t|/A_t$ (where A_t and E_t, respectively, represent actual demand and forecast error in period t), fails when demand is intermittent. We applied an alternative specification (e.g., Gilliland 2002) of MAPE as a ratio estimate, which guarantees a nonzero denominator:

$$\text{MAPE} = \left(\sum_{t=1}^{n} |E_t| \Big/ \sum_{t=1}^{n} A_t \right) \times 100 \cdot \qquad (1)$$

Eaves and Kingsman (2004) used the same three error statistics (MAPE, RMSE, and MAD) in comparing the performance of several methods (SES, Croston's, SBA, 12-month simple moving average, and the previous year's simple average) in forecasting demand for spare parts for in-service aircraft of the Royal Air Force (RAF) of the UK. They found SBA to provide the best results overall using MAPE, but the 12-month simple moving average yielded the best MADs overall. Willemain, Smart, Schockor, and DeSautels (1994) used median absolute percentage error (MdAPE), in addition to MAPE, RMSE, and MAD, as forecast accuracy measures to compare performance of SES and Croston's methods in intermittent demand forecasting. Noting that relative results were the same for all four measures, they reported only MAPEs.

2.2. Industrial Dataset and Partitioning

In the current study, we apply the four methods to an industrial dataset involving about 1500 items generally held in stock at a distribution center and a number of manufacturing plants of a firm operating in the professional electronics sector. The raw data consist of individual transactions as reported within the company's enterprise resource planning system, representing actual stock withdrawals. We initially aggregate the transactional data into weekly usage quantities, and further aggregate these usage quantities based on 13 four-week "months" in a calendar year. In this case, the monthly usage quantities do not constitute actual demand quantities, as the inventory on hand at the time of a transaction may not meet the required quantity. However, since demand is not traditionally tracked as well as actual usage in a transaction-based system, we treat monthly usage quantity as a surrogate measure of monthly demand.

This process yielded 66 months of "demand" data, which we broke down into initialization, calibration, and performance measurement blocks (as in Boylan, Syntetos, and Karakostas 2008) with each block consisting of 22 months in our study.

For each of the SES, Croston's and SBA methods, we selected α based upon the minimum MAPE attained in the calibration block for use as the smoothing constant in the performance block.

Given that the various SKUs represent end items, sub-assemblies, components, and spare parts that are used for building projects, retail sales, or servicing of professional electronic products, it is understandable that we found many of them to actually exhibit erratic or lumpy demand based on the earlier cited categorization scheme (Syntetos, Boylan, and Croston 2005).

In this paper, we report findings on a limited sample consisting of ten SKUs, with demand statistics presented in Table 1. These ten SKUs are not representative of our entire dataset. They were selected principally to demonstrate the approach we have taken, as well as to illustrate the results we have obtained thus far in both the empirical investigation of forecasting

performance and the empirical investigation of inventory control performance. The first nine SKUs are lumpy, while the last one is categorized as erratic (although its *ADI* of 1.27 is just below the cutoff of 1.32 as specified for lumpy demand).

Table 1: Sample of 10 SKUs

SKU #	1	2	3	4	5
Mean	10.97	1.44	0.71	9.74	2.82
Std Dev	13.30	2.23	1.76	13.82	7.19
CV^2	1.47	2.41	6.11	2.01	6.52
ADI	1.35	2.00	4.43	1.65	4.40
% of Zero Demand	27.3%	50.0%	80.3%	42.4%	80.3%
Demand Category	Lumpy	Lumpy	Lumpy	Lumpy	Lumpy

SKU #	6	7	8	9	10
Mean	3.47	4.85	9.27	3.03	6.37
Std Dev	5.01	6.76	19.35	8.03	7.85
CV^2	2.09	1.95	4.35	7.02	1.52
ADI	1.61	1.65	3.94	4.13	1.27
% of Zero Demand	37.9%	39.4%	77.3%	78.8%	24.2%
Demand Category	Lumpy	Lumpy	Lumpy	Lumpy	Erratic

2.3. Modeling of Intermittent/Lumpy Demand

A number of recent studies have referred to the use of a negative binomial distribution (NBD) to model the distribution of intermittent or lumpy demand items (e.g., Syntetos and Boylan 2006; Boylan, Syntetos, and Karakostas 2008; Syntetos, Babai, Dallery, and Teunter 2009). Syntetos and Boylan (2006) have argued that the NBD satisfies both theoretical and empirical criteria.

The NBD with parameters r and p, where $0 < p \le 1$ and $r > 0$, is given by the discrete density function (e.g., Mood, Graybill, and Boes 1974):

$$f(x;r,p) = \binom{r+x-1}{x} p^r (1-p)^x I_{\{0,1,2,\ldots\}}(x), \tag{2}$$

where the parameters p and r are a probability of "success" and a target number of successes, respectively. A realization x of the random variable X in this case represents a number of failures before the rth success is attained. The NBD has mean

$$\mu = E[X] = \frac{r(1-p)}{p} \tag{3}$$

and variance

$$\sigma^2 = V[X] = \frac{r(1-p)}{p^2}. \tag{4}$$

Since $V[X] = E[X]/p$, it follows that the variance of the NBD is greater than its mean.

When $r = 1$, the NBD reduces to a geometric (or Pascal) distribution with density

$$f(x;p) = p(1-p)^x I_{\{0,1,2,\ldots\}}(x). \tag{5}$$

To generate an NBD to approximate the distribution of a random variable with mean μ and variance σ^2, we simultaneously solve (3) and (4) and obtain:

$$\hat{p} = \frac{\mu}{\sigma^2} \tag{6}$$

and

$$\hat{r} = \frac{\mu^2}{\sigma^2 - \mu}, \tag{7}$$

as initial estimates of the NBD parameters (using the mean $\bar{x}$ and the variance s^2 of the 66-month actual demand time series as values of μ and σ^2, respectively). These expressions for $\hat{p}$ and $\hat{r}$, however, represent values in the set $\mathfrak{R}$ of real numbers, while the NBD parameter r is supposed to be integer-valued. In applying (6) and (7), we generally obtain a non-integer value of $\hat{r}$. Thus, in seeking to simulate the actual demand distributions, we have investigated the rounded up and rounded down values of $\hat{r}$ while adjusting the value of $\hat{p}$. However, we have found that, for many of the SKUs under study, it is not possible to obtain adjusted values of $\hat{r}$ and $\hat{p}$ that would lead to an NBD with mean, variance, CV^2, and *ADI* that are reasonably close to those of the actual demand distribution.

For a SKU with a proportion z of periods with zero demand is relatively high, we simulate the demand distribution by way of a two-stage process: a uniform distribution in stage 1 and an NBD in stage 2. Stage 1 involves an appropriately determined probability z_1 of zero demand, taking into consideration both z as well as the NBD in stage 2. We determine the mean and variance of the nonzero demands in the actual distribution and use these to calculate first approximations of the NBD parameters in stage 2. We test rounded up and down values of the parameter estimate $\hat{r}$ and refine the parameter estimate $\hat{p}$ as the values of mean, variance, CV^2, and *ADI* of the actual and simulated demand distributions are compared.

We attempted to simulate demand distributions of the SKUs under study applying our two-stage approach, performing 100 runs each consisting of 100 four-week "months", for a total of 10,000 months in each experiment. We used AnyLogic as our simulation platform, but with some code written in Java to handle mathematical modeling which could not be readily undertaken within the standard AnyLogic library. We accordingly selected z_1 for stage 1 and the parameters r and p of the NBD in stage 2 based on what appeared to yield the best combination of values of mean, variance, CV^2, and *ADI* of the simulated distribution in comparison with those of the actual distribution.

Generally, we found the percentages of zero demand and *ADI*s generated by our simulation approach to be reasonably close to those of the actual demand data. However, because the mean and standard

deviation of the simulated distribution using a combination of values of z_1, r and p are usually not fully consistent with those of the actual distribution, we tended to favor standard deviation and CV^2, which measure variability of demand sizes, in selecting the final parameter values.

Table 2 summarizes the parameters as determined and used for simulating the demand distributions of our sample of ten SKUs, applying our two-stage approach. In the currently reported sample of ten SKUs, the CV^2 values for the simulated distributions are greater than, but less than 120% of, CV^2 values for the actual data – except for SKUs 2 and 6 (with simulated values being 98% and 96% of actual CV^2). The percentages of zero demand and *ADI*s are generally close to the actual demand distribution values.

Table 2: Simulated Demand Distributions for Sample of 10 SKUs Using Two-Stage Approach

SKU #	1	2	3	4	5
Actual Demand Distribution					
Mean	10.97	1.44	0.71	9.74	2.82
Std Dev	13.30	2.23	1.76	13.82	7.19
CV^2	1.47	2.41	6.11	2.01	6.52
ADI	1.35	2.00	4.43	1.65	4.40
% of Zero Demand (z)	27.3%	50.0%	80.3%	42.4%	80.3%
Simulated Demand Distribution					
Stage 1: Uniform					
z_1	21.9%	25.4%	79.1%	38.5%	80.0%
Stage 2: NBD					
r	1	1	7	1	2
p	0.0691	0.3300	0.6667	0.0640	0.1230
Two-Stage Results					
Mean	10.26	1.47	0.72	9.07	2.82
Std Dev	13.30	2.23	1.77	13.79	7.24
CV^2	1.71	2.36	6.27	2.34	6.77
ADI	1.38	1.98	4.89	1.72	4.96
% of Zero Demand	27.3%	49.5%	79.9%	42.2%	80.3%

SKU #	6	7	8	9	10
Actual Demand Distribution					
Mean	3.47	4.85	9.27	3.03	6.37
Std Dev	5.01	6.76	19.35	8.03	7.85
CV^2	2.09	1.95	4.35	7.02	1.52
ADI	1.61	1.65	3.94	4.13	1.27
% of Zero Demand (z)	37.9%	39.4%	77.3%	78.8%	24.2%
Simulated Demand Distribution					
Stage 1: Uniform					
z_1	25.1%	30.6%	77.3%	77.1%	14.2%
Stage 2: NBD					
r	1	1	5	1	1
p	0.1711	0.1263	0.1085	0.0740	0.1171
Two-Stage Results					
Mean	3.55	4.70	9.09	2.87	6.45
Std Dev	4.98	6.80	19.35	8.08	7.96
CV^2	2.01	2.14	4.80	8.16	1.53
ADI	1.61	1.66	4.38	4.62	1.32
% of Zero Demand	37.7%	39.7%	77.1%	78.8%	24.4%

3. FORECASTING PERFORMANCE

3.1. Forecast Accuracy: Performance Block

The exponential smoothing constant α selected from among the candidate values (0.05, 0.10, 0.15, or 0.20) for each of the SES and SBA methods, based upon the minimum MAPE in the calibration block, are shown in Table 3. When accordingly applying SMA13, SES, and SBA to actual demand data in the performance block, the resulting error statistics are likewise reported in the same table. There does not appear to be a method that exhibits superior performance across the ten SKUs.

Table 3: Error Statistics when Applying Various Methods to the Performance Block

SKU #	1	2	3	4	5
Smoothing Constants Selected in Calibration Block					
SES	0.20	0.20	0.20	0.10	0.20
SBA	0.05	0.05	0.05	0.10	0.05
MAPE					
SMA13	90.36%	160.84%	150.43%	153.09%	162.60%
SES	85.89%	157.94%	163.77%	144.16%	151.70%
SBA	82.83%	196.25%	122.93%	140.08%	169.21%
Best MAPE	SBA	SES	SBA	SBA	SES
MAD					
SMA13	9.570	0.804	1.846	10.647	4.287
SES	9.097	0.790	2.010	10.025	3.999
SBA	8.772	0.981	1.509	9.742	4.461
Best MAD	SBA	SES	SBA	SBA	SES
RMSE					
SMA13	11.486	0.946	2.665	12.026	5.792
SES	10.925	0.876	2.752	11.423	5.702
SBA	10.723	1.076	2.681	10.914	5.683
Best RMSE	SBA	SES	SMA13	SBA	SBA

SKU #	6	7	8	9	10
Smoothing Constants Selected in Calibration Block					
SES	0.20	0.05	0.05	0.15	0.20
SBA	0.05	0.05	0.05	0.05	0.20
MAPE					
SMA13	102.62%	85.84%	177.91%	284.23%	98.83%
SES	110.62%	89.14%	191.34%	261.61%	98.25%
SBA	122.19%	90.04%	201.22%	255.94%	100.45%
Best MAPE	SMA13	SMA13	SMA13	SBA	SES
MAD					
SMA13	2.332	6.594	9.490	5.168	5.301
SES	2.514	6.848	10.176	4.756	5.270
SBA	2.777	6.795	9.906	4.653	5.388
Best MAD	SMA13	SMA13	SMA13	SBA	SES
RMSE					
SMA13	3.312	9.064	15.419	6.718	5.977
SES	3.409	9.488	14.924	6.293	6.063
SBA	3.245	9.797	14.956	6.234	6.156
Best RMSE	SBA	SMA13	SES	SBA	SMA13

3.2. Forecast Accuracy: Simulated Demand

When applying the methods to the simulated demand distributions, however, we see in Table 4 that the SBA method outperforms SMA13 and SES overall across the three error statistics in nine out of the ten SKUs. This suggests the overall superiority of SBA over a sufficiently longer time frame.

4. INVENTORY CONTROL PERFORMANCE

Traditionally, demand forecasting and inventory control have been treated independently of each other (Tiacci and Saetta 2009; Syntetos, Babai, Dallery, and Teunter 2009). However, demand forecasting performance, as assessed using standard statistical measures of accuracy may not necessarily translate into inventory systems efficiency (Syntetos, Nikolopoulos, and Boylan 2010).

Table 4: Error Statistics when Applying Various Methods to Simulated Demand Distributions

SKU #	1	2	3	4	5
Smoothing Constants Selected in Calibration Block					
SES	0.20	0.20	0.20	0.10	0.20
SBA	0.05	0.05	0.05	0.10	0.05
MAPE					
SMA13	97.67%	114.02%	165.42%	116.67%	163.50%
SES	99.39%	115.98%	163.89%	115.58%	163.96%
SBA	93.82%	111.92%	154.96%	113.61%	163.54%
Best MAPE	SBA	SBA	SBA	SBA	SMA13
MAD					
SMA13	10.118	1.681	1.203	10.551	4.576
SES	10.299	1.710	1.192	10.451	4.604
SBA	9.716	1.646	1.124	10.268	4.578
Best MAD	SBA	SBA	SBA	SBA	SMA13
RMSE					
SMA13	13.899	2.319	1.844	14.499	7.602
SES	14.136	2.363	1.825	14.293	7.643
SBA	13.544	2.259	1.800	14.087	7.264
Best RMSE	SBA	SBA	SBA	SBA	SBA

SKU #	6	7	8	9	10
Smoothing Constants Selected in Calibration Block					
SES	0.20	0.05	0.05	0.15	0.20
SBA	0.05	0.05	0.05	0.05	0.20
MAPE					
SMA13	104.36%	108.09%	156.34%	169.37%	93.82%
SES	106.08%	104.73%	156.68%	168.40%	94.92%
SBA	101.60%	102.67%	156.54%	163.98%	91.49%
Best MAPE	SBA	SBA	SMA13	SBA	SBA
MAD					
SMA13	3.728	5.106	14.209	4.8362	6.046
SES	3.791	4.944	14.199	4.8097	6.118
SBA	3.624	4.847	14.121	4.6655	5.897
Best MAD	SBA	SBA	SBA	SBA	SBA
RMSE					
SMA13	5.212	7.113	20.108	8.4369	8.262
SES	5.285	6.940	19.665	8.4355	8.383
SBA	5.046	6.935	19.529	8.1122	8.248
Best RMSE	SBA	SBA	SBA	SBA	SBA

A periodic review inventory control system has been recommended in dealing with intermittent demand (e.g., Sani and Kingsman 1997; Syntetos, Babai, Dallery, and Teunter 2009). Some recent intermittent demand forecasting studies (e.g., Eaves and Kingsman 2004; Syntetos and Boylan 2006; Syntetos, Babai, Dallery, and Teunter 2009; Syntetos, Nikolopoulos, Boylan, Fildes, and Goodwin, 2009; Syntetos, Nikolopoulos, and Boylan 2010; Teunter, Syntetos, and Babai 2010) that evaluate both forecasting and inventory control performance have used the order-up-to (*T*,*S*) periodic review system, where *T* and *S* denote the review period and the base stock (or 'order-up-to' level), respectively.

We assume in the current study a (*T*,*S*) inventory control system with full backordering. Inventory is reviewed on a monthly basis ($T = 1$). For most SKUs, the reorder lead time is more or less one month; we thus set $L = 1$. Let I_t and B_t, respectively, denote the on-hand inventory and backlog at the time of review *t*. The literature on inventory control suggests a safety stock component in order-up-to levels to compensate for uncertainty in demand during the "protection interval" $T+L$. For each demand series, we calculated the standard deviation, s_{cal}, of monthly usage quantities during the calibration block. We apply a "safety factor" *k* as a multiplier of s_{cal} to obtain a safety stock level of $k \cdot s_{cal}$. This safety stock determination is more or less similar to the $z \cdot \sqrt{T+L} \cdot \sigma_d$ suggested when daily demand during the protection interval is assumed to be identically and independently normally distributed with standard deviation σ_d (e.g., Silver, Pyke, and Peterson 1998). With the safety stock component, the replenishment quantity to order is

$$Q_t = \sum_{t+1}^{t+T+L} F_j + k \cdot s_{cal} - I_t + B_t . \quad (8)$$

We simulate performance of the (*T*,*S*) inventory control system based on the two most commonly specified service level criteria (Silver, Pyke, and Peterson 1998). One service criterion is a target average *probability of no stockout* per review period. The other service measure is a target average fraction of demand to be satisfied from stock on hand, also called a *fill rate* (FR) which has considerably more appeal for practitioners.

With a 95% target FR, averages of inventory on hand arising in the simulation experiments from the use of SMA13, SES, and SBA are reported in Table 5. We find that, for all of the ten SKUs, the average inventory on hand is consistently lowest when SBA is the forecasting method applied.

Table 5: Average Inventory on Hand for a 95% Fill Rate

SKU #	1	2	3	4	5
SMA13	34.77	6.24	6.33	39.50	29.03
SES	35.47	6.40	6.16	38.73	28.14
SBA	33.61	6.07	6.15	37.56	27.54

SKU #	6	7	8	9	10
SMA13	13.51	18.32	59.00	37.85	20.02
SES	13.87	17.72	57.01	37.74	20.34
SBA	13.07	17.69	56.48	37.39	19.86

For the same 95% target FR, the means of total backlogs over 100 months (as averaged over 100 replications) are reported in Table 6. We find these 100-month means to be roughly equal across the forecasting methods.

Under a 95% target probability of no stockout, SBA likewise generally outperforms SMA13 and SES with respect to average inventory on hand. This is exhibited in Figure 1 where all indices (with SBA as

base) are above 100, with an index of 98.3 when using SES for SKU 8 being the only exception.

Table 6: Mean 100-Month Backlogs for a 95% Fill Rate

SKU #	1	2	3	4	5
SMA13	52.01	7.45	3.71	45.60	14.45
SES	52.10	7.58	3.80	45.78	14.36
SBA	52.12	7.58	3.85	45.79	14.43

SKU #	6	7	8	9	10
SMA13	17.82	24.09	45.43	15.22	31.84
SES	17.90	24.51	45.57	15.21	32.12
SBA	18.07	24.29	46.18	15.30	32.13

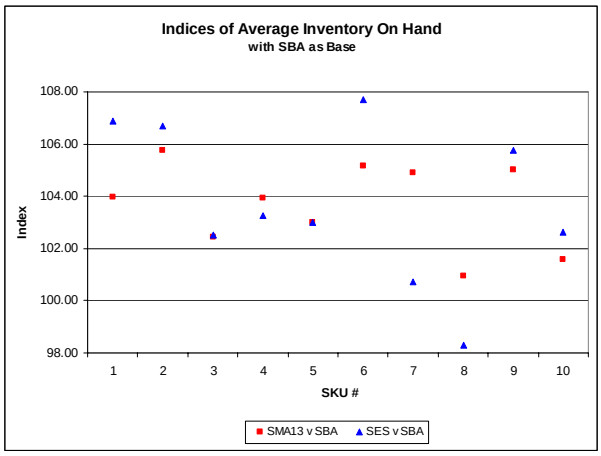


Figure 1: Average On-Hand Inventory Indices, with SBA as Base, for a 95% Probability of No Stockout

Likewise, with a 95% target probability of no stockout, the means of total backlogs over 100 months are reported in Table 7. None of the means arising from the use of SBA appears to be substantially higher than the mean corresponding to the use of either SMA13 or SES.

Table 7: Mean 100-Month Backlogs for a 95% Probability of No Stockout

SKU #	1	2	3	4	5
SMA13	75.49	14.72	12.21	80.99	47.66
SES	74.96	15.03	12.12	78.53	49.63
SBA	75.48	14.96	12.45	78.46	48.21

SKU #	6	7	8	9	10
SMA13	30.35	38.97	85.51	68.71	44.94
SES	30.40	39.25	84.69	68.76	45.62
SBA	30.76	39.51	79.06	70.06	45.49

5. CONCLUSION AND FURTHER WORK

We have devised a two-stage approach, involving uniform and negative binomial distributions, which allows modeling of the actual demand distribution, even when it is lumpy. Our work departs from earlier studies which have merely argued that the NBD satisfies both theoretical and empirical criteria, and accordingly assumed that an NBD adequately captures the behavior of intermittent demand. We believe that the simulated demand distributions arising from our two-stage modeling approach would more closely approximate the actual demand distributions of the SKUs under consideration.

In empirically investigating the forecasting methods on the performance block (the final 22 months of the 66-month actual distribution) using three traditional statistical measures of forecast accuracy, we found none of the methods under consideration to be consistently superior to the others. However, when the methods are tested over considerably more time periods (100 replications of 100 months using our two-stage approach), SBA is found to be the best performing method overall in terms of statistical accuracy.

We then proceeded to apply the demand estimates arising from the different forecasting methods, on the basis of the simulated demand distribution generated for a given SKU. We assumed a (*T*,*S*) periodic review inventory control system with full backordering, with a one-month review period and a one-month replenishment leadtime. Using either a target FR or a target probability of no stockout as customer service level criterion, we have found SBA to yield the lowest average levels of inventory on hand in almost all cases. At the same time, the frequency of backorders under SBA is comparable to those using the other forecasting methods.

The observations reported here are based on a very limited sample of ten SKUs from the industrial dataset. At the time of the conference, we expect to report more robust findings – based upon our analysis of forecast accuracy and stock control performance over a larger number of SKUs as well as other customer service levels.

REFERENCES

Altay, N., Rudisill, F., Litteral, L.A., 2008. Adapting Wright's modification of Holt's method to forecasting intermittent demand. *International Journal of Production Economics*, 111 (2), 389-408.

Babai, M.Z., Syntetos, A.A., and Teunter, R., 2010. On the empirical performance of (*T*,*s*,*S*) heuristics. *European Journal of Operational Research*, 202 (2), 466-472.

Boylan, J.E., Syntetos, A.A., and Karakostas, G.C., 2008. Classification for forecasting and stock control: a case study. *Journal of the Operational Research Society*, 59 (4), 473-481.

Croston, J.D., 1972. Forecasting and stock control for intermittent demands. *Operational Research Quarterly*, 23 (3), 289-304.

Eaves, A.H.C. and Kingsman, B.G., 2004. Forecasting for the ordering and stock-holding of spare parts. *Journal of the Operational Research Society*, 55 (4), 431-437.

Ferrari, E., Pareschi, A., Regattieri A., and Persona, A., 2006. Statistical management and modeling for demand of spare parts. In: H. Pham, ed. *Springer Handbook of Engineering Statistics*. Würzburg, Germany: Springer-Verlag, 905-929.

Gilliland, M., 2002. Is forecasting a waste of time? *Supply Chain Management Review*, 6 (4), 16-23.

Gutierrez, R.S., Solis, A.O., and Mukhopadhyay, S., 2008. Lumpy demand forecasting using neural networks. *International Journal of Production Economics*, 111 (2), 409-420.

Johnston, F.R. and Boylan, J.E., 1996. Forecasting for items with intermittent demand. *Journal of the Operational Research Society*, 47 (1), 113-121.

Mood, A.M., Graybill, F.A., and Boes, D.C., 1974. *Introduction to the Theory of Statistics*, 3rd ed. New York, USA: Mc-Graw-Hill.

Mukhopadhyay, S., Solis, A.O., and Gutierrez, R.S., 2011. The Accuracy of Non-traditional versus Traditional Methods of Forecasting Lumpy Demand. *Journal of Forecasting* (forthcoming; DOI: 10.1002/for.1242).

Sani, B. and Kingsman, B.G., 1997. Selecting the best periodic inventory control and demand forecasting methods for low demand items. *Journal of the Operational Research Society*, 48 (7), 700-713.

Schultz, C.R., 1987. Forecasting and inventory control for sporadic demand under periodic review. *Journal of the Operational Research Society*, 38 (5), 453-458.

Silver, E.A., Pyke, D.F., and Peterson, R., 1998. *Inventory Management and Production Planning and Scheduling*. New York, USA: John Wiley & Sons.

Syntetos, A.A. and Boylan, J.E., 2001. On the bias of intermittent demand estimates. *International Journal of Production Economics*, 71 (1-3), 457-466.

Syntetos, A.A. and Boylan, J.E., 2005. The accuracy of intermittent demand estimates, *International Journal of Forecasting*, 21 (2), 303-314.

Syntetos, A.A., Boylan, J.E., and Croston, J.D., 2005. On the categorization of demand patterns. *Journal of the Operational Research Society*, 56 (5), 495-503.

Syntetos, A.A. and Boylan, J.E., 2006. On the stock control performance of intermittent demand estimators. *International Journal of Production Economics*, 103 (1), 36-47.

Syntetos, A.A., Nikolopoulos, K., Boylan, J.E., Fildes, R., and Goodwin, P., 2009. The effects of integrating management judgement into intermittent demand forecasts. *International Journal of Production Economics*, 118 (1), 72-81.

Syntetos, A.A., Babai, M.Z., Dallery, Y., and Teunter, R., 2009. Periodic control of intermittent demand items: theory and empirical analysis. *Journal of the Operational Research Society*, 60 (5), 611-618.

Syntetos, A.A., Nikolopoulos, K., and Boylan, J.E., 2010. Judging the judges through accuracy-implication metrics: The case of inventory forecasting. *International Journal of Forecasting*, 26 (1), 134-143.

Teunter, R.H., Syntetos, A.A., and Babai, M.Z., 2010. Determining order-up-to levels under periodic review for compound binomial (intermittent) demand. *European Journal of Operational Research*, 203 (3), 619-624.

Tiacci, L. and Saetta, S., 2009. An approach to evaluate the impact of interaction between demand forecasting method and stock control policy on the inventory system performances. *International Journal of Production Economics*, 118 (1), 63-71.

Willemain, T.R., Smart, C.N., Shockor, J.H., and DeSautels, P.A., 1994. Forecasting intermittent demand in manufacturing: a comparative evaluation of Croston's method. *International Journal of Forecasting*, 10 (4), 529-538.

AUTHORS BIOGRAPHY

Adriano O. Solis is an Associate Professor of Logistics Management and Management Science at York University, Canada. After receiving BS, MS and MBA degrees from the University of the Philippines, he joined the Philippine operations of Philips Electronics where he became a Vice-President and Division Manager. He went on to obtain a PhD degree in Management Science from the University of Alabama. He was previously Associate Professor of Operations and Supply Chain Management at the University of Texas at El Paso. He has been a Visiting Professor in the Department of Mechanical, Energy, and Management Engineering, at the University of Calabria, Italy. He has published in *European Journal of Operational Research*, *International Journal of Simulation and Process Modelling*, *Information Systems Management*, *International Journal of Production Economics*, *Journal of Forecasting*, *Computers & Operations Research*, and *Journal of the Operational Research Society*, among others.

Letizia Nicoletti is currently PhD student at MSC-LES, Mechanical Department of University of Calabria. Her research interests include inventory management Modeling and Simulation. She is also applying Modeling & Simulation based approaches for training in complex systems, specifically marine ports and container terminals. She also actively supports the local organization of many international conferences in the modeling and simulation area.

Somnath Mukhopadhyay, an Associate Professor in the Information and Decision Sciences Department at the University of Texas at El Paso, has published in journals like *Decision Sciences*, *INFORMS Journal on Computing*, *Journal of Forecasting*, *Communications of the AIS*, *IEEE Transactions on Neural Networks*, *Neural Networks*, *Neural Computation*, *International Journal of Production Economics*, and *Journal of World Business*. He received MS and PhD degrees in Management Science from Arizona State University. He was a visiting research assistant in the parallel distributed processing

group of Stanford University. He has over 15 years of industry experience in building and implementing mathematical models.

Laura Agosteo, **Antonio Delfino**, and **Mirko Sartiano** are students in the *Laurea Magistrale* (master's degree) program in Management Engineering at the University of Calabria's Department of Mechanical, Energy, and Management Engineering. They obtained their first degrees in Management Engineering in 2010. They were recently exchange graduate students/visiting research scholars at the School of Administrative Studies, York University, Canada.

A MODEL OF A BIOFILTER FOR MECHANICAL PULPING WASTE-WATER TREATMENT

Stefano SAETTA[(a)], Lorenzo TIACCI[(b)], Markku TAPOLA[(c)],Sara HIHNALA[(d)]

[(a)(b)]Dipartimento di Ingegneria Industriale Università degli Studi di Perugia
[(c)]Meehanite Technology, Finland

[(a)]stefano.saetta@unipg.it, [(b)]lorenzo.tiacci@unipg.it, [(c)]Markku.Tapola@meehanite.org, [(c)]Sara.Hihnala@meehanite.org

ABSTRACT

VOC emissions are released from mechanical pulping process and can be abated by different cleaning technologies. VOC emissions origin both from grinding process and waste water treatment. In Europe the estimated annual total NMVOC emissions from mechanical pulping process is approx. 7.000 tons.

Biofiltering is one of the most promising abatement techniques available. This because its reduced CO_2 emissions and its long life.

In the paper a pilot application to a mechanical pulping waste water system is considered.

Mechanical pulping emissions are Volatile Organic Compounds, VOCs. Today VOCs are a main issue in environmental control.

The role of Biofiltering in VOC abatement is pointed out in the paper.

In the paper a model and a pilot implementation of the Biofilter are introduced. The results of the biofilter used in waste-water treatments are then shown.

Keywords: VOC emissions, Mechanical Pulping, VOCLESS

1. INTRODUCTION

Volatile Organic Compound (VOC) emissions are limited by international standards such as the EU VOC solvents emissions directive 1999/13/EC or Section 183(e) of the Clean Air Act in the United States).

One industry with a VOC emission problem is the pulp and paper industry that is not yet regulated by VOC emission limits.

In order to reduce VOC emissions, there are several solutions. One of these is the use of Biofilters.

Such a solution is relatively new and is based on the oxidations of VOCs by a population of microorganisms.

Since its first application in the 1970s (Reinluft), there has been a growing interest in Biofilter applications and their modeling and simulation.

This is because Biofilters reduce CO_2 emissions and they have low operating costs and a long working life. They are simple to manage and they, unlike incinerators, are $C0_2$ free.

Modeling and simulation have been considered in many environmental sectors (Bruzzone et al., 2009a) (Bruzzone et al. 2009b) and also in the field of biofilters. In Jacob et al. (1996), a dynamic simulation of biofilters is considered. The aim is to build and validate an analytical model based on the balance equations, from which a partial differential and algebraic equations are derived. The model is validated for the processes of denitritation and denitrification using data from a pilot plant. The model is solved using numerical methods by applying the methods of lines and an approximated method to reduce model complexity. The results obtained are good, but required more deeper analysis to be used in controlling biofilters.

In Mohseni and Allen (2000), a mathematical model for a mixture of hydrophilic and hydrophobic VOCs is considered. This mixture refers clearly to the pulp and paper industry, where such emissions are typical. The two main processes of VOC diffusion through a biofilm and their degradation are modeled.
The model parameters are determined by experimental data from a pilot plant.

The results of the model are then compared to data in the literature giving better results.

Noteworthy is the fact that the presence of hydrophilic VOCs can affect the abatement of hydrophobic VOCs. This model could be useful for use in the pulp and paper industry and also in mechanical pulping. However, analysis and modeling has been limited to only 2 substances and for softwood only.

In Easter et al. (2005), the use of biotechnologies in the abatement of VOCs and odours from wastewater is considered. In particular, the collection system is also considered because of its role in odour diffusion.

In the paper, design parameters such as Empty Bed Residence Time, EBRT:

$$EBRT = \frac{AD}{Q}\ [s],$$

where A is biofilter surfaces m^2, D is biofilter bed depth m, Q is odor flow m^3/s.

Also, volumetric Loading Rate is considered:

$$LR = \frac{Q}{A}\left[\frac{m}{s}\right]$$

Other parameters such as temperature, ph and moisture control, airflow distribution, leachate and drenage , and media selection are considered.

VOC concentration is not treated because in this study airflow composition is known and then parameters are computed.

Results are considered for different kinds of media and for both biofilter and biotowers. It is important to know the composition of the air stream in input of the biofilters, in the design process.

In the paper, a model and a pilot implementation of the biofilter are shown.

The results of the biofilter used in waste-water treatments are then considered.

Abatement emissions are considered a concern in mechanical pulping waste-water VOCs.

Biofilter tests are part of a LIFE+ project, VOCless Pulping Waste Waters (LIFE09 ENV/FI/000568).

In the following paper, after a brief description of the LIFE project, the biofilters developed in the project are shown.

Then results of a field test at a Mechanical Pulp plant are also given.

2. THE VOCLESS PULPING WASTE WATERS

The VOCless Pulping Waste Waters project is concerned with VOC emissions and odour problem in Mechanical Pulping. In particular, this project considers the VOC and odour emissions from mechanical pulping waste-waters.

The project starts from the results of a previous VOCless project where all VOC emissions from mechanical pulping in general were considered. During this project, VOC measurements showed that emissions from waste-waters accounted for the same amount as that emitted directly from the pulping process.

Therefore, the aim of this project is find the best technological solution to minimize such an amount of VOC emissions from waste-waters.

In order to do this, the pilot testing of several solutions was carried out. The pilot biofilters were designed, implemented, and tested for this project.

The VOCless project faces the increasing problem of VOC emissions reductions demanded by International Legislation. VOCs are known to be the cause of smog formation, odour problems, and other harmful effects.

The interesting aspect of the considered project is that in mechanical pulping VOCs are natural, so they are not controlled by VOC legislation (that applies only to anthropic VOCs). Nevertheless, mechanical pulping VOCs are emissions from industrial production, so their emissions should be considered also in this case.
Therefore, the only solution, since VOCs are natural is an end-of-pipe technology.

3. DESCRIPTION OF THE BIOFILTER

In the following, the biofilters considered are presented. The following part of the paragraph is part of the Meehanite report from the VOCless Waste Waters project. Biofilters lend themselves to all waste gas cleaning applications involving air pollutants that are readily biodegradable. Biodegradation of the pollutants is accomplished by micro-organisms colonising on a solid support media. The typical support media employed are chopped wood and wood bark, composts or other origins, fibrous peat and heather that may be combined with one another or other structure-giving materials. Moreover, inert materials exhibiting large inner surface areas (lava, expanded clay) and hence, the ability to support a large population of microorganisms are employed as support media. All these materials are normally arranged as randomly packed beds through which the waste gas flows. As the waste gas passes through the bed of media, the pollutants are sorbed onto the surface of the filter media where they are degraded by micro-organisms. For optimum growth and metabolic activity, the micro-organisms rely on defined environmental conditions (moisture, pH, oxygen content, temperature, nutrients, etc.) which must be controlled within narrow limits. As micro-organisms are affected by changes in their environment, they may require some time for acclimation before developing their full activity after the biofilter start up or changes in the operating conditions. Transport of the pollutants from the gas phase through the aqueous phase surrounding the filter media and from there to the bacterial cells involves the following individual steps:

- Mass transfer to gas/liquid interface
- Absorption into the liquid phase
- Mass transfer through the liquid phase to the bacterial cell
- Sorption and degradation by the cell

Factors governing the rate of reaction (degradation rate) include:

- Concentration and types of waste gas component
- Type, number, and activity of the micro-organisms colonising on the respective filter media
- Temperature
- Moisture content of the waste gas and the filter media
- pH value
- Solubility of the waste gas components
- The type and concentration of any reaction products accumulating in the filter media.

The net conversion of pollutants in the filter bed is determined by the rate of reaction, the residence time of the gas in the biofilter, and the concentration of the pollutant in the crude gas. If the biological reactions proceed relatively fast and the pollutants to be removed are sparingly soluble, transport processes of the reactants from the gas phase to the inner surface of the filter media may become rate-limiting.

The filter media may also serve as a source of nutrients and nutrient salts for the microorganisms. In applications involving waste gases poor in nutrients, intermittent operating conditions and waste gases with high organics loads, the micro-organisms rely on additional nutrient supply. This can be partly accomplished via the nutrient salts present in the filter media. However, in some cases subsequent fertilization of the filter media may be needed. These parameters of the biofilters were observed and some additional nutrients were added.

Another important function of the filter media is to provide the micro-organisms with sufficient moisture. Otherwise the micro-organisms would have to be continuously sprayed with water. The filter media has an optimum moisture level at the time it is being placed in the biofilter. In order to preclude rapid bed dry-out, the waste gas to be treated must be humidified to the maximum feasible level before being admitted to the biofilter. After pre-conditioning in a humidifier, the waste gas is normally about 95% water-saturated. This means that it will take up the remaining 5% as it passes through the filter bed causing the latter to dry out. Filter bed dry-out can be counteracted by providing for additional irrigation of the filter media. A particular problem encountered with organic filter media is that it develops increasing hydrophobic effects as the moisture content decreases. This means that material that has become overly dry is no longer amenable to humidification. In most cases, it has to be removed. Overly wet filter media, on the other hand, will lead to water-logging. Under these conditions, the pores of the filter media are filled with water and thus blocked for the gas flow. This does not only affect the removal efficiency but also leads to oxygen depletion and hence insufficient oxygen supply for the microorganisms. While part of the microbial population can adapt its metabolism to anaerobic conditions, this results in metabolic end products which are similar to those formed by decaying organic materials and have a very unpleasant smell. This effect can be reversed by drying the filter media. Because of these facts the humidity of the biofilters was carefully observed by an online data logger and regular pilot plant check-up visits.

3.1. Filter characteristics

For the filter media to accomplish its manifold functions, it must exhibit specific material properties. It may well be possible that the specific requirements rule out one another. In such cases, an optimum trade-off tailored to the specific application will have to be made in media selection: A uniform structure favours uniform flow distribution and hence uniform contacting conditions. This, in turn, improves the degradation efficiency of the biofilter and minimizes the pressure drop. A sufficient pore volume reduces the pressure drop across the bed of media and hence the energy consumption. In addition, it improves the drainage effect of the media and prevents water logging. Depending on the type and condition of the filter media, the void volume may range between 20% and 80%. A large inner surface of the bed of media can be achieved by selecting either fine-sized media or media with a large pore volume. The large inner surface is required to ensure good sorption performance of the biofilter on the one hand and create the conditions necessary for a high microbial density on the other. Basically, all surfaces of the media particles, except for extremely small pores, are suitable to support microbial growth. Here, it should be noted that a fine-sized material causes a higher pressure drop than a coarse material. As the filter media inevitably undergoes decomposition and becomes progressively finer in the course of biofilter operation, the use of fine-sized media from the beginning results in a shorter service life.

3.2. Bio filter control and maintenance

Biochemical measures mainly relate to the addition of supplemental nutrients to the filter media and buffering undesirable changes in pH. The objectives of both measures are identical. They must reach the filter media surface, influence it and have a long-term effect in order to minimize treatment intervals. For this reason, granular additives that have a depot effect are preferred for both measures. For pH adjustment, water-soluble granular and powdered lime was added to the filter media.

Ph and moisture were monitored regularly during the pilot site visits by manual measurements. Additional powdered lime dissolved in water was added periodically in both bio-filters.

Temperature monitoring of the filter material and raw gas was done by online measurement at several points with thermal couples. A channel heater was assembled to control the raw-gas temperature.

3.3. Description of the pilot plant

Two small-scale biofilter pilot plants were built up for the pilot tests to be carried out at the aerobic waste-water treatment plant at the Stora Enso Anjala pulp mill. Both pilot plants were of the same size but the content of the filter material varied between them. The Meehanite pilot biofilter plant is in the scale of 1 to 500 of the full-size application. The first site for the pilot biofilter was the air flow of the cooling tower with an airflow of some 10 m^3/s. The air flow of biofilter pilot plant is designed for an air flow rate of 70 - 140 m^3/h**.** A schematic drawing of a the pilot biofilter is presented in Figure 1.

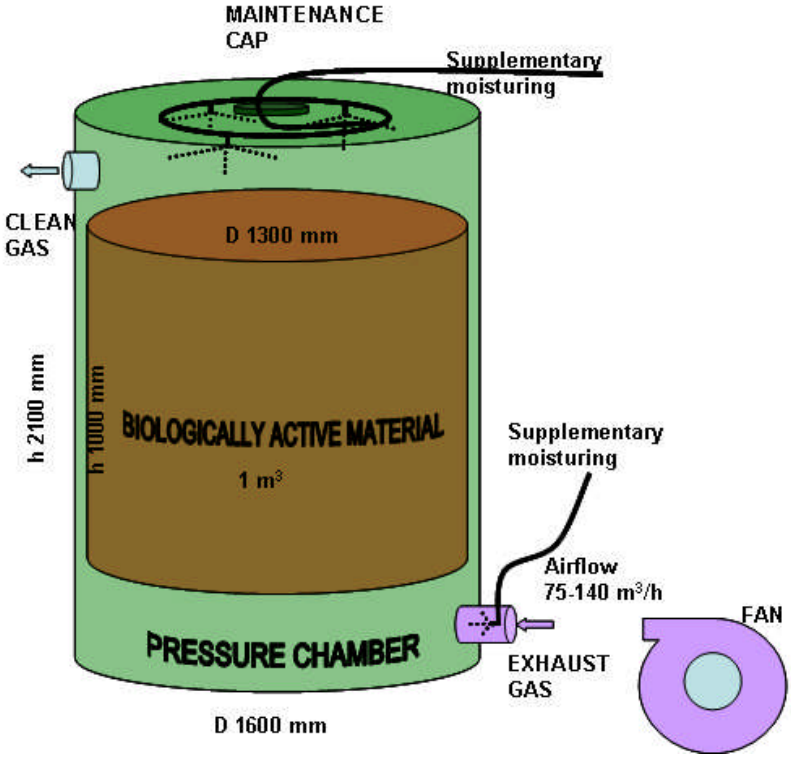


Figure 1. Schematic drawing of the principle features of the compact filter as a piloting biofilter

The exterior dimensions of the cylindrical shaped pilot biofilter are as follows: the diameter 1.6 m, and the height 2.1 m. The air flow direction is vertical from bottom to top. The surface area of biofilter bed is 1.3 m^2 and the bed thickness can vary from 1 000 to 1 500 mm. The biofilter contains a pre-filter stage as well. This acts as a humidifying and normalizing scrubber, and avoids unpleasant pH peaks that originate from the waste-waters of the refinery and grinding plants. The waste-water itself is neutralized already in the pump station, and the filter material pH is controlled regularly and modified when needed. The inlet exhaust gas was preheated when required by the channel heater (25-30 °C). A photograph of the Meehanite biofilter pilot plant is presented in Figure 2.

Figure 2. The Meehanite biofilter pilot plant.

3.4. Experimental Results

The 2 biofilters were used for the abatement of emissions from the cooling tower and from the micro bed bioreactors. Data from the two points of emissions are shown in Table1.

Table 1 . Parameters of the biofilter pilot plants.

Pilot site	**Cooling tower** 5.7.-19.7.2011.		**Microbed bioreactors** 20.7.-16.8.2011	
Biofilter 1 / Biofilter 2	**Bio 1**	**Bio 2**	**Bio 1**	**Bio 2**
Air flow temperature (°C) before biofilter	30-40		30-40	
Air flow temp. °C after biofilter	28-33	26-34	28-34	20-34
Biofilter bed temp. (°C)	25-32	26-32	26-34	26-34
Humidity (%)	99.9		99.9	
Pressure (Pa) approx.	2300		200-2500	
Air flow (m^3/h)	20		66	

Table 2: Biofilters performances at the waste-waters treatment plant

Site	Unit	VOC-concentration			Cleaning rate
		date	before	after	
			mg/Nm³	*mg/Nm³*	*%*
Cooling towers	Biofilter 1	6.7.2011	3.7	4.8	*
		19.7.2011	2.5	0.63	75
	Biofilter 2	5.7.2011	3.3	0.70	79
		19.7.2011	3,.4	0.66	81
MBR-1	Biofilter 1	21.7.2011	11.1	3.5	68
		2.8.2011	18.1	4.7	74
		17.8.2011	29.9	13.6	55
		30.8.2011	25.0	13.9	*
		31.8.2011	12.9	2.9	78
	Biofilter 2	21.7.2011	12.1	4.4	64
		2.8.2011	10.0	2.8	72
		17.8.2011	22.0	13.0	*
		31.8.2011	11.7	6.6	*

* malfunction

CONCLUSIONS

Two pilot biofilters were tested at a aerobic waste water treatment Mill in Finland. Pre-measurements were made and the pilot biofilter plants were placed at the cooling tower and the MBBR basin where the highest VOC concentrations appeared. Two small-size biofilter pilot plants were constructed. Different compositions and moisture of bacteria carrier material were tested. The pilot tests began in May and were completed at the end of August.

At the sites, the incoming VOC concentrations were different being higher at the MBBR site. An unexpected concentration of methane for aerobic biological wastewater treatment was detected. That presented a challenge in interpreting the measurement results.

The cleaning efficiency rate of the biofilters was on average 80% at the cooling tower, and the cleaning efficiency rate of the biofilters at the MBBR basin was about 70%. Cleaning rates were better with higher initial concentrations and averaged towards 70 to 75%. The cleaning rates probably fluctuated because of the varying composition and texture of the bacteria carrier material in the biofilters.

Overall, the biofilter tests showed good performance and sustainable working. It is obvious that the biobed material must be prepared and matured in good time prior to the actual exposure for a continuous run. The bioreactor material should always be reserved for the loading of an extra filter batch in the case of a malfunction in the process.

The test results showed that no extra pre-filter e.g. neutralization of the inlet gas is needed for the biofilters. The maximum specific air flow of 100 m³/h per square metre of filter area cannot be exceeded and a minimum depth of one metre of filter material is needed to reach an acceptable 80% cleaning efficiency.

As a future development, a model could be tested in other waste-waters treatment plants and it could optimised by using simulations.

ACKNOWLEDGMENTS

The research mentioned in this paper is partially funded by the European Union, within Life+ Program, Project: "Life09 ENV/FI/000568, abatement of VOC load from waste water treatment in mechanical pulping". The authors thank all the partners in the project: Meehanite Technology Oy Finland, Università di Perugia, Italy, AX Consulting Ltd. Finland, Formia Emissions Control Oy, Finland, Stora Enso Publication Papers Oy Ltd Anjala Mill, Finland, and Desinfinator Oy Ltd. Finland.

.

REFERENCES

Bruzzone, A.G., Tremori A., Massei M., Tarone, F. 2009a, Modeling Green Logistics, *AMS'09: Third Asia International Conference on Modelling & Simulation*, pp. 543-548, May 25-29, Bandung/Bali, Indonesia

Bruzzone A.G., Massei M., Tarone F., Longo F., 2009b, Environmental Sustainability: A Case Study In The Automotive Sector, Proceedings of INGIND17 Summer School, 15th-19th September, Porto Giardino, Monopoli, Italy

Easter C., Quigley C., Burrowesa P., Witherspoon J., Apgar D., 2005, Odor and air emissions control using biotechnology for both collection and wastewater treatment systems, *Chemical Engineering Journal 113 93–104*

Jacob, J., Pingaud, H., Le Lann, J.M., Bourrel, S., Babary, J.P., Capdeville, B., 1996, Dynamic simulation of biofilters, *Simulation Practice and Theory 4 335-348*

Mohseni M., Allen D. G., 2000, Biofiltration of mixtures of hydrophilic and hydrophobic volatile organic compounds, Chemical Engineering Science 55 1545-1558

Reinluft Environmental Engineering Company Ltd., www.reinluft.de [August, 19th, 2012]

Vocless Pulping Waste Waters, www.voclesspulping.com [August, 19th, 2012]

MODELS & INTERACTIVE SIMULATION FOR CIVIL MILITARY INTEROPERABILITY IN HUMANITARIAN AID AND CIVIL PROTECTION

Agostino Bruzzone[(a)], Alberto Tremori[(a)], Francesco Longo[(b)], Michele Turi[(c)], Giulio Franzinetti[(d)]

[(a)] DIME, University of Genoa, Italy - www.itim.unige.it
[(b)] MSC-LES, Mechanical Dept, University of Calabria, Italy – www.msc-les.org
[(c)] DIMS, University og Genoa, Italy – www.dims.ingegneria.unige.it/
[(d)] Lio Tech Ltd UK – www.liotech.co.uk

[(a)]agostino@itim.unige.it, alberto.tremori@simulationteam.com, [(b)]f.longo@unical.it
[(c)]michele.turi@simulationteam.com ,[(d)]giulio.franzinetti@liotech.eu

ABSTRACT

Modelling and Simulation based on innovative approaches and solutions could support training as well operational planning in humanitarian crisis management. Armed Forces could provide great expertise in Crisis Management to Civil Protection Agencies and, at the same time, need to be trained themselves to interoperate with civil organization in Humanitarian Support Operations. This paper describes on going researches devoted to develop innovative interoperable simulation models to enhance current capabilities in this field. Authors are focusing in particular on agent based simulation, human behavioral models and interoperability among different simulators to recreate complex crisis scenarios where the population represents the main critical target of any kind of Civil Protection operations. A specific attention is devoted to analyze simulation computational workload issues.

Keywords: Intelligent Agents Computer Generated Forces (IA_CGF), Human Behavioral Models, Interoperable Simulation Crisis Management, Simulation Computational Workload

1. INTRODUCTION

Our Armed Forces, based on current situation and international agreements, are more and more involved in Humanitarian Aid and Civil Protection operations for Natural and human-made Disasters in stabilization operations and to support and help population. Such involvement is focused, for every event, on providing personnel, means and vehicles and, in particular, to provide expertise in crisis management and coordination of "joint" coalitions (that is a further issue for Civil Military Interoperability, usually due to cultural differences).

Another critical issue in terms of military skills that Armed Forces could provide (with potentially huge benefits in the area of Humanitarian Aid and Civil Protection) is the personnel training and education at all the different hierarchical levels. In this area Armed Forces usually own consolidated methodologies and tools that could be brought to Civilian Agencies: Modelling & Simulation (M&S). M&S is a critical technology that could be used for Crisis Management training.

Nevertheless there are also several gaps that must be filled to provide a really breakthrough technology and an innovative methodology for training Civil Protection by using M&S.

In particular the research described in this paper focuses on how M&S, starting from the military experience in training, should be improved to be successfully used in Humanitarian Aid and Civil Protection considering some of the most critical factors that have to be improved in current simulations (Amico et al. 2000):

- Human Behaviour Modelling: the goal in this case is to recreate realistic complex scenarios where humans are the very real target of the operations. New models must be studied and developed to reproduce population behaviour and reactions to disasters and humanitarian relief operations (Bruzzone et al. 2011,2007).
- Models Interoperability: a crisis has to be regarded as a very complex scenario; the correct recreation of such scenarios by computer simulation models requires to consider several elements (i.e. the natural events, the economical impacts, the operative issues…); this clearly compel to develop real interoperable simulation where different federated simulators are in charge of reproducing different aspects of reality in the most complete, realistic way.
- Exercises repeatability and costs control: the goal in this case is to provide real benefits to both Military Commands and Civil Authorities in joint crisis management training therefore it is required to provide tools easy to be deployed with reasonable costs. In order to solve this problem (simulation based exercises reproducing extended complex scenarios, with thousands or million of

people involved in the crisis), this project proposes the use of Computer Generated Forces based on Intelligent Agents (IA_CGF) with autonomous behaviour to reduce the staffs number involved in CAX (Computer Assisted Exercise), preparation and management and, at the same time, to rapidly study and generate new scenarios and scripts to face different kind of events.

- Usability and Users engagement: the "serious game" approach for training is growing in diffusion in the military community and there are a lot of specific areas, during a crisis, that could be reproduced by a serious game to train specific personnel capabilities

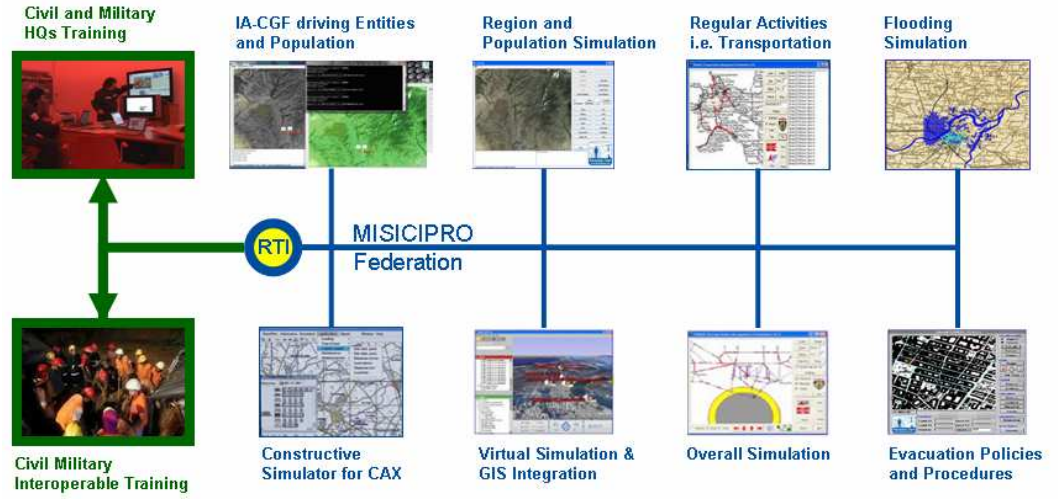


Figure 1: Sample of a Federation

There is a further area of development emerging from the results of the proposed research project: *innovative usage of M&S not only for training but also for operational planning and dynamic crisis management*. The researches described in this paper will pose the basis for the development of innovative models to be used for crisis management both in planning and execution phases. In fact we may suppose to provide different analytical supports for dynamic crisis management; in particular there are natural disasters where you can have a kind of "on-line" support from models to follow and analyze the event evolution over the time (i.e. flooding where decision makers have to predict the evolution of the scenario and to plan and coordinate interventions accordingly).

The general goal of this project is to provide a concrete support in disasters and crises management with the aim of developing a simulation that considers Civil-Military Interoperability and allows to demonstrate the results of the studies. To fulfill these goals a demonstrator will be developed during the activities. Another important concept is to create a virtual loop bringing to the Civil Protection Community the experience of the Military in Modeling and Simulation for training and to educate officers to co-operate with civilian considering the cultural and operative barriers that exits between the two communities. Accordingly to this another intent of the on-going researches described in this paper is to create a create a community involving military experts and Civil Protection personnel in training activities related to this context. This community could support in definition of innovative VV&A /Verification, Validation and Accreditation), a critical area where authors have several previous experiences (Bruzzone 2011, Tremori 2009).

2. PREVIOUS EXPERIENCES

The first step of this project is to overview the current state of the art above all in terms of ongoing research projects in other countries. Indeed we can cite several on-going activities at NATO level to explore Civil Military Interoperability. To mention one of the most important, we can cite the German experience where German Armed Forces and Federal Office of Civil Protection and Disaster Assistance (BBK) are jointly using the GESI constructive simulator (from CAE) in the BBK-owned Academy of Crisis Management, Emergency Planning and Civil Protection (AKNZ – Akademie für Krisenmanagement, Notfallplanung und Zivilschutz) in Bad Neuenahr. This project can be regarded as a good example to understand how to leverage on military experience and the starting point to provide innovative methodologies and solutions to enhance training and readiness of Civil Protection agencies.

For what concern the existing technological gap to Prof. Agostino Bruzzone's research team, in the last years, has worked on several projects to develop new computer generated forcers managed by intelligent agents (IA_CGF) (Bruzzone et al. 2011, 2010, 2009, 2008,2007) and based on human behavioral models to create complex simulated scenarios where agents act and react based on behavioral rules in an autonomous

way, considering the flowing of events and the perception of the events.

The first important project we can cite is PIOVRA (Polyfunctional Intelligent Operational Virtual Reality Agents), an European Defense Agency funded project. PIOVRA main goal was just to develop a new generation of CGF able to simulate "Intelligent" behavior. The final demonstrator based on the PIOVRA findings was used to recreate a urban scenario with civil disorders. From this project different research activities that represent solid base for the project described in this paper followed. Among the other it is necessary to cite another on-going project CAPRICORN (CIMIC And Planning Research In Complex Operational Realistic Network).

CAPRICORN project is going to be completed and it is providing final results and deliverables about the usage of agent driven simulation based on intelligent agents and Human Models to simulate Stability and Reconstruction operation, Civil Military Cooperation (CIMIC) and psychological operations (PSYOPS) where the clear final target is population and civilian.

Based on the researches in this area in 2010 the Authors worked with the US-Joint Force Command to re-create the Port O' Prince area during the 2010 Earthquake in Haiti. For that demonstration a IA_CGF Non Conventional Framework (IA_CGF NCF)developed by authors was federated with other systems (JTLS, JCATS, VBS2…) for representing the humanitarian support operation led by Coalition Forces in Haiti, and specifically in Port O'Prince. The IA_CGF NCF represented the 200.000 people population of the town before the disaster and after the operations, analyzing and representing evolution of human factors for every person in the area. The other simulators were used to re-create the different level of traditional operations, considering operative matters such as logistics or troops deployment, In this area authors have several experiences in the application of M&S also to the industrial sector and the described project could largely benefit from these works (Bruzzone 2004, Curcio&Longo 2009). This demonstrator is a very good sample of an HLA federation with different simulators specifically used to recreate different elements of the scenario and in particular with models based on Intelligent Agents to reproduce autonomous behavior of the population.

Authors have also specific competencies in modelling crisis scenarios. An important project was completed with Dartmouth College for the reproduction of a Katrina like scenario to demonstrate the possibility to model a national crisis and to simulate a wide emergency; the project successful demonstrated the simulation of an hurricane impact on the transportation layers of Louisiana state considering traffic cargo, evacuation activities, etc. Another project related to Crisis Management was PANDORA (PANdemic Dynamic Objects Reactive Agents) by a joint simulation project involving USA, European and Australian R&D Centers (DIPTEM, Dartmouth College, CRiCS).: it addressed the dynamics of the spreading of a Pandemic and experiments on H1N1 flu A virus. An evidence-based approach was used whereby statistical data (census) and ethnographic surveys were sources for the model and this was integrated with Human Factors representing the psychological and social parameters impact on people behaviors and their reaction to containment measures and policies. The models evaluated the efficacy and cost benefit of various mitigation strategies such as school closures, target anti-viral prophylaxis and other mitigation measures, level of absenteeism, and its impact on commerce, industry, economy and functioning of society as well as population attack rate, risks related to specific groups and on flows across State borders.

Another research project that should be mentioned is CIPROS (CIVIL Protection Simulator). As the name itself suggests CIPROS presented a a modular approach for Civil Protection that integrates GIS and Simulation. By the models developed during this project it was possible to generate Crisis Dynamic Web Sites for supporting training and information share. CIPROS included simulation of explosions, hazardous material fallout and flooding (Bruzzone et al. 1996).

3. THE PROJECT

To complete a general overview of this research we start from the innovative methodologies and technological enhancements the will be investigated and developed and in particular the three main pillars of the project described in this paper:

- Agent Driven Simulation and in particular based on Intelligent Agents Computer Generated Forces (IA_CGF)
- Human Behaviors Models: to provide a behavioral base for any models where the impact of humans is a critical element.
- Interoperable Simulation: and specifically such simulators will be federated under the IEEE 1516 High Level Architecture with the aim of creating a Federation of Simulators.

To provide evidence on the relevance of the approach proposed in this project, one of the first steps is the definition of the scenario to be modeled as a demonstration case study (see next paragraph). The successive step will be an accurate survey of the state of the art; the results of this survey will help in understanding which interoperable simulators from the Military area are available to support disasters and crises management with the aim of enhancing Civil-Military Interoperability. As additional step, it will be also possible to define also the demonstrators (simulators or serious games) to be developed in order to complete the above mentioned Federation. This simulator (or serious game) will provide an effective training environment with a realistic reproduction of the disaster scenario accordingly to the defined training targets.

By this project it is envisaged to provide a support to all the M&S related areas: from personnel training and education to operative analysis of different scenarios both for Defence for "Support Operations" and to support population through Civil Protection for disaster relief operations.

Needs to train such capabilities are reflected both in national standard policies for Peace Mission deployment and in specific Civil Protection Department operations. These policies will define conceptual references for the development of this project.

In particular specific operative capabilities that will be supported by this project are:

- Forecast: causes analysis to identify risks and critical areas
- Prevention: damages reduction and mitigation with a gradual improvement of population consciousness
- Relief: interventions aimed to ensure different types of assistance to populations
- Other Actions: different actions to remove obstacles and rubbles to restore normal life conditions;
- Personnel Training: military and/or civilian both collective and individual, in particular for procedures use;
- Procedures Development: development of innovative methodologies and techniques for the DMP (Decision making Process), crucial element for crisis management (Tremori 2011)

Accordingly to the above described capabilities to obtain an effective representation of complex scenarios (as required by the goal of this project), with involvement of military forces and civilians, it is required to define properly all the Political, Military, Economical, Social, Informatics and Infrastructural parameters (PMESII). Therefore models to be developed for this project shall include among the others the following characteristics:

- Entities for representing Civilian Population with related familiar and social networks
- Socio-Economic-Ethnic-Religious aspects for each entity
- Actions that could have an impact on population

4. SCENARIO AND COMPUTATIONAL ISSUES

In this paragraph we briefly summarize two of the possible scenarios that we are studying to provide an effective tool for training Civil Protection agencies and, at the same time, Armed Forces to interoperate in crisis management. A particular effort is done by authors to predict and face the different computational issues that emerge from complex scenarios where large amount of human beings are involved and, at the same time, specific incidents and all the related events has to be reproduced.

Based on previous experiences and on available simulators we could plan two kinds of exercise scenarios based on different kind of disasters: natural and manmade. In these two macro-categories we are planning to investigate and work on:

- Flooding and/or Earthquake in an highly populated area (see samples of existing models in Figure 2).
- Environmental Emergency with hazardous materials fallout.

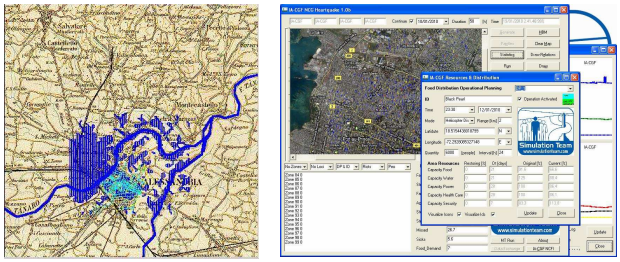


Figure 2: Sample s of existing models for Flooding and Earthquake disaster management

The above mentioned scenarios imply different problems for what concern simulators computational workload.

First of all there is the large number of people involved in the scenario and all the behavioral rules that must be considered to accurately reproduce the population. Among the different phenomena that must be considered to reproduce population behavior there are:

- The combination of Macro-Micro models to represent human behavior at social and single individual level and the combination of these two levels (Meso models).
- The parental and social relationships and networks that create counter-flow phenomena (i.e. if a family member is in the dangerous area you an agent will run towards the risky zone and against the general flow of entities running away).
- Leadership: in a group of people during any kind of event could emerge a leader: its leadership can be based on recognizable, official authority (i.e. a police officer) or on unofficial., single perception based authoritativeness (i.e. people know that on guy lives in the area and knows the places).
- Knowledge and perception of events that influence the decision of agents during the incident evolution

These and other phenomena must be consider at single agent level; in previous experiences (i.e. US Joint Force Command project for Haiti earthquake scenario) were simulated populations of over 200.000 units, considering all the relationships (parental and social), the perception and knowledge of the events, diffusion and sharing of information and other topics that influence single and, accordingly, population behavior.

For wider area, with higher number of people involved in scenario software engineering solutions will

be studied to face the exponential increasing of computational workload.

Another important kind of models that will be federated to realistically reproduce the different listed crisis scenarios are specific simulators to reproduce evolution or effects of the disaster. For the scenarios under evaluation at this stage we can see the following specific models to be federated:

- Diffusion and fall-out of hazardous materials: in previous projects such kind of simulators were developed by the authors to reproduce diffusion and fall-out over an area considering wind streams, terrain orography, buildings geometry.
- Flooding dynamics: accurate models must be developed to reproduce the evolution of a flooding and all the related effect on the terrain, transportation network, productive facilities…
- Earthquake Models: specific simulators must be used to appropriately model impact of an earthquake over a specific area. For instance are undergoing collaborations with the research centers (i.e. INGV - Istituto Nazionale di Geofisica e Vulcanologia) for federating different existing models to recreate an earthquake disaster scenario.

The above mentioned models imply severe issues for what concern computational workload there are, indeed, to be considered several elements (i.e. fluid-dynamic or structural) to obtain realistic results.

Finally we think it is important to be stressed that the above described issues imply increased complexity in time management in the Simulators HLA Federation: indeed we must consider that different models, representing different level of complexity run with different time rules (i.e. fast-time for constructive system at theater level, real-time for virtual systems and slow-time for CFD, structural or fluido-dynamic simulators).

5. PROJECT OBJECTIVES

As already highlighted in the introduction, the models developed in this project and the federation will be used, initially, for training, with the perspective to become in a further step the base to start new developments for operational planning purposes. In both cases, models and federation will have a positive impact in terms of:

- Support to Decision Makers that usually have access to partial or insufficient information.
- Reduction of time for decision making (time is usually very limited during the disaster management).
- Better use of the available resources (also resources are usually limited during the disaster management).
- Capability of considering multi-dimensional Problems involving population.
- Study and solve Simulation Computational Workload issues related to the large number of autonomous agents (i.e. population) and the characteristics and dynamics of the simulated incidents.

It is important also to underline models interoperability via the IEEE HLA standard for distributed simulation. As already mentioned disasters can be considered as very complex scenarios; the HLA choice can provide multiple benefits not only in terms of simulators interoperability, re-usability and composability but also in terms of future possible integrations (additional models to be integrated within the federation for investigating specific scenarios or recreating particular phenomena). This is the case, for instance, of flooding where specific models could be part of the HLA federation.

Furthermore, the use of interoperable simulation based on HLA provides the users with simulation available on-demand, easy share of resources among different agencies and involved offices in the Exercises or in operations planning.

6. CONCLUSIONS

We can conclude that the aim of the researches described in this paper will be to rapidly recreate different scenarios able to reproduce the constrains and the complexity of the different and elements of the scenario it self, this in terms of population behavior and impacts and evolutions of the specific simulated incident to prepare Civil and Military agencies to Crisis Management. This kind of joint training will be fundamental also considering all the cultural constraints of Civil-Military Interoperability. Furthermore the use of distributed simulation based on HLA combined with Computer Generated Forces with autonomous behavior and Intelligent Agents will allow the development of innovative solutions at reasonable costs with the proper attention to realism and fidelity.

REFERENCES

Amico Vince, Guha R., Bruzzone A.G. (2000) "Critical Issues in Simulation", Proceedings of SCSC, Vancouver, July

Avalle L, A.G. Bruzzone, F. Copello, A. Guerci, P.Bartoletti (1999) "Epidemic Diffusion Simulation Relative to Movements of a Population that Acts on the Territory: Bio-Dynamic Comments and Evaluations", Proc. of WMC99, San Francisco, January

Bocca E., Pierfederici, B.E. (2007) “Intelligent agents for moving and operating Computer Generated Forces”, Proc.of SCSC, San Diego July

Bruzzone A.G., Massei M. Tremori A., Bocca E., Madeo F., Tarone F, (2011) “CAPRICORN: Using Intelligent Agents and Interoperable Simulation for Supporting Country Reconstruction”, Proceedings of DHSS2011, Rome, Italy, September 12 -14

Bruzzone A.G. Tremori A., Massei M. (2011) "Adding Smart to the Mix", Modeling Simulation & Training: The International Defense Training Journal, 3, 25-27, 2011

Bruzzone A.G., Massei M., Tremori A., Bocca E., Madeo F. (2010) "Advanced Models for Simulation of CIMIC Operations: Opportunities & Critical Issues Provided by Intelligent Agents", Proceedings of DIFESA2010, Rome, Italy, April

Bruzzone A., (2010) “Human Behaviour Modelling as a Challenge for Future Simulation R&D: Methodologies and Case Studies”, invited speech at Eurosim 2010, Prague

Bruzzone A.G., Massei M. (2010) "Intelligent Agents for Modelling Country Reconstruction Operation", Proceedings of AfricaMS 2010, Gaborone, Botswana, September 6-8

Bruzzone A., Madeo F., Tarone F., (2010) “Modelling Country Reconstruction based on Civil Military Cooperation”, Proc. of I3M2010, Fez, Oct.

Bruzzone A.G., Massei M., Tremori A., Bocca E., Madeo F. (2010) "Advanced Models for Simulation of CIMIC Operations: Opportunities & Critical Issues Provided by Intelligent Agents", Proceedings of DIFESA2010, Rome, Italy, April

Bruzzone A.G., Reverberi A., Cianci R., Bocca E., Fumagalli M.S, Ambra R. (2009) "Modeling Human Modifier Diffusion in Social Networks", Proceedings of I/ITSEC2009, Orlando, Nov-Dec

Bruzzone A.G., Frydman C., Cantice G., Massei M., Poggi S., Turi M. (2009) "Development of Advanced Models for CIMIC for Supporting Operational Planners", Proc. of I/ITSEC2009, Orlando, November 30-December 4

Bruzzone A.G., Frydman C., Tremori A. (2009) "CAPRICORN: CIMIC And Planning Research In Complex Operational Realistic Network" MISS DIPTEM Technical Report, Genoa

Bruzzone A.G., A. Tremori, M. Massei (2008) "HLA AND HUMAN BEHAVIOR MODELS" Proceedings Of MMS2008, April 14 - 17, 2008, Ottawa, Canada

Bruzzone A.G. (2008) "Intelligent Agents for Computer Generated Forces", Invited Speech at Gesi User Workshop, Wien, Italy, October 16-17

Bruzzone A.G., Scavotti A., Massei M., Tremori A. (2008) "Metamodelling for Analyzing Scenarios of Urban Crisis and Area Stabilization by Applying Intelligent Agents", Proceedings of EMSS2008, September 17-19, Campora San Giovanni (CS),Italy

Bruzzone A.G., Massei M. (2007) "Polyfunctional Intelligent Operational Virtual Reality Agent: PIOVRA Final Report", EDA Technical Report

Bruzzone A.G., Brano E., Bocca E., Massei M. (2007). Evaluation of the impact of different human factor models on industrial and business processes". SIMULATION MODELING PRACTICE AND THEORY, vol. 15, p. 199-218, ISSN: 1569-190X

Bruzzone A. G et al. (2004) "Poly-Functional Intelligent Agents For Computer Generated Forces", Proceedings of the 2004 Winter Simulation Conference Washington D.C., December

Bruzzone A.G., (2004). Preface to modeling and simulation methodologies for logistics and manufacturing optimization . SIMULATION, vol. 80 , p. 119-120 , ISSN: 0037-5497, doi: 10.1177/0037549704045812

Bruzzone A.G., Giribone P., Mosca R. (1996). Simulation of Hazardous Material Fallout for Emergency Management During Accidents. SIMULATION, vol. 66 (.6), p. 343-355, ISSN: 0037-5497Caussanel J., Frydman C., Giambiasi N., Mosca R. (2007) “State of art and future trend on CGF” Proceedings of EUROSIW2007, Santa Margherita, Italy, June

Curcio D, Longo F (2009). Inventory and Internal Logistics Management as Critical Factors Affecting the Supply Chain Performances. INTERNATIONAL JOURNAL OF SIMULATION & PROCESS MODELLING, vol. 5(4), p. 278-288, ISSN: 1740-2123

Fletcher M., (2006) “A Cognitive Agent-based Approach to Varying Behaviours in Computer Generated Forces Systems to Model Scenarios like Coalitions”, Proceedings of the IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and its Applications,

Kallmeier V., Henderson S., McGuinness B., Tuson P., Harper R., Price S. Storr J. (2001) "Towards Better Knowledge: A Fusion of Information, Technology, and Human Aspects of Command and Control", Journal of Battlefield Technology, Volume 4 Number 1.

Lichtblau, D., et al., (2004) “Influencing Ontology,” ex-tended abstract in Behavior Representation in Modeling and Simulation Conference

Rietjens S.J.H., M. Bollen, (2008) “Managing Civil-Military Cooperation”, Military Strategy and Operational Art

Nacer A., Taylor A., Parkinson G. (2007) “Comparative Analysis of Computer Generated Forces” Artificial Intelligence, Ottawa

Rehse P., (2004) “CIMIC : Concepts, Definitions and Practice”, Heft 136, Hamburg, June

Thagard, P., (2000) Coherence in Thought and Action, Cambridge, MA: MIT Press

Tremori A., Baisini C., Enkvist T., Bruzzone A.G., Nyce J. M. (2012), “Intelligent Agents and Serious Games for the development of Contextual Sensitivity”, Proceedings of AHFE 2012, San Francisco, US, July

Tremori A., Bocca E., Tarone F., Longo F., Poggi S. (2009)"Early Testing Procedures For Supporting Validation Of Intelligent Agents For Simulating Human Behavior In Urban Riots", Proceedings of MAS2009, Tenerife, September

RENEWABLE ENERGY SOURCES: ADVANCED SOLUTIONS FOR FLOATING PHOTOVOLTAIC SYSTEMS

Mirabelli Giovanni[(a)], Nicoletti Letizia[(b)], Pizzuti Teresa[(c)] , Stumpo Pierluigi[(d)]

[(a)(b)(c)]Department of Mechanica Engineering, University of Calabria
[(d)]Freelance Professional

[(a)] g.mirabelli@unical.it, [(b)]letizia_nicoletti@libero.it, [(c)]teresa.pizzuti@unical.it, [(d)] mailto:p.stumpo@alice.it

ABSTRACT

In times of crisis, the need to achieve higher levels of competitiveness and reduce costs has driven many companies to use renewable energy sources and take advantage of the related incentives. In addition, the need to reduce delivery service costs and keep good quality standards is noticed by local authorities whose monetary transfers from the central government have been decreased. Moreover Land reclamation and drainage authorities are not exempt from the considerations made above; in fact, they are required to pursue their mission toward the land and the associated farms regardless of the monetary transfers received from the Government.

Within this framework, this research work is intended to find new opportunities and new ways to exploit the available resources better. This study has been carried out in collaboration with The University of Calabria, The Land reclamation and drainage authority of Cosenza and the Association Fattorie del Sole - Coldiretti Calabria.

In particular, it aims to prove that solar powers from innovative photovoltaic systems, wich are installed on the watersheds managed by the Italian Land Reclamation Consortia, can provide tangible benefits.

To this end it has been shown that the most important factors to be considered are: the production capacity for each investigated location, the systems structural design and the verification of economic and financial sustainability.

Keywords: Renewable energy, Photovoltaic systems

1. INTRODUCTION

Photovoltaic systems suitability for electric power production has been widely demonstrated (Ciriminna et al. 2011). As a matter of facts, in the last few years the installation of such systems has grown exponentially in Italy (from 2010 to 2011 the electric power increased from 3.470 MW to 11.340,4 MW, GSE).

This result has placed the Italian market first in Europe for installations number, electric power produced and turnover. The plant designs that have been used are both rooftop installations (which require an authorization process quite simple) and ground installations that require more costly and complex authorization processes (Camilleri, 2011). Lately, floating photovoltaic systems (including both traditional and Floating, Tracking, Cooling, Concentrators) have been proposed and are still under study; they allow using bodies of water for energy production. In Italy such systems can be dated back to 2008 (Solarolo RA) and since then many companies and research centres are seeking for more efficient and effective engineering solutions. Some of them have also developed patented in-house solutions. Certainly, mirrors of water exploitation results in higher costs and requires more complex structural solutions compared to conventional installations (on roofs or on ground) but allow achieving relevant advantages:

- There is no need to reduce the areas intended for agricultural crops;
- They allow obtaining considerable performances enhancements (in summer and in winter). In summer, the mitigating effect of the water allows the panels to work at lower temperatures whereas, in winter, the panels defrost more quickly therefore daily production can be increased.
- They allow using simpler and more effective solar tracking systems; as a result the average yield can be increased by 20%.
- They allow reducing the water evaporation: floats and panels have a covering effect that limits the direct exposure to sunlight and the catchment areas overheating.
- They allow reducing theft or damage risks owing to the lower accessibility.
- They allow an easier maintenance; water availability speeds up surfaces cleaning therefore the efficiency of panels can be kept high.
- They allow obtaining greater incentive rates thanks to their innovative nature.

2.

3. THE ENERGETIC ISSUES OF THE LAND RECLAMATION AND DRAINAGE AUTHORITY OF COSENZA

The Land reclamation and drainage authority of Cosenza covers an area of 110,000 hectares, includes 46,000 associate companies, encompasses 32 unicipalities and its energetic expenditures can be attributed to 40 sites (i.e., administrative headquarters, representation offices, operational offices, support offices, lifting systems, water scooping).

The assessment of the energy consumptions reveals that the installed electrical power capacity is about 7 MW and costs for power supply are about 500,000 €/year. These data highlight the need for technical solutions, able to reduce the energy consumptions, and for real opportunities of in-house energy production. In order to address this problem organically and evaluate different options, a group of experts from the Mechanical Department of the University of Calabria and from the Land reclamation and drainage authority was set up. The working group, after evaluating different options, find out that solar energy is the most viable choice that could take advantage from the numerous reservoirs managed by the Consortium.

Therefore the identification and analyzes of the main features of the reservoirs located all over the consortium territory, was the first step of this research.

By doing so, approximately 30 catchments were detected and examined in order to identify the most interesting sites that were selected considering: catchments size, geographical location, possibility of an easy connection with electricity networks, topography of the area, etc. Table 1 reports the main reservoirs, their extent and the potential energy production of the photovoltaic systems that can be installed on.

Table 1: The catchments under study

Municipality	Catchment denomination	Area (mq)	Capability (kWp)
Corigliano Calabro	Bacino Mandria del Forno	5.570	200
Corigliano Calabro	Vasca 4	795	53
Trebisacce	Vasca Saraceno	35.000	2.300
Trebisacce	Vasca 9 – Rovitti	800	53
Rocca Imperiale	Vasca 5	4.290	330
Rocca Imperiale	Vasca 6	3.050	200
Cerchiara	Vasca del Caldanello	2.300	155
Amendolara	Vasca 6	700	47

Afterwards the energetic capability of the selected catchments was calculated (Cucumo et al, 1994). To this end, the best computational and simulation tools, currently available, were used and the features of a broad spectrum of photovoltaic panels (marketed in the main international markets) were considered. With regard to the computational approach, it is worth saying that the main underlying tools were PVGIS and SAM; the results obtained by applying these tools have been accurately analyzed and compared.

The former (PVGIS) is currently used in Europe and it allows having an estimate of the energy output from the photovoltaic system taking into account the geographical location, the type of system (power, exposure, etc.) and the potential losses (Suri et al., 2008). The latter (SAM - System Advisor Model) has been developed by the National Renewable Energy Laboratory (NREL) in collaboration with the Sandia National Laboratories and the U.S. Department of Energy within the SETP project framework (Gilman and Dobos, 2012). SAM is more flexible since it allows inserting and managing more data and therefore its results are more accurate and reliable.

Moreover to achieve the research goals above mentioned, 14 different types of photovoltaic panels (whose main features are reported in table 2) have been investigated and compared.

Table 2: Features of the analyzed photovoltaic panels

Panel type	Power	Efficiency	Typology	Weight	Area
Aleo S166 185	185 W	13,42%	Mono – Cri	17 Kg	1,378 mq
BP Solar BP 585	84 W	13,14%	Mono – Cri	7,7 Kg	0,634 mq
BP Solar BP 4175	174 W	13,88%	Mono – Cri	15,4 Kg	1,260 mq
Sanyo HIP-190 BE11	180 W	14,09%	Mono – Cri	15,5 Kg	1,277 mq
BP Solar SX 170B	170 W	13,52%	Poli – Cri	15 Kg	1,258 mq
Evergreen ES195	195 W	13,05%	Poli – Cri	18,2 Kg	1,494 mq
Kyocera Solar KC130 TM	130 W	14%	Poli – Cri	11,9 Kg	0,929 mq
Mitsubishi Pv UD190 MF5	190 W	13,78%	Poli – Cri	17 Kg	1,383 mq
SunPower SPR – 210	210 W	16,89%	Poli - Cri	15 Kg	1,244 mq
Suntech STP180s 24AC	180 W	14,09%	Mono - Cri	15,5 Kg	1,277 mq
Solyndra SL200-220	220 W	16%	-	31,8 Kg	2,485 mq
SunPower E20/333	333 W	20,40%	Mono – Cri	18,6 Kg	1,631 mq
Suntech STP 280-24 Vd	280 W	14,40%	Poli – Cri	27 Kg	1,940 mq
Sanyo HIT-H250E01	250 W	18%	Mono – Cri	16,5 Kg	1,386 mq

At this stage the technical features of the panels (dimensions, weight, efficiency, power) together with the size of the reservoirs under study have drawn the attention on the panel SunPower E20/333. Therefore the two previously mentioned software tools have been applied to assess the capability of this particular panel for each considered catchment. Results are reported in Table 3.

Analyzing the data reported on table 3 it is possible to notice that, for each reservoir, the results obtained by using SAM are greater than those obtained by using PVGIS, the mean difference is 13,9%.

These findings allow estimating the total potential energy: 4,235,400 kWh per year using PVGIS and 4,780,933 kWh per year by using SAM. These results have been the starting point for the subsequent economic evaluations.

4. TECHNICAL AND STRUCTURAL DESIGN.

The design of the floating structures, allowing The stability and the functionality of the solar panel, has been set up to obtain the following results:

- modularity of building and assembling;
- low weight;
- low cost;
- high structural stiffness;
- weather resistance;
- easy assembly and maintenance;
- easy to move, connect and wire;
- high durability (25 years min.);
- optimal exploitation of free surfaces;
- maximum energy performance through the reduction of shading.

The initial data covered: the weight of the panels, their size, the standard and accidental loads, the dynamic actions of the water, the disruptive actions of the wind, the flotation coefficients (DM14-01-2008), the durability of materials, the sunpath study (UNIEN10002-1:2004), the shape of the reservoirs.

In order to solve the complex structural problem, a 3D truss has been realized, by means of steel tubular elements connected to form several square base pyramids, joined using other tubular elements as shown in Figure 1(CNR 10011–1997).

Table 3: Estimate of annual electricity production for each reservoir (SunPower E20/333)

Site	PVGIS [kWh] year	SAM [kWh] year	Gap
Corigliano Bacino Mandria del forno	250.000	285.064	14,03%
Corigliano Vasca 4	66.100	75.109	13,63%
Trebisacce - Vasca Saraceno	2.940.000	3.292.988	12,01%
Trebisacce - Vasca Rovitti	67.800	75.361	11,15%
Rocca Imperiale - Vasca 5	408.000	470.037	15,21%
Rocca Imperiale - Vasca 6	247.000	288.340	16,74%
Cerchiara - Vasca del Caldanello	196.000	225.226	14,91%
Amendolara - Vasca 6	60.500	68.808	13,73%

In this way, the goals of modularity and easy workability have been reached; in fact different series of linear assembling could be set up in laboratory by opportunely yielding the linear elements, together with the surface treatments, while the final assembling could be easily reached using different types of patented structural joints, ensuring a very low corrosion level due to a very few in-situ operations.

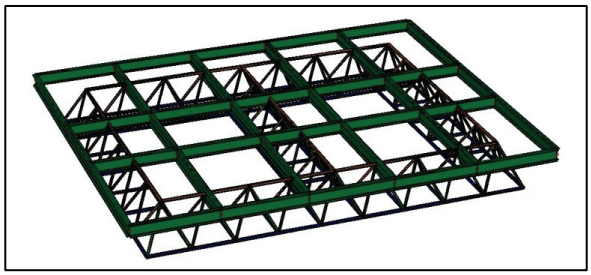


Figure 1 - Structural solid scheme.

Once made and designed the structure, it has been modeled by means of linear two-joints finite elements, connected in joints which allow the complete transmission of axial and shear stresses, together with a small percentage (20%) of flexural and torsion ones; in this way it has been taken into account that hinge node is a theoretical abstraction, and actually a transmission of moment stresses is active between the elements of a truss (Badalamenti et al., 2008). Then, defined both standard and accidental loads, has been taken into account also the anchorages to the basin bottom, and a dynamical analysis has been carried out by means of a SAP-type FEM software, which solves the dynamical equilibrium equation (1):

$$M \cdot \ddot{u} - C \cdot \dot{u} + K \cdot u = p \qquad (1)$$

In equation (1) $\boldsymbol{M}$ is the mass tensor, $\boldsymbol{C}$ is the damping tensor and $\boldsymbol{K}$ is the stiffness tensor, while $\boldsymbol{u}$ and $\boldsymbol{p}$ are the displacement and load vectors, respectively. The former equation has been solved in terms of eigenvalues and eigenvectors of the structure, corresponding to the natural resonance frequencies and the relative displacements, and 150 vibrating modes have been computed, to take into account the higher resonance frequencies also. The static and dynamic solutions have been combined following the instructions of the Italian building standards (N.T.C. 14-01-2008 and the related rule book), using the standards combination coefficients and considering the loss of elasticity as the limit state, in order to avoid corrosion, located to the yielding points above all. However, the achieved stress levels (see Figures 2, 3, 4, and 5) are low if compared with the maximum elastic stresses of the chosen material, and the material weight is the main load condition for these structures, which are specifically designed to be installed in fresh water basins, such as small lakes or artificial water basins for agricultural irrigation (Chakrabarty, 2010).

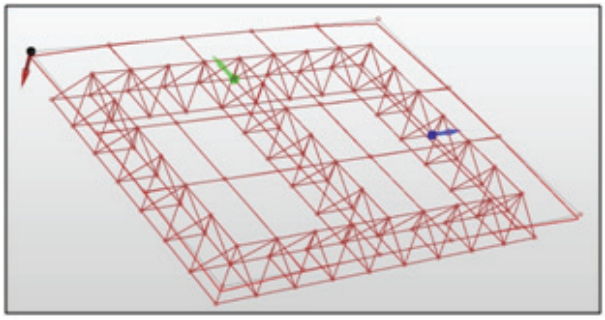


Figure 2 – Structural FEM model.

The reached solutions refer to two modular structures ensuring an installed nominal power of 3 kWp and 5 kWp each. It is evident that composing these basic modules it comes to saturate the available surfaces of the various basins, reaching the estimates of performances already shown in table 1.

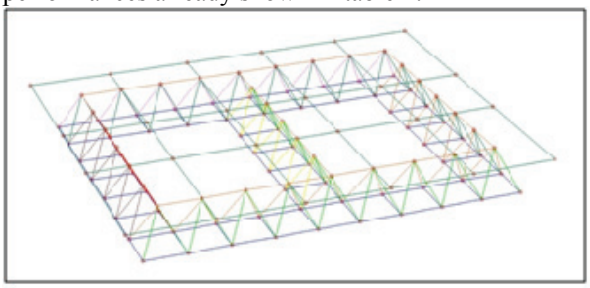


Figure 3 – Static displacements

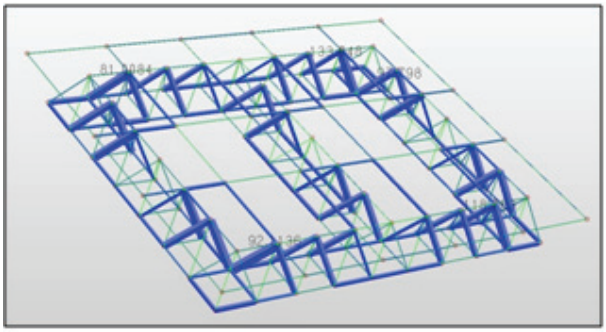


Figure 4 – Axial stresses [kg].

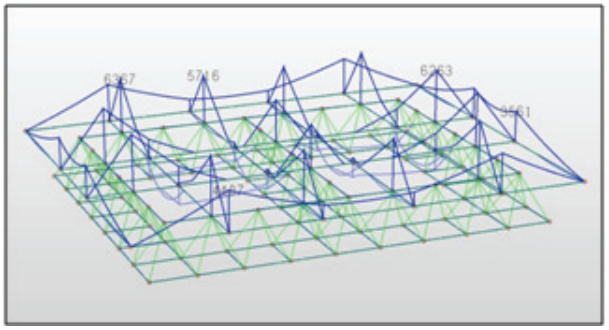


Fig. 5. - Flexural stresses [kg·cm].

The essential features of the modular structures are summarized in Table 4, while the figures 6, 7, 8, 9, and 3 show views in plant and section of the designed structures.

Table 4 – Main features of the designed structures

Module	3 kWp	5kWp
Dimension [m]	4 x 5	6,5 x 5
Main structure	3D truss	3D truss
Secondary structure	IPN 140	IPN 140
Material	S350 type Steel	S350 type Steel
Total weight	1250 kg	1850 kg
Bridging service	Orsogrill type panels	Orsogrill type panels
Surface treatment	Hot dip galvanizing (triple treatment)	Hot dip galvanizing (triple treatment)
Solar panels	10 panels	15 panels
Floaters	12 units	16 units
Cost	5.000 €	9.000 €

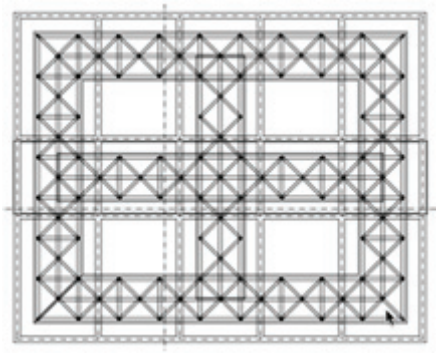


Figure 6 – Layout of the 3 kW module

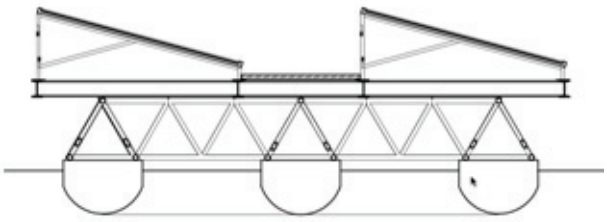


Figure 7 – Section of the 3 kW module
Figure 8 – 3 kW Module

5. ECONOMICAL EVALUATIONS

In order to verify the project efficiency and financial sustainability further research activities have been carried out and the economical outcomes have been estimated.

Revenues have been assessed considering both the incomes from energy sale and the incentives provided on the Italian territory through the Fifth Energy Bill approved in July 2012 (DM 5 July 2012).

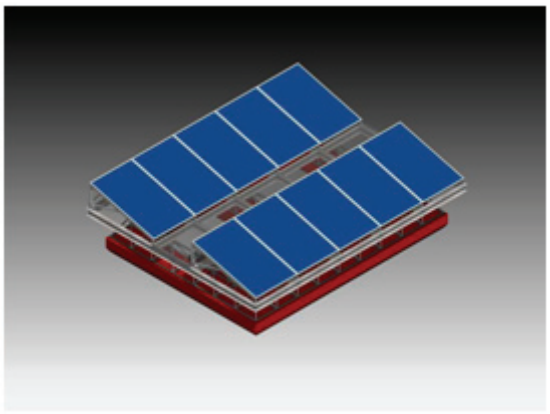


Figure 8 – 3 kW Module

The current rules in this field establish that incentives vary depending on the installed power. In particular, there are two forms of incentive. The former is directly provided by the Energy Services Operator (GSE) and aims to reward the use of renewable sources.

On the other hand, the latter rewards those subjects who use the self-produced energy without feeding it into the grid. However, it should be noted that the guaranteed minimum price, which is paid to the producer when the system is connected to the grid, is defined and periodically updated by the Authority for Electricity and Gas (AEEG). These prices are paid by the GSE only for the first 2 MWh of power fed into the grid in one year. Furthermore the incentive policies of the GSE for the floating photovoltaic systems are more rewarding than those applied for traditional systems. In this study the ministerial tables (for systems that have been installed since September 2012) have been referred to. In tables 5 and 6 the incentives have been reported (DM 2011).

Table 5 – Gse Incentives , From September 2012

Power Of The System	Minimum Guaranteed Incentive
[KWP]	[€/KWH]
1< P ≤ 20	0,288
20< P ≤ 200	0,276
P > 200	0,255

In order to estimate the energy produced in one year, the panels efficiency has been reduced by 0.9 % compared with the previous year. This calculation procedure estimates 18% reduction of the panels overall efficiency over the last operating year. The cost assessment allows identifying two different categories: set-up costs and operating costs. (Li Calzi et al., 1995) In addition, to make the economic findings more usable, a "parametric cost" has been introduced; to this end the single kWp of installed power was used as a reference. In this way, it is possible to find out aseptic monetary values that can be used to estimate the total costs changing the potentials of the system.

Table 6 – Incentive For Self-Consumption From September 2012

Power Of The System	Minimum Guaranteed Incentive
[KWP]	€/KWH
1< P ≤ 20	0,186
20< P ≤ 200	0,174
P > 200	0,153

The obtained parametric costs are reported in table 7 whereas table 8 provides the overall costs for installing such systems in the analyzed basins (Longo et al., 2006).

Table 7 – Parametric costs [€/kWp]

Description	[€/kW]
3kW Floating module	1850
5 kW Floating module	1750
Panels	1000
Wiring and hookings	240
Structure and connections	230
Special systems and lighting	260
Inverter and wiring	300
Increase for connection in MT	200

Table 8 – Estimated installation costs [€]

Site	[€]
Corigliano Calabro – Vasca 4	200.000
Corigliano Cal. – Bacino Mandria del Forno	756.000
Trebisacce – Vasca Saraceno	8.694.000
Trebisacce – Vasca Rovitti	200.000
Rocca Imperiale – Vasca 5	1.247.000
Rocca Imperiale – Vasca 6	756.000
Cerchiara di Cal. – Vasca del Caldanello	586.000
Amendolara – Vasca 6	177.000
Totale investimento	12.616.000

Besides set-up costs, operating costs including maintenance, insurance and general expenses have to be considered. These costs are estimated as a percentage, ranging from 0.075% to 0,095%, of the initial capital expenditure. Furthermore the calculation procedure being adopted in this work is based on the following assumptions:

- the total revenue includes the incentives related to the self-consumption and those granted by the GSE;

- the initial outlay is the total cost of the system and can be obtained multiplying the parametric costs by the installed power;
- the operating financial requirements are calculated as percentage of the overall set-up cost;
- cash flows are calculated as the difference between income and expenditure.

Moreover the economic assessments have taken into account the current economic scenario, therefore two main options have been considered. The former consists of supporting the investment with both equity and borrowed capital; the equity capital is 20% of the total set-up costs, and the borrowed capital stems from a fifteen-year mortgage with a 6% interest rate. The latter consists of financing the whole amount with equity. In addition, the first option is based on an amortization schedule where the depreciation charge is constant; thereby the debt can be wiped out before the service life is exhausted and the profits can cover the risk capital since the first years. These options have been assessed by using two well-known and well-established methods: the Net Present Value (NPV) and the Discounted Cash Flow Rate of Retourn (DCFRR) (Simeoni and Vigolo, 2006; Iotti and Bonazzi, 2012). Tables 10 and 11 report the results that have been obtained for the basin Vasca 4 located in Corigliano Calabro.

Table 10 – Discounted cash flows $\boldsymbol{i}$ = 7.82% (Equity Level = 20%)

Year	Revenues (€)	Costs €)	Fk €)	Fk dis.(€)	∑ Fk dis. €)
0	0	200.000	-200.000	-200.000	-200.000
1	33.799	15.435	18.364	17.029	-182.971
2	33.495	15.023	18.473	15.884	-167.087
3	33.194	14.585	18.608	14.837	-152.250
4	32.895	14.122	18.773	13.880	-138.369
5	32.599	13.631	18.968	13.005	-125.364
6	32.306	13.110	19.195	12.204	-113.161
7	32.015	12.558	19.457	11.470	-101.690
8	31.727	11.973	19.754	10.799	-90.891
9	31.441	11.353	20.088	10.183	-80.708
10	31.158	10.695	20.463	9.619	-71.089
11	30.878	9.999	20.879	9.101	-61.988
12	30.600	9.260	21.340	8.625	-53.363
13	30.324	8.477	21.847	8.189	-45.174
14	30.051	7.647	22.404	7.787	-37.388
15	29.781	6.767	23.014	7.417	-29.971
16	29.513	5.835	23.678	7.076	-22.894
17	29.247	5.835	23.412	6.488	-16.406
18	28.984	5.835	23.149	5.949	-10.457
19	28.723	5.835	22.888	5.454	-5.003
20	28.465	5.835	22.630	5.000	-3

Figure 4 illustrates how the cash flows recover the initial investment for both the options under study; the actualization index has been set equal to the minimum guaranteed return of the investment (7%).

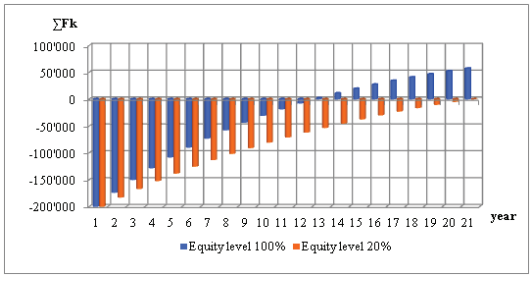


Fig. 4 – Discounted cash flows ($\boldsymbol{i}$ = 7 %)

6. CONCLUSIONS

The research activities that have been carried out allow pointing out the technical feasibility of floating photovoltaic systems as well as the related benefits. The proposed solution is easy and inexpensive to put in practice, safe, easy to inspect and maintain.

Furthermore, some simple constructive measures allow load-bearing structures (platforms flotation) to be assembled with different kinds of commercial photovoltaic panels in order to fill the available spaces fully.

However, it should be noted that today in Italy there is a unified regulatory framework therefore the various local authorities grant authorizations for the exploitation of reservoirs based on autonomous rules and regulations. Furthermore in implementing implementation of such photovoltaic plants the availability of water resources for civil purposes (i.e. fire-fighting systems) has to be ensured.

Table 11 – Discounted cash flows $\boldsymbol{i}$ = 11.64% (Equity Level = 0%)

Year	Revenues (€)	Costs €)	Fk €)	Fk dis. €)	∑ Fk dis.(€)
0	0	200.000	-200.000	-200.000	-200.000
1	33.799	5.835	27.964	25.047	-174.953
2	33.495	5.835	27.660	22.191	-152.762
3	33.194	5.835	27.359	19.659	-133.103
4	32.895	5.835	27.060	17.416	-115.687
5	32.599	5.835	26.764	15.429	-100.258
6	32.306	5.835	26.471	13.668	-86.590
7	32.015	5.835	26.180	12.108	-74.482
8	31.727	5.835	25.892	10.725	-63.757
9	31.441	5.835	25.606	9.501	-54.256
10	31.158	5.835	25.323	8.416	-45.840
11	30.878	5.835	25.043	7.454	-38.386
12	30.600	5.835	24.765	6.603	-31.783
13	30.324	5.835	24.489	5.848	-25.935
14	30.051	5.835	24.216	5.180	-20.756
15	29.781	5.835	23.946	4.588	-16.168
16	29.513	5.835	23.678	4.063	-12.105
17	29.247	5.835	23.412	3.598	-8.507
18	28.984	5.835	23.149	3.187	-5.320
19	28.723	5.835	22.888	2.822	-2.497
20	28.465	5.835	22.630	2.499	2

The Reclamation Consortia, may achieve significant advantages from these systems: they can reduce the cost of the services provided to its members (i.e. water supply for irrigation purposes) and acquire financial resources to maintain their own infrastructures or support new ones internally.

The economical evaluations have shown that profitability ratios range from 5% to 12% with estimated payback periods of about 10 years , either by obtaining external loans or by using equity capital.

Therefore the financial commitment required to implement such initiative is high extremely important and hardly feasible in the short period. A good solution could be to start the process focusing on small installations that require lower financial contributions and then consider larger systems taking advantage of the acquired experience.

In conclusion, it should be noted that if the outcomes are considered from an entrepreneurial perspective may appear not particularly rewarding. The point of view of public authorities, instead, may lead to better ratings because ethical and social are taken into account.

.

REFERENCES

Ciriminna R., Pagliaro M., Palmisano G., 2011, L' energia solare in agricoltura. Reddito economico e risanamento ambientale, Maggioli Editore, Dicembre 2011, ISBN 8838769265

Camilleri, C., 2011, Fotovoltaico. Analisi di fattibilità per un'edilizia sensibile. Dal protocollo di Kyoto al varo del quarto conto energia, Geva ed., 2011, ISBN: 8889323779

Cucumo, M., A., Marinelli, V., Oliveti G., 1994, Ingegneria Solare – principi ed applicazioni, Pitagora Editrice, Bologna 1994, ISBN: 8837107293

Suri, M., H., T., Dunlop, E., D., Cebecauer, T., 2008, Geographic Aspects of Photovoltaics in Europe: Contribution of the PVGIS Website, Joint Res. Centre, Eur. Comm., Ispra March 2008.

Gilman, P., Dobos, A., 2012, System Advisor Model, SAM 2011.12.2: General Description National Renewable Energy Laboratory (NREL) February 2012

DM14-01-2008. Norme tecniche per le costruzioni. Gazzetta Ufficiale del 04-02-2008.

UNIEN10002-1:2004:"Materiali metallici - Prova di trazione".

CNR 10011–1997: "Costruzioni di acciaio. Istruzioni per il calcolo, l'esecuzione, il collaudo e la manutenzione".

Badalamenti, V., Colajanni, P., La Mendola, L., Pucinotti, R., Scibilia, N., 2008, Indagine sperimentale su tralicci in acciaio di travi reticolari miste, Prooceedings of the XVII, Conference CTE, Rome, 6-8 November, 193-202, (2008).

Chakrabarty, J., 2006, Theory of Plasticity, Butterworth-Heinemann, Vol I, third edition, (2006).

DM 2011, Decreto Ministeriale del 05/05/2011, pubblicato sulla Gazzetta Ufficiale del 12/05/2011

Li Calzi A., Passannanti A., Lombardo A., Abbate L., Bellomo, V., Perricone L., 1995 Manuale di Ingegneria Gestionale – v. II: Strumenti complementari avanzati, Licam Editore, Palermo 1995.

Longo F. , Grande A. , Mirabelli G. , Papoff E., 2004 Aspetti tecnici ed organizzativi di processo, Proceedings of the National conference Animp, Treviso, 14-15 ottobre 2004

Simeoni F., Vigolo V.,2006 Modelli per la valutazione della convenienza degli investimenti industriali, Giuffrè ed. 2006, ISBN: 8814133182

Iotti M., Bonazzi G., 2012 Evaluation of Industrial Investments: Some Hypothesis of Methodological Improvement, International Journal of Business and Behavioral SciencesVol. 2, No.4; April 2012

AUTHORS BIOGRAPHY

Giovanni Mirabelli is currently researcher at the Mechanical Department of University of Calabria. He has published several scientific papers participating as speaker to international and national conferences. He is actively involved in different research projects with Italian and foreign universities as well as with Italian small and medium enterprises.

Letizia Nicoletti is PhD student at MSC-LES, Mechanical Department of University of Calabria. She took the Management Engineering degree, summa cum Laude, in December 2009 from the University of Calabria. Now She collaborates with the Industrial Engineering Section of the University of Calabria to research projects for supporting innovation technology in SMEs.

Teresa Pizzuti took her degree in Management Engineering, summa cum Laude, in July 2010 from the University of Calabria. She is currently PhD student at the Mechanical Department of University of Calabria. Her research activities concern the study and application of innovative methods in order to improve quality and logistics of industrial products, referring in particular to agri-food product.

INTELLIGENT SYSTEMS FOR THE CORE OF ANTHROPOCENTRIC OBJECTS AND ITS MODELING

B.E. Fedunov
State Research Institute of Aviation Systems, Moscow Aviation Institute
Moscow, Russia
E-mail: boris_fed@gosniias.ru

ABSTRACT
Classification and patterns of computer simulation systems of onboard intelligent systems are presented; these systems provide improvement of knowledge bases of onboard intelligent systems at particular stages of their development. In the computer simulation system of the intelligent information system "The situation awareness of the crew" (the first global control level of anthropocentric object) the presence of a man-operator in the simulation loop is mandatory, and for onboard online advisory expert systems for typical situations of a functioning an anthropocentric object (the second global control level of anthropocentric object), the operator work is simulated by a special a block of situation management.

Keywords: onboard intelligent systems, computer simulation systems for the development of there knowledge bases, algorithms of an activity of an operator.

1. INTRODUCTION

An anthropocentric object is a shell with a great number of macrocomponents implemented in it (the board of the anthropocentric object):

- the onboard measurement devices receiving information from the external world, in which the anthropocentric object operates, and from its onboard world;
- the core of the anthropocentric object in which three global control levels are distinguished: online assigning of the goal (the first global control level), determination of method for achieving online assigned of the goal (the second global control level), realization of this method (the third global control level). In the core the main role belongs to the team of operators (crew);
- the onboard executive devices acting on the external and onboard world.

It is convenient to describe operation of any anthropocentric object by the calculated set of its operation sessions; each of these sessions is characterized by the general goal of the session and the semantic network of typical operation situation in turn, each of these typical situations is represented by the semantic network of problem subsituations of this typical situation. All above are the content of the model "Stage" (another name "General task - global control level" (Fedunov 2010).The model "Stage" is using in the time the development of onboard algorithmic and indication providing of modern anthropocentric objects.

The practice of application and development of modern anthropocentric objects requires the creation of onboard intelligent systems supporting the crew in the course of solution of problems of the first and the second global control levels.

2. STRUCTURE OF OPERATOR ACTIVITY ON BOARD TECHNICAL ANTHROPOCENTRIC OBJECT AND CLASSIFICATION OF ONBOARD INTELLIGENT SYSTEMS OF ITS SYSTEMGENERATING CORE

The activity of the operator (crew) on board the anthropocentric object is represented in terms of the following components (the algorithm of the activity of the operator (the AlAcOp)).

The operator makes decisions on online problem, realizes its solution, and participates in different tracking operations as the element of the tracking system (Fedunov 2002a). All information necessary for operator activity is giving to him on the information-control field (ICF) of the crew cockpit and/or is giving to him via cockpit voice devices. The realization of decisions and participation in tracking operations is performed by the crew via control organs of the ICF. All elements of the operator activity are represent in the integrated way by the graph of operator decisions. The estimation of feasibility of the whole volume of this activity described in operator decisions graph is performed using the computer system "operator decisions graph estimation" (Abramov et al. 2006; Fedunov et al. 2006b).

Let us consider the capability of estimation of the time necessary for the operator for each component of his activity.

Each operator decision is related to one of the following types: π-decisions (perceptive-identification), ρ-decisions (speech-mental), and π-ρ-decisions (heuristic) (Fedunov 2002).

Each π-decision is characterized by the instantaneous reaction of the operator for a certain signal stimulus.

Time expenses of the operator on making such a decision consist of the time spent for detection and recognition of the corresponding signal-stimulus.

These decisions are represented in the operator decisions graph:

(i) by the composition of information or speech message (signal-stimulus) which are necessary for the operator to make decision;

(ii) by the output information: composition and sequence of manual operations necessary for realization of the received decision by the operator.

For the estimation of the time of the perception and the comprehension of the information by the operator the information is represented by the set of the symbols that are indicated on the display of of the ICF.
The time estimates necessary for it are introduced into the computer system named "The estimation of the operator decisions graph" from (Fedunov 2002a).
Each ρ-decision is characterized in the operator decisions graph by the following:

- input information including information of the information-control field of the cockpit according to which the operator should make this decision; the composition and duration of speech message, that is used in decision making, is informing to the operator by the cockpit speech device ;
- the decision structure is described by the number and composition of online perception units according to which the decision is made; the composition and sequence of elementary acts of the decision, that had maked, are described in terms of the indication symbols in the displies of the information-control field;
- output information are represented by the composition and sequence of manual operations that are necessary for realization of this decision.

The necessary time for the adoption and the realization of this decision is estimating into the system "The estimation of the operator decisions graph".
Each π–ρ-decision is heuristic. Upon design of the operator activity, it is characterized in the operator decisions graph as follows:

- (i) input information represented by the composition of information in the ICF of the cockpit according to which the operator should make this decision; the composition and duration of speech message, that is used in decision making, is informing to the operator by the cockpit speech device ;
- (ii) the necessary time for the adoption and the realization of this decision is estimated experimentally;
- (iii) output information is characterizing by the composition and sequence of the manual operations that are necessary for the realization of this decision.

The operator activity algorithms that connect with his participation in tracking processes are described in a rather general form at the stage of development of specifications of onboard algorithms. For estimation of the time that the operator had used on the tracking process the following assumptions are maded. It is assumed that the operator is working in discrete-continuous regime. He execution of tracking operations and after they had ended he is occupied with the making and realization of decision (decisions). After the decision (decisions) was realized, the operator comes back to the tracking process again. He eliminates the error of the tracking. This error accumulated in the time when the operator is taking decision (decisions). Time instants of operator diversion to tracking operations cannot break the process of the decision making and the process of the realization of this decision. The time, that the operator is spending on the tracking process, is represented by the dependence $\tau_{track} = f(\tau_{div})$, where τ_{track}.is the time of the removal of the tracking error by the operator, τ_{div}.is the time when the operator is taking decision (decisions). The operator decisions graph can contain several tracking types, each of these types is characterized by the own separate dependence $\tau_{track} = f(\tau_{div})$. The tracking processes can be nested each in other.
Finally, all above elements of operator activity are united by the conceptual model of operator behavior, online change of this model by the operator in the course of his activity requires certain time. This time is characterized by one quantity for all conceptual models.
For the development of the situation control block, whose description will be given below, it is required in the considered typical situation:

- the develop of the operator decisions graph for this typical situation;
- then use this graph to estimate the time that necessary for the operator to realize it;
- separate the semantic component of the operator decisions graph to realized it in the situation control block;
- determine the parameters of time delays for the supply of control signals, that produced in the situation control block, to the particular situation control blocks.

2,1. The example of the development of the fragment of the operator decisions graph.

Let consider the working the operator of the Antr/object (object O-N1) when the object O-N1 meets with the object O-N2. The object O-N1 intends the counter-action to the object O-N2. The object O-N1 have the counter-action of two types: PR1 and PR2. The operator solves this problem on the basis of the information that present him on the disply of the ICF (Figure 1).

If distance D between object O- 1 and O- 2 more D1= const, then operator of the object O- 1 must fix the counter-action PR1. Otherwise operator must fix the counter-action PR2. All of this work the operator makes on background of the spying for the mark (1) on the display (Figure 1).

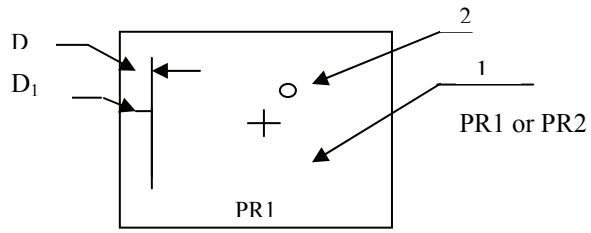


Figure 1: The display on the ICP.
D - the current distance between object O-N1 and the object O-N2.
D1 - const
1 - a cut-off signal PR1 or PR2,
2 - the mark, that is including operator in the tracking process

We compose of the fragment of the graph of the decisions of the operator (the ODG) and calculate on him temporary expanses of the operator for the execution of this work (Figures 2).

There are the following elements of the activity of the operator in the fragment of the ODG. The time for the execution of those elements are took up from (B. E. Fedunov. 2002a).

The AlAcOp-1. The ρ-decision: finding two segments (on the range scale: the segment between O-2 and O-1 - the left scale on fig.1) and the symbol (1) on the indicator. The time for it is 0.4 s. The elementary act of the adoption of this decision: comparison the lengths of two segments (D-0; D1-0). The time for it is 0.7s. The general time for the execution of the AlAcOp-1 is 0.4 + 0.7 = 1.1 s

The AlAcOp-2. The π-decision: finding PR1 in the indicator. The time for it is 0.2 s. The general time for the execution of the AlAcOp-2 is 0.2 s

The time for the realization of the AlAcOp - R1 decision is = 0.5 s.

The time for the realization of the AlAcOp - R2 decision is = 0.5 s.

The temporary expenses of the operator on the tracing process is determined on the approximation of the experimental dependency $\tau_{track} = f(\tau_{div})$:

$\tau_{track} = 0.5 + 0.2\ \tau_{div}$ if $0 \le \tau_{div} \le 1.5$ s; $\tau_{track} = 0.5 + 0.2\ 1.5 + 2\ (\tau_{div} - 1/5)$ if $1.5 \le \tau_{div}$.

The time of the tracing process for the dot 1.1 of the ODG is 0.8 +2(2.3 - 1.5) = 2.4 s.

The time of the tracing process for the dot 2.1 of the ODG is 0.5 +0.2 1.1 = 0.72 s.

The time of the tracing process for the dot 2.2 of the ODG is 0.5 +0.2 0.2 = 0.54 s.

The moment of the passing of the signal on the on-board executive devices from the beginning of the producing of the decision the AlAcOp-1(the delay of the signal) is 2.3 + 2.4 = 4.7 s.

The moment of the passing of the signal on the on-board executive devices from the beginning of the producing of the decision the AlAcOp-2(the delay of the signal) is 1.1 + 0.2 +0.72 + 0.54 = 2.56 s.

On the base of such material the block of situation management (the BSC) is developed for the system of the simulation modeling of the work of the BOSES TS

3. CLASSIFICATION OF ON-BOARD INTELLIGENT SYSTEMS

According to (Fedunov 2010) let us briefly describe onboard intelligent systems supporting the process of solution of problems of global control levels I and II (GLC-I and (GLC-II).

Let us consider the solution of problems of the GLC-I by the crew: online prescribe the current purpose at the execution operation session (online prescribe the typical situation). The motivation for this prescription cannot be formalized completely. The smaller part of these motives is weakly structured, and the larger part cannot be even verbally indicated. For making such decisions the crew uses heuristic π-ρ- decisions. These decisions of the crew are supported by the information from the information model of external and onboard situation. This model are shown the crew on the ICF of the cabin. It is created by the onboard intelligent information system "Situation awareness of the crew"(IIS SAofC) (Gribkov et al.2010).

Problems of the GLC-II, as a rule, are solving by the operator which is using ρ-decisions and π-decisions. This makes it possible to elaborate "onboard online advisory expert systems for typical situations of operation sessions" (BOSES TC) for these problems. BOSES TC supply the crew with the recommendations about the method of the achieving of the prescribed current purpose (problems of the GLC-II) [(Fedunov 1996; Fedunov 2002a).

4. THE CLASSIFICATION OF SIMULATION SYSTEMS FOR TESTING AND DEVELOPMENT OF KNOWLEDGE BASES OF ONBOARD INTELLIGENT SYSTEMS OF THE CORE OF AN ANTHROPOCENTRIC OBJECT.

Development of onboard intelligent systems has the following three stages:

(i) designing of the algorithmic shell of the intelligent system for the problems of the seted domain of a certain class of anthropocentric objects;

(ii) filling algorithmic shell of intelligent system by particular knowledge of the seted domain. As a result, the basic sample of intelligent system had oriented on the generalized (as a rule, most "rich") onboard information environment of anthropocentric objects of this class;

(iii) the adaptation of the basic sample of the intelligent system to the onboard information environment of the particular anthropocentric object of this class.

As a result, the adapted sample of the intelligent system is obtained. At the stages of creation of the basic and adapted samples of the onboard intelligent systems of the anthropocentric object, it is necessary to test and

develop their knowledge bases with the help of professional operators. For this purpose the simulation complexes with elements of real onboard systems and full scale information-control fields of the crew cockpit are creating (Fedunov 2002b). The development of these simulation complexes requires large financial and labor costs, and their using for the development of the basic intelligent systems is difficult both due to large duration of work with the basic sample and the limited possibility of the participation of the highly qualified operators. Therefore the computer simulation systems are elaborated for the modeling of the work the onboard intelligent systems. At these systems the work of the crew (operator) of the anthropocentric object simulation with the computer programs in a number of cases.

Two classes of computer simulation systems of are developed on the basis of the technical documentation of the simulated anthropocentric object:

- the simulation system of the GLC-I for the IIS with necessary inclusion of professional man- operator in the simulation loop,

- the simulation system of the GLC-II for each BOSES TC with simulation of the work of the professional man-operator by the situation control block (the computer program).

5. COMPUTER SIMULATION SYSTEM OF FOR DEVELOPMENT OF KNOWLEDGE BASES OF INTELLIGENT INFORMATION SYSTEMS "SITUATION AWARENESS OF THE CREW"

For the onboard intelligent system "Situation awareness of the crew" (IIS SAoC) the specific features of the simulation system are the presence the following components:

- mathematical models of the onboard measurement devices of the anthropocentric object and the onboard digital computer algorithms of the information processing;
- the simulation on the computer display of information frames of the crew cockpit that were designated for the situation awareness;
- the representation of the dynamics of the outward and inward situation;
- the simulation on the computer keyboard of the control organs on the anthropocentric object and its onboard equipment used by the crew in solution of problems of the GLC-I;
- the presence of the real operator in the simulation loop.

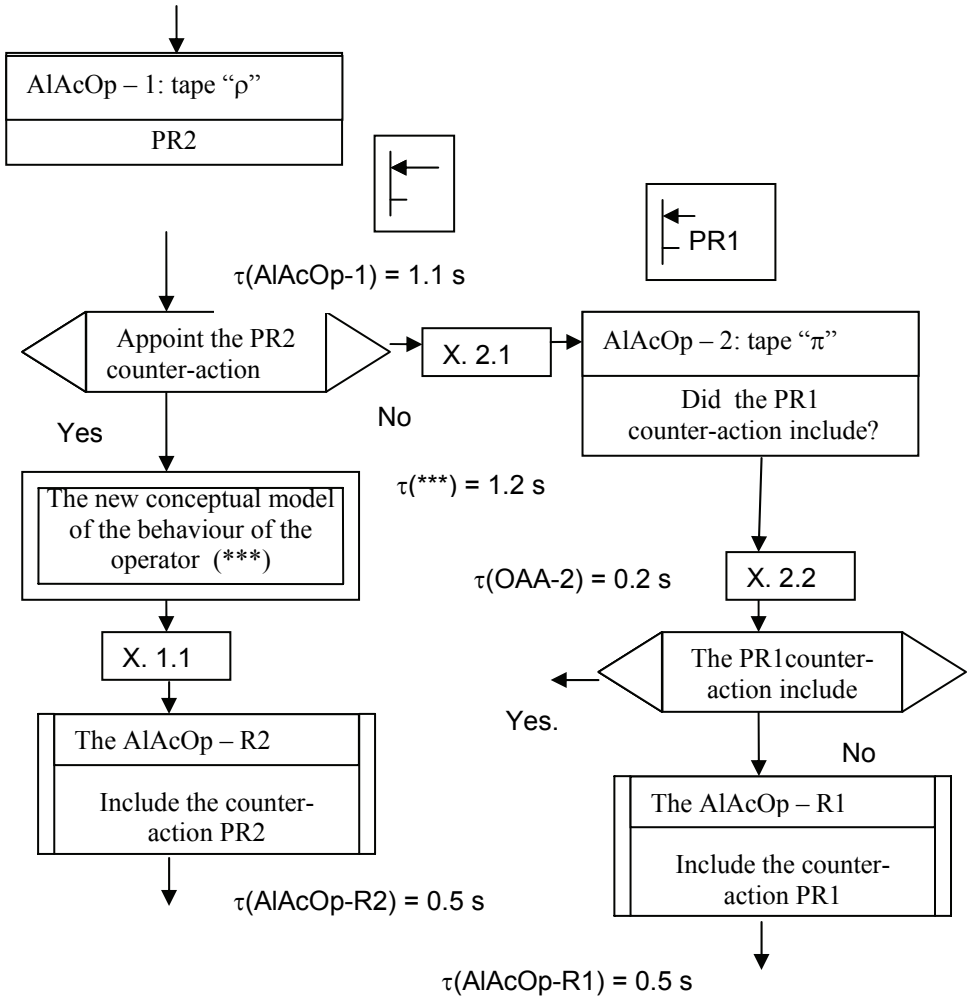


Figure 2: The fragment of the graph of the operator decisions.

That system will be called the simulation systems for the GLC-I. Note, out of the main types of solutions by the crew of problems of the GLC-I are heuristic solu tions follow that the presence of the operator in the simulation loop is mandatory.

Upon simulation by the simulation system for the GLC-I the correctness of assignment by the operator of typical situation, convenience of perception and comprehension of the model of external and onboard environment presented in the ICF are estimated.

6. THE COMPUTER SIMULATION SYSTEMS FOR THE DEVELOPMENT AND THE TESTING OF YHE KNOWLEDGE BASES OF THE INTELLIGENT SYSTEMS PROVIDING THE CREW WITH THE SOLUTION TO PROBLEMS OF THE GLOBAL CONTROL LEVEL II (Romanenko A. V. et al. 2010.

The solution of problems of the GLC-II, where the operator is not assuming heuristic solutions, is executed by the BOSES TS without a participation of the man-operator. The structure of such BOSES TS and the technology of their development were discussed in ((Fedunov 2009).

For the testing and the development of the knowledge bases of the BOSES TS the simulation system for the corresponding typical situation is designed. Here the activity of uhe man-operator is represented by the mathematical situation control block (the BSU).

6.1. The example of the functional blocks of the simulation system for typical situation are shown in Figure 3.

We will not describe all blocks, specified on Figure 3, but will describe only block of the situation management (BSU).The BSU simulates the work of the operator at the choice type situation, in which the operator does not use the heuristic decisions. The BSU is developed on the base of the ODG (shown on the figure 2) with preliminary cut-in in it the estimations of the temporary expenseses of the operator on each his AlAcOp.

The BSU includes itself:

- the choice of the current information for the conditions of the output on the actuated branch of the ODG:

- the definition of the manual acts of the operator in this branch ,

- summing of the temporary expenseses of the operator in the actuated branch of the ODG (from the begining of the actuated branch) for the definition of the delay of the control signal, that was produced in the actuated branch of the AlAcOp - R№*** (where AlAcOp - R№*** is any branch of the ODG on the figure 2);

- a delay of the sending the control signal on the on-board executive devices (the control signal was produced by the actuated branch the AlAcOp - R№***).

For the figure 3 the BSU is designed on the basis the ODG, that shown on the figure 2.

CONCLUSIONS .

The practice of the applications and the developments of the modern anthropocentric objects require the creation of on-board intelligent systems of two classes for its cores:

- intelligent information system "Situation awareness of the crew" (IIS SAofO), that is creating information support of the crew for the online prescription of the current goal of the operation session;
- the onboard online advisory expert systems for typical situations (BOSES TS), that present the crew with the method for the achieving of the current goal of the operation session in real time.

For testing and development of knowledge bases of these systems two classes of computer simulation systems are developed:

- the simulation systems for the first global control level (the simulation systems for the GLC-I) with necessary inclusion of professional man-operator in the simulation loop;
- the simulation system for the second global control level (the simulation systems for the GLC-II) for each BOSES TS with simulation of operation of professional man-operator by the situation control block. The simulation systems for the GLC-II allow to estimate and improve the knowledge base of the BOSES TS without the direct participation of qualified professional operators, and leave this resource for the stage of the development of the adapted BOSES TS.

These simulation systems are being developed on the basis of the technical documentation of the simulated anthropocentric object, that include itself the information about the information and control components about the information-control field, about the onboard measurement devices, about the onboard executive devices and about the composition and structure of the onboard digital computer algorithms.

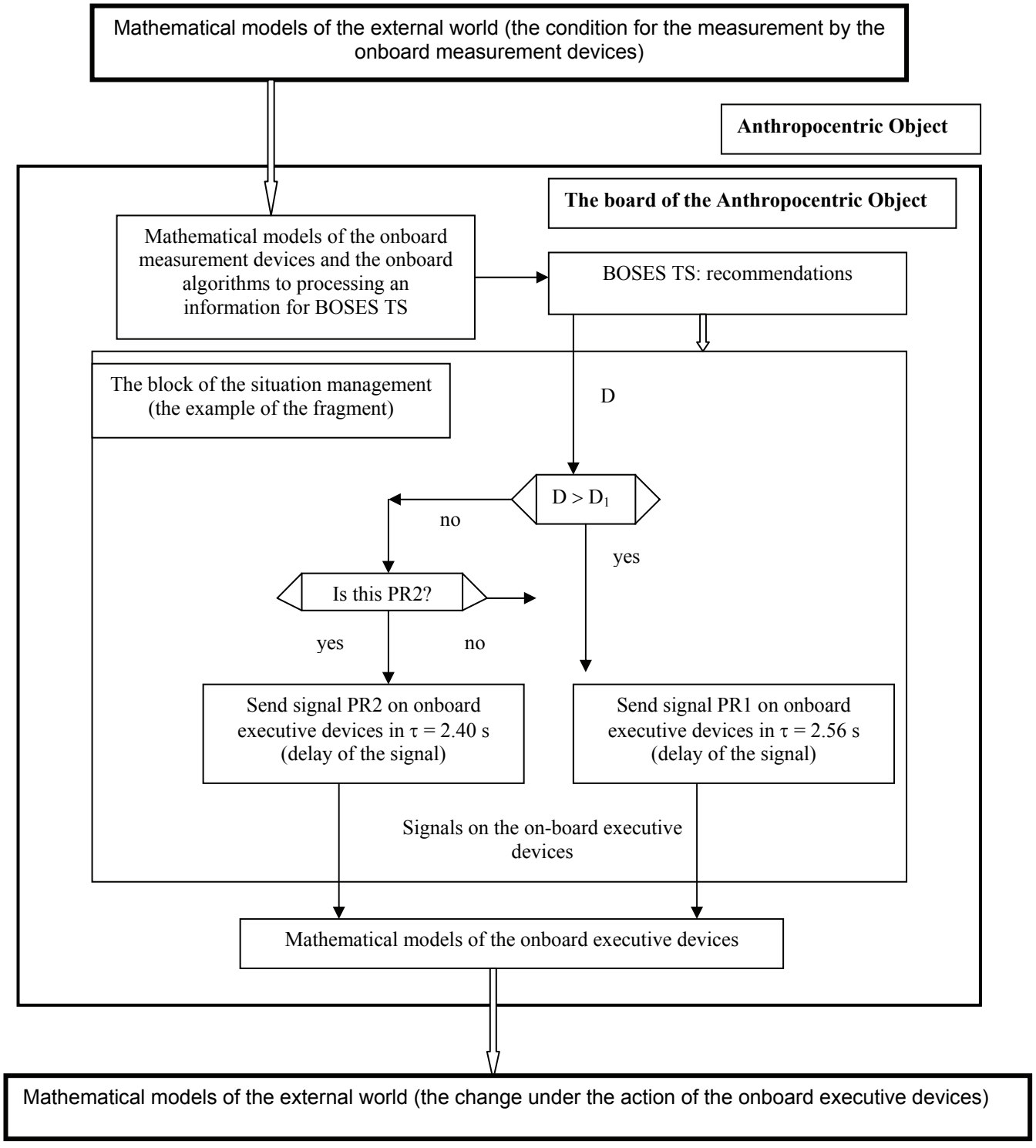


Figure 3: The system of simulation modeling of the work BOSES TS with replacement the Anth/object operator by the block of the situation management

REFERENCES

Abramov A. P., Vydruk D. G. and Fedunov B. E. 2006. A Computer System for Evaluating the Realizability of Algorithms of Crew Activity," Izv. Ross. Akad. Nauk, Teor. Sist. Upr., No. 4, 122–134 (2006). [in Russian]. [the English version journal is "Comp. Syst. Sci." 45 (4), 623–626 pp.].

Gribkov V.F., Fedunov B.E. 2010. The On-board information intellectual system "Situational privity of the crew combat plane". In the book of "Intellectual managerial system". By edit akad. RAN S.N.Vasil'ev. Izd. Machine building. 108-116 pp. [in Russian].

B. E. Fedunov. 1996. "Problems of the Development of On Board RealTime Advisory Expert Systems for Anthropocentral Objects," Izv. Ross. Akad. Nauk, Teor. Sist. Upr., No. 5 (1996). [in Russian]. [the English version journal is "Comp. Syst. Sci." 35 (5), 816–827 pp.

B. E. Fedunov. 2002a. "Technique of Estimating the Realizability of the Graph of Operator Decisions of an Anthropocentric Object when Designing Algorithms of Onboard Intelligence," Izv. Ross. Akad. Nauk, Teor. Sist. Upr., No. 3 (2002). [in Russian]. [the English version journal is "Comp. Syst. Sci." 41 (3), 437– 446 pp.]

Fedunov B. E. 2002b. *Onboard Online Advisory Expert Systems of Fifth Generation Tactical Aircrafts (Survey of Foreign Press)* (NITs GosNIIAS, Moscow, 2002) [in Russian].

Fedunov B.E.,Vidruk D.G. 2006 "Computer system of a crew activity algorithms realizability estimation. Modeling Methodologies and Simulation", "Key Technologies in Academia and Industrial." In Proceedings of 20th European Conference on Modeling and Simulation (ECMS2006).

B. E. Fedunov. 2009. "Basic Algorithmic Shell of Onboard RealTime Advisory Expert Systems for Operation Situations Typical for an Object," Izv. Ross. Akad. Nauk, Teor. Sist. Upr., No. 5, 90–101 (2009). [in Russian]. [the English version journal is "Comp. Syst. Sci." 48 (5), 752–764].

Fedunov B.E. 2010."Intelligent Support of the Crew On Board of Anthropocentric Object", "Mechatronic, Automation, Control", No 2, 62-70 pp. (in Russian).

Romanenko A. V., Fedunov B.E.2010. The computer simulation systems for the testing of the knowledge bases the onboard intelligent systems of the core of an anthropocentral object. Izv. Ross. Akad. Nauk, Teor. Sist. Upr., No. 6 (2010). 102-121. [in Russian]. [the English version journal is "Comp. Syst. Sci."]

AUTHOR BIOGRAPHIES

BORIS E. FEDUNOV, doctor of technical sciences, professor, had graduated the Moscow Aviation Institute (State University of Aerospace Technologies) in 1960 as aircraft designing engineer, Moscow State University in 1965 as mathematician. Now he works at the State Research Institute of Aviation Systems and delivers a lectures in Moscow Aviation Institute on syntheses of on-board algorithms systems of piloting aircrafts and applied systems analysis. He is the member of the scientific council of the Russian Association of an Artificial Intellect (collective member of European Coordination Committee on the Artificial Intellect)

Authors' Index