THE 17TH INTERNATIONAL CONFERENCE ON MODELING AND APPLIED SIMULATION

SEPTEMBER 17 – 19, 2018
BUDAPEST, HUNGARY

EDITED BY
AGOSTINO G. BRUZZONE
FABIO DE FELICE
CLAUDIA FRYDMAN
FRANCESCO LONGO
MARINA MASSEI
ADRIANO SOLIS

PRINTED IN RENDE (CS), ITALY, SEPTEMBER 2018

ISBN 978-88-85741-09-6 (Paperback)
ISBN 978-88-85741-07-2 (PDF)

ISBN 978-88-85741-09-6 (Paperback)
ISBN 978-88-85741-07-2 (PDF)

THE 17TH INTERNATIONAL CONFERENCE ON MODELING AND APPLIED SIMULATION, MAS 2018

SEPTEMBER 17 - 19 2018

BUDAPEST, HUNGARY

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

HUNGARIAN ACADEMY OF SCIENCES CENTRE FOR ENERGY RESEARCH

AUTONOMOUS UNIVERSITY OF BARCELONA

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

UNIVERSITY OF BORDEAUX

UNIVERSITY OF CYPRUS

DUTCH BENELUX SIMULATION SOCIETY

I3M 2018 INDUSTRIAL SPONSORS

CAL-TEK SRL

LIOTECH LTD

MAST SRL

SIM-4-FUTURE

I3M 2018 Media Partners

Inderscience Publishers – International Journal of Simulation and Process Modeling

Inderscience Publishers – International Journal of Service and Computing Oriented Manufacturing

IGI Global – International Journal of Privacy and Health Information Management (IJPHIM)

Halldale Media Group: The Military Simulation and Training Magazine

Halldale Media Group: The Journal for Healthcare Education, Simulation and Training

SAGE
Simulation Transaction of SCS

De Gruyter
International Journal of Food Engineering

MDPI - Sustainability

Euromerci: The Italian Monthly Logistics Journal

EDITORS

Agostino G. Bruzzone
MITIM-DIME, University of Genoa, Italy
agostino@itim.unige.it

Fabio De Felice
University of Cassino, Italy
defelice@unicas.it

Claudia Frydman
LSIS, France
claudia.frydman@lsis.org

Francesco Longo
DIMEG, University of Calabria, italy
f.longo@unical.it

Marina Massei
DIME, University of Genoa, Italy
massei@itim.unige.it

Adriano Solis
York University, Canada
asolis@yorku.ca

THE INTERNATIONAL MULTIDISCIPLINARY MODELING AND SIMULATION MULTICONFERENCE, I3M 2018

GENERAL CO-CHAIRS

AGOSTINO G. BRUZZONE, *MITIM-DIME, UNIVERSITY OF GENOA, ITALY*
MIQUEL ANGEL PIERA, *AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN*

PROGRAM CO-CHAIRS

FRANCESCO LONGO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
YURI MERKURYEV, *RIGA TECHNICAL UNIVERSITY, LATVIA*

THE 17TH INTERNATIONAL CONFERENCE ON MODELING AND APPLIED SIMULATION, MAS 2018

GENERAL CO-CHAIRS

ADRIANO SOLIS, YORK UNIVERSITY, CANADA
MARINA MASSEI, *LIOPHANT SIMULATION, ITALY*

PROGRAM CHAIR

FABIO DE FELICE, *UNIVERSITY OF CASSINO, ITALY*

MAS 2018 International Program Committee

Naamane Aziz, *Laboratoire des Sciences de l'Information et des Systèmes, France*
Edgar Alonso Lopez-Rojas, *Norwegian University of Science and Technology, Norway*
Eleonora Bottani, *University of Parma, Italy*
Youssef Bouanan, *Lab Gretha Université Bordeaux*, France
Agostino G. Bruzzone, *University of Genoa, Italy*
Alessandro Chiurco, *University of Calabria, Italy*
Gerson Cunha, *LAMCE COPPE UFRJ, Brazil*
Fabio De Felice, *University of Cassino, Italy*
David Del Rio Vilas, *Ferrovial Services, Spain*
Marco Frascio, *University of Genoa, Italy*
Claudia Frydman, *LSIS, France*
Yilin Huang, *Delft University of Technology, Netherlands*
Janos Sebestyen Janosy, *Centre For Energy Research Hungarian Academy Of Sciences, Hungary*
Andreas Körner, *Technische Universität Wien, Austria*
Pasquale Legato, *University of Calabria, Italy*
Francesco Longo, *MSC-LES, University of Calabria, Italy*
José M. Cecilia , *Universidad Católica San Antonio, Spain*
Marina Massei, *Liophant Simulation, Italy*
Rina Mary Mazza, *University of Calabria, Italy*
Letizia Nicoletti, *Cal-Tek Srl, Italy*
Antonio Padovano, *University of Calabria, Italy*
Horacio Emilio Pérez Sánchez, *Universidad de Murcia, Spain*
Antonella Petrillo, *University of Cassino, Italy*
Maria Celia Santos Lopes, *LAMCE/COPPE/UFRJ – Brasil*
Peer Olaf Siebers, *University of Nottingham, UK Georgios Ch*
Adriano Solis, *York University, Canada*
Giuseppe Vignali, *University of Parma, Italy*
Gregory Zacharewicz, *IMS Université Bordeaux 1, France*

Tracks and Workshop Chairs

Simulation for Strategic Engineering
Co-Chairs: Agostino G.Bruzzone, Simulation Team Genoa, University of Genoa (Italy); Francesco Longo, MSC-LES (Italy); Ian Mazal, Brno University of Defence (Czech Republic); Marina Massei, Liophant Simulation (Italy); Anna Franca Sciomachen, DIEC (Italy)

Inventory Management Simulation
Chairs: Adriano Solis, York University, Canada; Letizia Nicoletti, Cal-Tek Srl, Italy

Production Systems Design
Chair: David del Rio Vilas, Ferrovial Services, 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, Centre for Energy Research Hungarian Academy of Sciences (HAS), Hungary

The Interplay between Behavioural Game Theory and Agent-Based Social Simulation
Chairs: Peer-Olaf Siebers, Nottingham University, UK;

Simulation Interoperability in Enterprises
Chairs: Gregory Zacharewicz, MS Université Bordeaux 1 (France) ; Youssef Bouanan, Lab Gretha Université Bordeaux (France)

Control and monitoring of Hybrid Renewable Energy Sources and Systems for buildings
Chair: Naamane Aziz, Laboratoire des Sciences de l'Information et des Systèmes, France

Simulation Based Design
Chair: Yilin Huang, Delft University of Technology, Netherlands

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

Workshop on virtual and augmented reality
Chairs: Gerson Cunha, LAMCE/COPPE/UFRJ, Brasil; Maria Celia Santos Lopes, LAMCE/COPPE/UFRJ, Brasil

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 Católica San Antonio, Spain; José M. Cecilia Universidad Católica San Antonio, Spain

Workshop on Serious Games applied to Security, Crisis Management and Safety
Chairs: Francesco Longo University of Calabria, Italy; Agostino G.Bruzzone, University of Genoa, Italy

Human-centred and Human-Focused Modelling and Simulation (common track MAS-EMSS)
Chairs: Agostino G. Bruzzone, DIME, University of Genoa, Italy; Peer Olaf Siebers, University of Nottingham, UK

Towards a universal framework for the Intelligent Social Simulation-based Systems (ISSS)
Chair: Antonio Padovano, University of Calabria, Italy

Welcome Message 2018

After 16 successful editions, we are pleased and proud to announce The 17th International Conference On Modeling and Applied Simulation. Thanks to this regular appointment, a tight-knit community has been established along the years and a wealth of knowledge and experiences has been build up. Year by year, this community has been able to grow and to provide valuable R&D contributions influencing actual and future trends in the way M&S is perceived and deployed in different application areas.

Nowadays, M&S potentials are well clear among non-practitioners and M&S pervasiveness has grown dramatically. It opens the doors to new scientific endeavors and new advances that The International Conference On Modeling and Applied Simulation is eager to support and encourage. This is why from year to year the community of researchers, practitioners and scholars has put great efforts to raise new expectations working very hard toward greatest scientific achievements in terms of M&S approaches, paradigms, technologies as well as innovative prototypes.

This is why over time MAS has become a well-established and influential event even more attractive and growing in scope. In the general framework of The International Multidisciplinary Modelling & Simulation Multiconference, where MAS is co-located, all the attendees will have the possibility to get in touch with the latest advances of M&S methodologies, techniques and applications for Business, Finance and Security, Logistics, Supply Chain Management Production, Design, Training, etc. A great effort has been done to provide a rich technical program as well as social arrangements hoping that MAS 2018 could be a successful and fruitful experience. To this end, we would like to express our deepest gratitude to the Program Chairs for their excellent support in organizing the conference program, the Program Committee for thoroughly and timely reviewing the papers, the Local Organizing Committee members, that worked extremely hard for the conference organization and social events, the authors and all the participants to whom we wish a scientifically rewarding and a pleasant experience in Budapest, the Pearl Of Danube.

Marina Massei
University of Genoa,
Italy

Adriano Solis
York University
Canada

ACKNOWLEDGEMENTS

The MAS 2018 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 2018 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.

I3M 2018 INTERNAL STAFF

MATTEO AGRESTA, *SIMULATION TEAM, ITALY*
AGOSTINO G. BRUZZONE, *DIME, UNIVERSITY OF GENOA, ITALY*
ALESSANDRO CHIURCO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
RICCARDO DI MATTEO, *SIMULATION TEAM, ITALY*
JESSICA FRANGELLA, *CAL-TEK SRL, ITALY*
CATERINA FUSTO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
LUCIA GAZZANEO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
FRANCESCO LONGO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
MARINA MASSEI, *DIME, UNIVERSITY OF GENOA, ITALY*
LETIZIA NICOLETTI, *CAL-TEK SRL, ITALY*
ANTONIO PADOVANO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
ANTONIO REDA, *CAL-TEK SRL, ITALY*
CATALDO RUSSO, *CAL-TEK SRL, ITALY*
KIRILL SINELSHCHIKOV, *SIMULATION TEAM, ITALY*
CARMINE TOTERA, *CAL-TEK SRL, ITALY*
MARCO VETRANO, *CAL-TEK SRL, ITALY*
BEATRICE ZACCARO, *SIMULATION TEAM, ITALY*

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

Index

DRONE BASED INSPECTION SERVICES IN INDUSTRIAL CONTEXTS: RISK ASSESSMENT AND MARKET OPPORTUNITIES

Giuseppe Aiello[(a)], Fabrizio Hopps[(b)], Mario Venticinque[(c)], Domenico Santisi[(d)]

[(a),(b)] Università di Palermo – Dipartimento Di Innovazione Industriale e Digitale - DIID
[(c)] Institute for Agriculture and Forestry Systems in the Mediterranean of the National Research Council of Italy (CNR-ISAFOM).
[(c)] Leonardo Engineering and Research s.r.l.

[(a)]giuseppe.aiello03@unipa.it, [(b)]fabrizio.hopps@unipa.it, [(c)]ventimario@inwind.it, [(d)]d.santisi@leonardoengineering.net

ABSTRACT
Unmanned Aerial Vehicles (UAVs) is an emerging technology with the potential of introducing disruptive innovation in a large number of industrial and civil applications. In particular, many studies have highlighted the achievable advantages for “three Ds” (i.e., dull, dirty, or dangerous) missions, which actually represent their natural market niche. A huge potential market in such sense is constituted by high-risk industrial installations including Oil&Gas, chemical, power generation, shipbuilding, etc. where significant improvements can be achieved in terms of safety and ergonomics. This paper proposes a comparison between traditional inspection approaches and innovative drone-based services in high risk industrial contexts, with the objective of obtaining a clear picture of the safety risks involved.

EXTENDED ABSTRACT
Unmanned Aerial Vehicles (UAVs), commonly known as drones, is a new technology that can significantly outperform existing solutions in many commercial and industrial contexts such as surveillance, firefighting, inspection of industrial plants as well as agriculture, logistics, disaster recovery, etc.. In particular, UAVs are often preferred for “three Ds” (i.e., dull, dirty, or dangerous) missions, which actually represent their natural market niche. As a consequence of the important potential benefits in these business areas, a significant global market growth is inevitably expected in the near future, involving Europe, the US and China in a role of market leaders. A recent economic analysis of the drones value chain reported in the European Drones Outlook Study by SESAR (2016) reveals a potential European market exceeding 10 billion euros annually by 2035 and further growing past 15 billion euros annually by 2050. In particular, value added services ("Drones-As-A-Service") are expected to represent the largest market opportunity, since the high number of typologies of UAVs and the recent advances in sensing and monitoring technologies, make the landscape of possible applications extremely ample and variegated. Referring to the industrial context, UAVs can efficiently perform various tasks (e.g. inspections, monitoring, surveillance, etc.), drastically lowering the exposure of human operators to safety risks and health threats. A huge potential market in such sense is constituted by high-risk industrial installations including Oil&Gas, chemical plants, power generation, shipbuilding, etc. where significant improvements can be achieved in terms of safety and ergonomics.
Coherently with this objective, this paper proposes a comparison between traditional inspection approaches and innovative drone-based services in high risk industrial contexts, with a specific focus on the issues related to safety. In fact, industrial plant structures and equipment are constantly monitored through frequent periodic inspections within appropriate maintenance plans, because a failure can be an extremely dangerous event with catastrophic hazardous consequences. Nevertheless, several major industrial accidents occurred in the past, such as the ConocoPhillips, UK in 2001 and, more recently, the Caribbean Petroleum terminal in Bayamon, Puerto Rico in 2009 and the Chevron, US in 2012. A detailed analysis of recent accidents occurred in EU and OECD countries, has been recently issued by the Joint Research Centre of the European Commission (Wood et al, 2013). The report takes into account more than 400 accidents and reports material degradation as a frequent cause of failure, attributing to poor maintenance the main responsibilities. To avoid failures, the risks related to material degradation such as corrosion in metallic structures, are strictly monitored. Industrial plants, hence, normally require frequent inspections, which are generally carried out by external service providers. Although safety is the most important element decsion makers take into account when selecting a service provider in critical industrial contexts, in some cases, it is not easy to have a clear perception of the risks, particularly when a new technology is employed. In this paper a detailed risk assessment approach is proposed referred to traditional (scaffolding and rope access) and drone based inspection services. Traditional inspection methods are a well-known cause of hazards, mainly related to the working at height condition. The risk associated to these operations, are generally considered

very high due to the relatively high probability of occurrence and to the severity of their consequences severe consequences like fatalities, injuries or illnesses.
Determining the risk profile of an inspection is more complex when dealing with drone-based operations, since this is a new technology and the regulatory system, particularly in the EU, is still under development and frequently updated. In addition EU regulation concerning the safety of flying UAVs is extremely fragmented, since it relies on national regulatory systems which do not benefit from mutual recognition. This clearly constitutes a strong limitation from a market point of view since it prevents operators to perform EU-wide activities. In the following, the risks related to drone operations will be discussed with reference to the Italian regulation, which follows the same regulatory principles of many other countries in EU, although some slight differences may apply. According to the regulation, the general risk profile of a drone system is related to few main hazards, namely: midair collisions with other aircrafts, and ground impact with people or structures. The most frequent causes of hazard are failures, human errors and environmental conditions. In the following paragraph such elements are discussed considering the risk assessment procedures prescribed by the regulations.

System reliability and criticality of operations

The reliability UAVS results from their technical features, and requires a preliminary distinction between Remotely Piloted Aircraft Systems (RPAS), and unmanned autonomous systems (UAS). The term drone", which stands for Dynamic Remotely Operated Navigation Equipment and is commonly used to broadly address RPAS. According to the regulations, a risk score (C_{RPA}) is assigned to the type (manual or automatic) of flight control. The risk profile of RPAS is generally considered less critical compared to UAS, thus risk contribution of 0.1 is assigned in case of manual operations and 0.5 in case of automatic operations. The risk contribution of drones, falling in the class of RPAS, is thus equal to 0.1.
The other preliminary element of risk assessment is related to the characteristics of the operations performed by means of the drone. Safety regulations, in fact, typically distinguish between critical/non-critical operations and offensive/inoffensive drones. The critical nature of operations depends on their location: those operations conducted in areas where a midair collision is impossible and an impact on the ground cannot cause fatal injuries to people or severe damage to the infrastructures are classified as “non critical”. Non-critical operation hence do not involve, even in the event of system failures and malfunctions, overflights of urban areas and infrastructures, restricted areas, transport systems and industrial plants. Non critical activities, in addition, must be performed within "V70" air volumes, at an adequate safety distance from congested areas, in daylight conditions, and outside airfield traffic zones (ATZ). Such activities must be conducted in visual line of sight (VLOS) and at a minimum distance of 8 km from the perimeter of an airport and from the paths of approach/take-off. The difference between critical and non-critical operations is important since they undergo a different authorization process. For non-critical operations the assessment of operational risks is demanded to the operator, who must simply submit a declaration of compliance to civil aviation authority (ENAC in Italy). Critical operations, on the contrary, require a specific authorization from the aviation authority, which is granted only after satisfactory assessment of the related risks.
A further distinction must be made between offensive and non-offensive drones. Such classification is related to the weight of the UAV and to other technical features such as: maximum wing surface, maximum wing loading, etc. ENAC regulation applies to vehicles with a total weight less than 150 kg, and classifies them in Very Light (300 g <MTOW< 4kg), Light (4 kg <MTOW< 25kg) kg, Heavy (MTOW > 25kg). Critical operations with UAVs under 25 kg over urban areas have been recently allowed, provided specific safety conditions are met (ENAC, 2016), however flying over groups of people remains prohibited in any case. On the contrary, operations with RPAS whose maximum take-off mass is less than or equal to 2 kg are always considered non-critical, provided that the RPAS design ensures its inoffensive nature, as assessed by ENAC. Recently a new regulatory framework has been issued for RPAS with MTOW less than 300g (mini drones), which are considered intrinsically inoffensive. These systems can operate freely in urban or industrial contexts, provided they do not overfly crowded areas.
According to the regulations, hence, flying a RPAS in an industrial plant must be considered critical operation, unless an intrinsically inoffensive drone is employed.

Ground impact

Ground impact (g.i.) risk refers to the possibility of a drone crashing on humans or structures on the ground. The probability of fatalities in such case is closely related to probability of impacting persons, which, in turn, depends upon the population density in the area of operations. In particular, in non-populated areas the contribution to risk related to human impact can be considered null, while in populated areas, the probability of impacting people is calculated by the following equation:

$$E(fatalities) = N_{exp} P(fatality|exposure) \quad (1)$$

Where N_{exp} is the number of people exposed to the crash event, which assuming a uniform population density, can be calculated as the product of the expected crash area (A_{exp}) by the population density (ρ).

$$N_{exp} = A_{exp} \cdot \rho \quad (2)$$

For the determination A_{exp} the regulations refer to 3 zones: the area of operations, the buffer area, and the adjacent area. The area of operations is the area directly interested by the flight plan, the buffer area is a safety distance around the area of operations, and the adjacent area is the area which is not involved in the RPAS crash event, unless in case of uncontrolled flight. The determination of the safety (buffer) area is a topic frequently addressed in the literature, and several theoretical and empirical models have been proposed by researchers. The theoretical models generally applied in the industrial practice, are based on different assumptions: steep descent, uncontrolled glide and parabolic fall. In the steep geometric model (Weibel and Hansman, 2003, Clothier and Walker, 2006; Dalamagkidis et al 2008), the aircraft instantly loses its lift (e.g. due to a failure in the wings or in the propulsion system) and the crash area can thus be approximated by the frontal area of the aircraft augmented by a small buffer to account for the width of an average human person (approx. 0,25m, see for example Waggoner, 2010). The Uncontrolled gliding model (Lum and Waggoner, 2011) assumes a total loss of power on a fixed wing aircraft, meaning loss of thrust and control. Since the airframe is intact, the aircraft starts gliding at some angle γ, typically dependent upon the lift to drag ratio of the air vehicle. The impact area thus consists of a rectangle as wide as the wing span of the vehicle ($W_{aircraft}$) and as long as the descent from the top of a theoretical standing person's head to the point of impact. Finally, a third theoretical model is obtained considering a parabolic fall trajectory from the apogee with an initial horizontal velocity V_{0X}. In such case, the UAV will maintain its initial velocity until the ground crash (neglecting the air friction). The radius of the area interested by the fall event ca thus be calculated according to the formulas reported in the following table.

Model	Radius Formula
Free fall	$A_{exp}^{steep} = R_{UAV} + R_{human}$
Uncontrolled glide	$A_{exp}^{glide} = (W_{UAV} + 2R_{human})\left[L_{UAV} + 2R_{human} + \frac{H_{person}}{\tan(\gamma)}\right]$
parabolic	$A_{exp}^{parabolic} = V_{0x}\sqrt{\frac{2h}{g}}$

Table 6 – theoretical ground impact models

ENAC (2015) regulations for critical operations prescribe the use of flight termination system designed to allow the pilot to completely shut down the RPAS power system in order to minimize the risk of uncontrolled flight. For RPAS, hence, the buffer area can be calculated according to following formula, based on a parabolic fall model augmented by the time required for the pilot to activate the flight termination system (3s):

$$r = V_{0x}\sqrt{\frac{2h}{g}} + V_{0x}T \quad (3)$$

For example if we refer to a commercial mini drone such as the DJI Spark, which is one of the smallest system currently available on the market, the following parameters are set: V_{0x}=max horizontal velocity (including maximum wind speed)=20 m/s; h=50 m; $R_{UAV}=0.2\ m$, $R_{human}=0.3m$, $\gamma = 30°$ (for a quadcopter)and Length (L_{UAV})= $0.143m$. The resulting radius of the buffer area is approx. 125m. Considering also the operations area, the total potential crash area may easily extend for a radius of approx. 200m from the center of operations. In order to evaluate the number of people exposed to the ground impact risk, the density of population in this area must be estimated. This is generally done referring to institutional/statistical data of the territory involved. In case of nonhomogeneous population density areas, the area is subdivided in homogenous circular sectors and the weighted sum is calculated. Finally, the degree of protection provided by existing structures (e.g. buildings) in the area overflown is taken into account by means of a shelter factor ranging from 0,1 (industrial areas) to 1 (no obstacles.). In the case the drone is operated into an industrial area, and no houses or offices are present in a range of 120m from the area of operations. The corresponding ground risk can therefore be considered null. In addition, as a prevention measure the safety area can be fenced and interdicted to workers, so that no human operators (except the pilot) will be present in the area.

mid-air collision risk

The second critical hazard of UAV operation is related to midair collisions, which refers to the case of an Unnamed Aerial System (*UAS*) facing a Manned Aircraft (MA) in a potential collision trajectory. Quantitative risk assessment models to evaluate likelihood of a mid-air collision event based on the gas particle model are dated back in the 70's although more recent formulations can be found in the literature (Lum and Waggoner, 2011). Applying this model for UAV risk analysis in practice is considerably complicated, and rarely required in practice, because in operations performed within V70 airspaces, the possibility of mid-air collisions is negligible. This is generally the situation when the drone is employed to inspect the surface of a structure in an industrial plant, unless in case of inspecting tall towers or chimneys. To ensure operations within V70 airspaces, a retention cable may be employed in order not to exceed the maximum altitude. In such situations, the midair collision risk can be actually considered negligible.

Human errors and environmental conditions

Finally an additional source of risk can be represented by human factors and environmental conditions. In order to control such risk factors, the regulations prescribes that pilots attend a specific training in order to get qualified for operations. According to the risks related to environmental conditions, it will be part of the risk assessment process to select the dates of inspection taking into account the wind speed and the external temperature. The operations will thus be scheduled in dates when environmental conditions are compatible with the limits prescribed by the RPAS manufacturer, and adequate for operating in safety.

Level of acceptable risk

Once the hazards have been analyzed, and the related probabilities of occurrence have been determined, the risk assessment procedure for RPAS prescribes a final step based on Target Safety Standard (TLS) approach, well established in the civil aviation. The methodology consists in comparing the risk evaluated for a specific mission with a reference maximum acceptable value defined as TLS, generally referred to the likelihood of the worst possible outcome (fatality). The metric used to measure the TLS is fatalities per flight hour (FH), which is coherent statistics already in use by the Authorities. For small RPAs the Italian regulation considers an acceptable risk of 10^{-6} per FH referred to the ground impact, and an acceptable risk of 10^{-6} per FH. According to this method, the mission risk, is calculated as the event probability multiplied by its impact. In the case of ground impact, the mission risk is the multiplication of the RPAS falling in the expected crash area, multiplied by the number of people present in the area:

$$R_c = P_c N_{exp} \tag{4}$$

The Safety Objective (*SO*), associated to the acceptable mission risk is thus calculated as:

$$SO = \frac{E_c}{P_{GI}} C_{SAPR}, \text{ for the case of Ground Impact} \tag{5}$$

$$SO = \frac{E_c}{P_{MAC}} C_{SAPR}, \text{ for the Midair Collision} \tag{6}$$

Where *Ec* is the expected casualty measured in number of victims per flight hours (10^{-6} acceptable risk), and C_{sapr} is equal to 0.1 and 0.5 for manual and automated operations respectively.

The Safety Objective (*SO*) must finally be compared to the acceptable probability of a top (catastrophic) event (*PTE*), which is considered equal to 1 unless evaluated empirically by means of a statistical analysis, and, if the following condition is not verified specific mitigation measures must be enforced (e.g. parachute, retention cable, etc.)

$$\boldsymbol{SO > PTE} \tag{7}$$

Clearly, as stated before, the employment of mini-drone with a MTOW below 300g and all the features that make it inoffensive in a V70 airspace, both the probabilities related to ground impact and the midair collision are negligible, as discussed before, hence condition (7) is always satisfied and the corresponding risk can be considered null.

Explosion risk

Finally, when operating in industrial areas, particular attention must be paid to the issues related to explosion risk. It is hence mandatory that the drone is compliant with the Atex regulations, which ensures it does not constitute into a potential source of ignition in presence of flammable gases. This is actually a substantial limitation since the RPAS currently available on the market generally are not ATeX compliant, except for some rare exceptions.

Results and Discussion

The objective of the study is to investigate the possibility of actually performing drone based inspections in high-risk industrial plants. This inspection system, in fact, can lead to a significant risk reduction in for the inspection workers, and a cost effective alternative compared to traditional inspection techniques. In order to meet such objectives, a first issue that must be investigated concerns the employment of an intrinsic inoffensive drone with MTOW less than 300g. which can easily be found in the market. The employment of such system drastically simplifies the risk assessment process, thus speeding up the authorization process from the Civil Aviation Authority (ENAC in Italy). Concerning the ground impact risk, it depends on the location where the inspection activities are carried out, and, specifically, on the density of population in the operations and buffer areas. Such risk is generally low, for industrial inspection activities, since the access to the operation area can be restricted, and the population density in the safety range is generally low. In particular, this risk can be realistically null if the plant is located in an isolated industrial area. Finally, the midair collision risk, is related to the presence of a manned aircraft in a potential collision trajectory with the drone. If the inspection activities are carried out by means of an inoffensive drone, this risk can be neglected, while in other cases the employment of a retention cable ensures operations are carried out in the established airspace volume. The last element to consider is the explosion risk, which requires the drone to be compliant with the Atex regulation. This condition, in fact, can hardly be met by an inoffensive drone. Some systems are actually present on the market, but their costs and operating features frequently make their operations impossible or non convenient. The conclusion, hence, is that the risks related to a drone based inspection can actually be negligible compared to traditional approaches, although the Atex regulation drastically limits the areas of operations for commercial drones. The employment of an Atex compliant inoffensive drone could however solve the problem.

Another important element to take into account is related to the capability of a drone inspection system of obtaining reliable information about the condition of the structures inspected, and the comparability of the results obtained by the different inspection methods. Traditional inspection method involve a team of skilled operators that examines the structure from a scaffold, generally performing a preliminary visual inspection of the surface. Subsequently, more advanced tools such as Phased Array Ultrasonic Technique or Low Frequency Electromagnetic Technique (LFET), can be applied if required. Information gathered from a visual inspection (which still is probably the most important of all non-destructive tests) may hence give a preliminary

indication of the condition of the structure and allow the formulation of a subsequent more detailed inspection program.
An inspection carried out by drone system, instead, is capable of providing a digital image of the structure inspected, but no additional information can be determined for example about the wall thickness. The digital image however can be multi-spectral thus allowing to analyze the heat distribution on the surface, or, for example, the presence of leaks. The digital information obtained by a drone inspection, in addition, can easily be archived for further reference, and it can post-processed into ortho-photos and digital 3D models with textured surfaces. Such information however can be effectively exploited within decision making processes related to the management of the structures.
The drone based inspection, hence, does not provide the same results, of traditional inspection techniques, and the quality of the information obtained is thus comparable only to a limited extent. In particular,, it must be highlighted that the accuracy and the quality of information obtained by a drone inspection is largely dependent upon flight-path and the features (e.g. resolution) of the camera employed. The flight-plan in fact should be designed with the aim of keeping the drone as close as possible to the structure, provided all the safety prescriptions are respected. For such reason the flight path must always be discussed by the service provider and the facilities manager. The imagery obtained by a drone based inspection can be finally post-processed in a typical photogrammetric workflow including setting ground control points, point cloud and digital surface model generation, and image ortho-rectification and mosaicking. An example of the results obtainable by such approach is given in the following figure 1, which has been obtained by post-processing the data gathered in an experimental inspection on the storage tanks of an abandoned industrial site. The system employed is a commercial DJI Spark mini drone, equipped with a weight reduction kit to achieve an extremely low (300g) maximum takeoff weight (MTOW), which makes in inoffensive. This drone comes equipped with a camera featuring a wide-angle lens with 25mm equivalent focal length allowing for stabilized video at 1080p/30 frames per second, and 12MP still images. Finally, The Spark mini drone has a Maximum Flight Time of approx. 16 minutes, a GPS satellite positioning system and is capable of reaching a maximum speed of 14 m/s. Another significant limitation of this system is represented by the environmental conditions, since the operating temperature range of 0-40°C and the maximum wind speed must not exceed 10 m/s. The minimal dimensions of this system also allow it to fly in small environments (e.g. into a tank or inside an indoor area) which would definitely be challenging task for bigger drones with less control sensitivity.

Figure 1 - 3D model of the site inspected

In order to highlight the presence of critical situations, for example related to corrosion, advanced post-processing methodologies can be employed. The topic of corrosion detection from images, in fact, has been recently investigated by several researchers and different image processing techniques have been proposed mainly based on Texture analysis and filtering, edge detection, image segmentation, etc.

Conclusions

Drones is an emerging technology with the potential of introducing disruptive innovation in a large number of industrial and civil applications. This paper in particular focuses on the safety issues related to drone based inspection services in industrial plants. For such purpose, the safety risks related to the traditional approaches and the drone based inspection have been compared. The results show that the new technology has the potential to drastically reduce the safety risks, but the achievement of such objective, significantly depends upon the specific features of the site inspected, and upon the technical features of the drone employed. In particular the employment of an inoffensive drone may significantly simplify the authorization process, thus speeding up the operations. The drawback is that the limited operating range of such systems may hamper the inspections due, for example, to the environmental conditions. In addition the video capturing devices which typically equip such systems are quite limited, due to the necessity of keeping the weight low. The most critical issue, however, concerns the compliance with the Atex regulations for explosion risks. In a critical industrial context such as oil and gas, for example, the drone could in fact ignite an explosion if it is not Atex compliant. Considering that the commercial offer of Atex compliant drones is very limited and expensive, this is probably the most significant element to take into consideration.
Another important issue is related to the quality of the information acquired by means of the drone based inspection service, which depends upon the capabilities of the sensing devices installed. The devices currently installed on commercial systems nowadays may easily allow for the generation of detailed 3d textured models, which can be stored and analyzed anytime. However, the identification of incipient failures or critical situations which are still in at an experimental stage, requires high quality images to provide precise and reliable results, and current commercial devices can be inadequate for such purposes. Generally speaking, the state of the art of commercial systems nowadays still lacks of adequate

software and hardware features to allow effective decision support.
Finally, additional problems that nowadays hamper a substantial market spread of drone based inspection services is related to the regulatory system, which, particularly in the EU, mainly relies on national regulations which do not benefit from mutual recognition.
Nevertheless drone technology is undergoing significant improvements, and we may expect in the near future that new aerial vehicles and sensing devices will benefit of enhanced capabilities to better support human operators in decision making. Current research trends are also focused on autonomous decision making and cooperative control strategies. Such features may turn common drones into highly efficient cyber-physical system (CPS) capable of operating autonomously or in swarms in dangerous situations (search, rescue, surveillance, fire prevention, anti-terrorism, etc.). This will also allow to raising the safety standards of many industrial operations.
Concerning the monitoring devices, the possibility of prototyping new advanced sensor systems would eventually lead to the possibilities of combining data sets, from different sources in a unified representation, exploiting the potential of modern data fusion methodologies. In year 2017, for example, the deployment of the world's first UT (Ultrasonic Thickness Testing) integrated UAV system has been announced (see for example http://www.texodroneservices.co.uk/) although no experimental evidences have been provided. The value of archiving the information obtained, however, should also be considered in terms of evidences to provide to third parties and local authorities when necessary.
In conclusion, the study proposed demonstrated that the drone technology can actually be the most economic and safe solution for monitoring industrial plants. However, there are still important shortcomings that limit the spread of such technologies and the market opportunities particularly in high-risk industrial contexts.

AUTHORS BIOGRAPHY

Prof. Giuseppe Aiello is an associate professor at the Department of Industrial and Digital Innovation of the University of Palermo in Italy. He is expert of Logistics and supply chain management and has a significant background in industrial engineering, including layout design, plant management and logistics in several industrial contexts including naval, chemical, power generation and Oil&Gas. The most relevant research topics concern the solution of complex optimization problems involving uncertain information by means of heuristic and meta-heuristic approaches. Recently, the research interests are mainly focused the development of decision support systems (DSS) and control rooms based on multisensor data fusion for industrial applications and on Cyber-Physical Systems supporting human operators in difficult or dangerous tasks. He has also been responsible of research projects granted by national and international institutions. In cooperation with enterprises for the development of prototypal devices for industrial applications.

Dr. Fabrizio Hopps is a PhD in aeronautical engineering and an expert professional in the field of Occupational Health, Safety and Ergonomics. He is a consultant for the Italian National Institute for Insurance against Accidents at Work, with a specific interest towards research activities and new technology for safeguarding workers against physical injuries and occupational diseases. He has also a significant experience on industrial risk assessment in high risk industrial contexts. He has also a researchers and professor of industrial safety at the Department of Industrial and Digital Innovation of the University of Palermo in Italy. His research topics are mainly focused on risk assessment and evaluation as well as multi criteria decision making methods.

Dr. Mario Venticinque is an electronic engineer with a consolidated experience in the development of electronic systems and devices based on innovative technologies. He is a consultant for many national and international enterprises and a researcher Institute for Agriculture and Forestry Systems in the Mediterranean of the National Research Council of Italy (CNR-ISAFOM). He is expert in the development of embedded industrial systems involving sensors and monitoring devices. Recently he has also cooperated with research institutions for the development of smart technologies and wireless sensor networks (WSN) for industrial applications. He has also been a team leader for several research projects granted by national and international institutions.

Dr. Domenico Santisi is a PhD in Aeronautics and Aerospace engineering with a substantial background of high risk industrial contexts, occupational safety, and plant management. He is also a former Navy officer, and a still active airline pilot and instructor with more than 40 years' experience in the civil and military aviation. He has also gained relevant skills in the development of manned and unmanned aerial vehicles by participating the development and test stages of prototypes of fixed and rotary wing vehicles. In his career he has cooperated with important national and international companies and research institutions including the Massachusetts Institute of Technology, and several Universities all over the world. More recently, his professional activity has focused on the field of Remote Piloted Aircrafts (RPAs), where he is currently operating as a pilot and instructor. He is also committed in the development of innovative services and technologies for the drone industry in cooperation with several international companies and institutions.

REFERENCES

SESAR Joint Undertaking, 2016, European drones outlook study. Unlocking the value for Europe.

www.sesarju.eu/sites/default/files/documents/reports/European_Drones_Outlook_Study_2016. pdf

Wood, M.H, Vetere Arellano, A.L. and Van Wijk, L., 2013, Corrosion-Related Accidents in Petroleum Refineries: lessons Learned from Accidents in EU and OECD Countries, EU 26331 EN

ENAC, 2017, Ente Nazionale per l'Aviazione Civile, Remotely Piloted Aerial Vehicles Regulation Issue No. 2 dated 16 July 2015, Revision 3 dated 24 March 2017

Weibel R.E., Hansman R.J., 2003, Safety considerations for operation of small unmanned aerial vehicles in civil airspace. In: MIT joint university program quarterly meeting.

Clothier, R.A., and Walker, R.A., 2006, Determination and Evaluation of UAS Safety Objectives. Proceedings 21st International Unmanned Air Vehicle Systems Conference, 18.1–18.16.

Dalamagkidis, K., Valavanis, K.P., and Piegl, L.A., 2008, Evaluating the Risk of Unmanned Aircraft Ground Impacts. 16th Mediterranean Conference on Control and Automation, June 25–27, IEEE,709–716.

C. W. Lum and B. Waggoner, 2011, A Risk Based Paradigm and Model for Unmanned Aerial Systems in the National Airspace, in Proc. AIAA Infotech@Aerospace Conference, St. Louis, MO, 1–31.

ENAC - Ente Nazionale per l'Aviazione Civile, 2016, aeromobili a pilotaggio remoto con caratteristiche di inoffensività 2016: https://www.enac.gov.it/La_Normativa/Normativa_Enac/Linee_Guida/info148190721.html

Waggoner, B., 2010, Developing a Risk Assessment Tool for Unmanned Aircraft System Operations, in Department of Aeronautics and Astronautics, University of Washington.

ENAC, Metodologia di valutazione del rischio in operazioni RPAS per autorizzazioni e permessi di volo non geografici, Edizione 1, 10 dicembre 2015.

METHODS OF CONSTRUCTING THE SYSTEMS OF MONITORING THE TECHNOLOGICAL PROCESS EXECUTION

V. A. Kholopov[(a)], A. G. Shmeleva[(b)], E. N. Kashirskaya[(c)], E. V. Kurnasov[(d)]

[(a),(b),(c),(d)]Moscow Technological University (MIREA), Moscow, Russia

[(a)]holopov@gmail.com, [(b)]annshmeleva903@gmail.com, [(c)]liza.kashirskaya@gmail.com, [(d)]kurnasov@mirea.ru

ABSTRACT

The paper describes the development of practical methodology for constructing the systems of monitoring the technological processes execution in digital machinery production. The methodology is primarily designed for controlling the machining technological processes in the metal-cutting workshop. In order to implement the system of monitoring the execution of technological processes in low-volume/ high-variety production the paper sets the tasks of defining the initial system condition, identifying the most significant factors, evaluating the probability of completing technological operations, modeling the flow of each technological process in the workshop. The paper also justifies the applicability of the expert evaluation method, the Pareto method, and neural network usage for developing the system. A flow chart of the simulation modeling of monitoring the technological process implementation is suggested; the chart demonstrates the interrelation between the methods of analyzing the technological processes described in the paper. Basing on the simulation modeling results it is possible to draw the conclusion that all technological processes will be completed during the work shift and also to identify the possible causes of failures before the production cycle is complete.

Keywords: process automation; technological process; simulation model; monitoring; expert evaluation method; Pareto analysis; neural networks; forecasting; process state graph; mechanical engineering

1. INTRODUCTION

The development of software for automation of high-tech equipment control in digital machinery production (Romanov, Romanov, Kharchenko, and Kholopov 2016; Kashirskaya, Kurnasov, Kholopov, and Shmeleva 2017; Kashirskaya, Kholopov, Shmeleva, and Kurnasov 2017) identified a problem connected with controlling the technological processes of machining details in a metal-processing workshop. The successful execution of such technological processes in low-volume/ high-variety production is very uncertain when several technological processes are simultaneously carried out on the same set of production equipment. This uncertainty complicates the organization of planning the workshop load and further control over the technological processes carried out in the workshop since it is impossible to guarantee the planned execution of all technological processes on the site. In order to achieve efficient control over the workshop performance in digital machinery production the authors suggest constructing a system of monitoring the completion of technological processes. Several tasks need to be accomplished in order to construct such a system. This paper considers the following four tasks:

- first task – identifying the combination of factors and the degree of their influence on the faults and failures of the technological equipment and the technological operations;
- second task – identifying the most important factors that influence the execution of technological operations of a specific technological process;
- third task – evaluating the probability of implementing the technological operations of this technological process taking into account the whole combination of all technological operations performed in the workshop;
- fourth task – modeling the process and calculating the probability of executing each technological process in the workshop taking into account the probability of completing technological operations.

Different methods of solving the above mentioned tasks were considered in works (Kashirskaya, Kurnasov, Kholopov, and Shmeleva 2017; Kashirskaya, Kholopov, Shmeleva, and Kurnasov 2017), and the choice of the most efficient ones was rationalized. This paper considers the justification of their application for developing a system of monitoring the execution of the technological process.

2. METHOD OF EXPERT EVALUATIONS OF THE INITIAL SYSTEM CONDITION

The initial data for solving the first task is the parameters of the initial condition of the system that includes the technological equipment and the executed technological processes. The first priority goes to identifying the combination of all factors influencing the faults and failures of the equipment. An efficient method of initial study of heterogeneous data, structuring and identification of the reasons causing emergency situations is the expert evaluation method. The idea behind the method is to identify the factors influencing the execution of technological operations and to prioritize them. For each piece of technological equipment expert evaluations form the largest possible list of factors influencing the execution of the technological operation. After that, the probability of their occurrence is evaluated.

The main stages of the method include:

- forming a group of experts,
- studying the characteristic of the technological operations (roadmaps),
- forming an array of factors,
- expert evaluation of the factors,
- identifying the factors that have the most influence on how the technological operations are carried out,
- expert evaluation of the probability of faults in the technological operations if the identified factors occur,
- forming the probability matrix for the occurrence of emergency situations.

The expert evaluation method is highly dependent on the professional qualities of polled experts and implies thorough verification of their competency during the process of forming the expert group.

In order to form the array of factors it is expedient to use non-connected ranks, i.e. give strictly ranged evaluations without repeating the values. This will make it possible to identify the important and secondary parameters in the system of monitoring the implementation of the technological process. The methodology of expert evaluation implies forming a weighted conclusion based on the detailed study of the characteristics of the technological operations (Kashirskaya, Kurnasov, Kholopov, and Shmeleva 2017).

The precision of factor evaluation depends on the character of expert's statements. It should be taken into account that opposite opinions within the expert group significantly reduce the precision of the conducted evaluation. In order to assess the consistency of opinions the concordance coefficient is used, which – not taking into account the connected ranks – is calculated using the formula:

$$W = \frac{12\sum_{i=1}^{n}\left(x-\bar{x}\right)^2}{m^2\left(n^3-n\right)} \quad (1)$$

where m – number of experts; n – number of objects; x – the total number of scored points; $\bar{x}$ – the mean average of all scored points.

The concordance coefficient varies between 0 and 1. The concordance coefficient sufficient for accepting the solutions suggested by the expert group should be above 0.7, meaning that at least 70% of the expert group members should be unanimous in their evaluations.

Basing on the expert evaluation method the authors suggest using the procedure of forming the maximum list of factors influencing the execution of the technological operation (Fig. 1).

Table 1 provides a variant of the result data array for one technological operation when using the expert evaluation procedure.

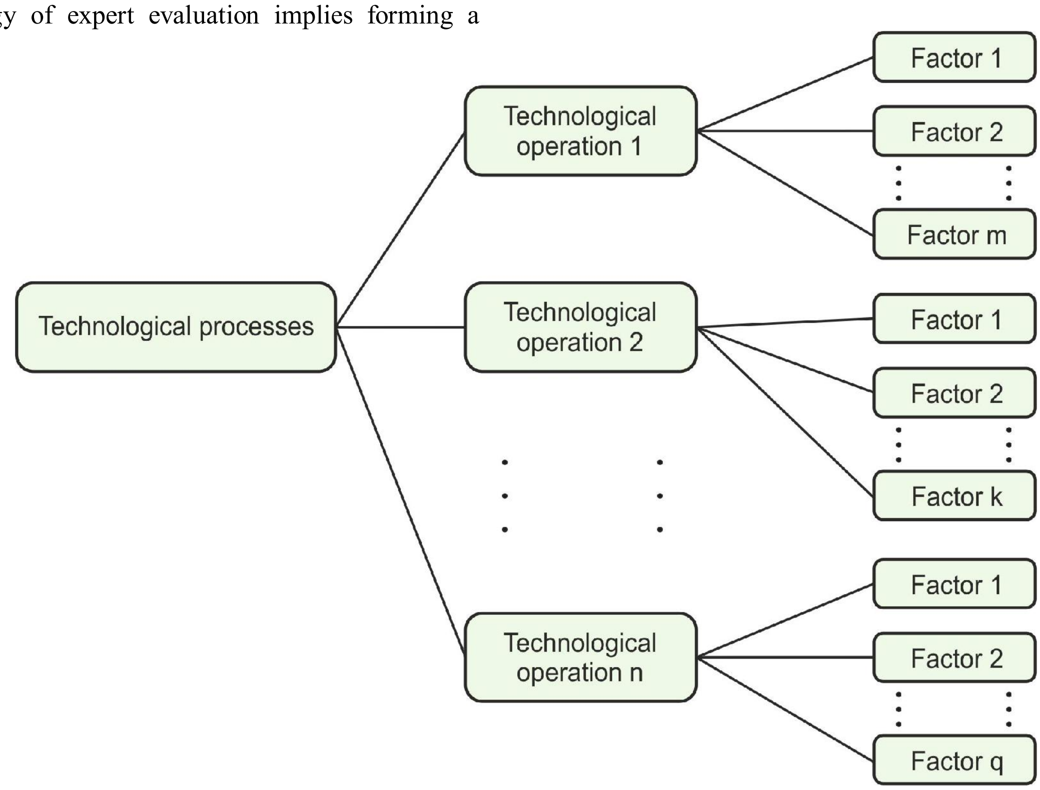


Figure 1: Forming a list of factors

Table 1: Data received after completing the expert evaluation for one technological operation

Number of objects							4		Number of experts		8
	E1	E2	E3	E4	E5	E6	E7	E8	Total number of scored points	Mean average of a factor	Probability %
Factor 1	1	1	1	1	1	1	1	1	8	1.00	10
Factor 2	3	3	2	3	3	2	3	3	22	2.75	28
Factor 3	4	4	4	4	4	4	4	2	30	3.75	37
Factor 4	2	2	3	2	2	3	2	4	20	2.50	25
	Mean average of all scored points								20		
	Concordance coefficient								0.78		

The expert evaluation method that has been considered in details in (Kashirskaya, Kurnasov, Kholopov, and Shmeleva 2017) allows forming an array of factors, assessing the influence of each of them, identifying the most important ones, ranging them and defining the probabilities of operation failures in the technological process.

3. STATISTICAL ESTIMATION OF THE MOST IMPORTANT FACTORS

The universal method for solving the second task – evaluation of the most important factors causing emergency situations – is the Pareto analysis (Kashirskaya, Kurnasov, Kholopov, and Shmeleva 2017) that is used in various technical systems (Desineni, Berndlmaier, Winslow, Blauberg, and Chu 2007; Kalach, Romanov, and Tripolskiy 2016; Dmitriev, Khalyasmaa, Doroshenko, and Romanov 2016). It enables identifying the main reasons why the emergency situations occur. Application of the method results in a bar chart that characterizes the most important factors of faults and failures in a technological operation according to the accumulated statistical information; the chart is a convenient tool to perceive the expert evaluation results.

The procedure of forming the minimal list of significant factors influencing the faults and failures of a technological operation when carrying out a specific technological process is provided in Fig. 2. Significant factors include: low-quality workpieces (defective items); unaccounted specifics of operating the equipment; non-critical situations solved by the operator, power supply failures; control software failures; low-quality machine equipment (defective items); improper machine assembly; machine downtime due to technical reasons, e.g. failure of the feed mechanism; improper observation of the production technology etc.

The statistical data of the registered factors (Kashirskaya, Kurnasov, Kholopov, and Shmeleva 2017) that lead to improper execution of the technological operation during the observed time period is provided in Table 2. The data gathered for the technological operation makes it possible to conduct quality analysis of the studied characteristics.

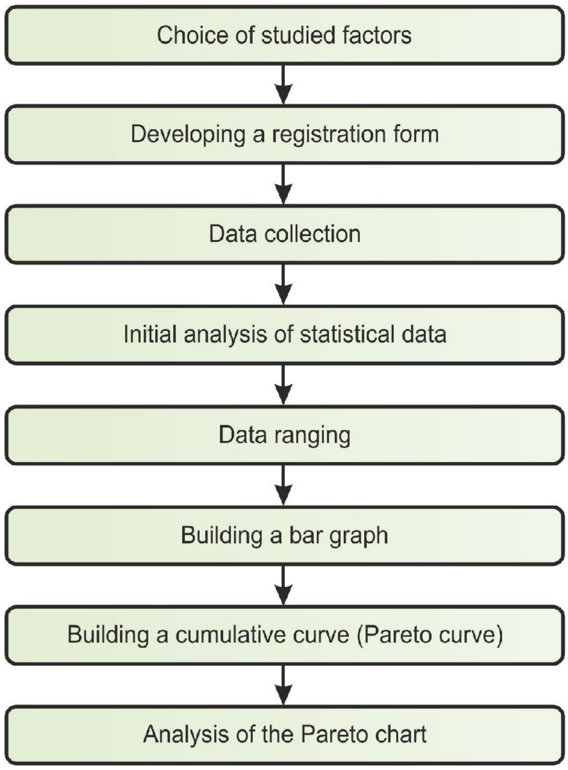


Figure 2: Pareto analysis conducted in relation to the technological operation

Table 2: Example of destructive factors evaluation

	Quantity	Failures, %	Total percentage, %
Factor 4	73	39.04	39.04
Factor 1	56	29.95	68.98
Factor 7	17	9.09	78.07
Factor 6	15	8.02	86.10
Factor 3	8	4.28	90.37
Factor 2	7	3.74	94.12
Factor 5	6	3.21	97.33
Factor 8	5	2.67	100
	187	100	

Fig. 3 shows a chart that visually reflects the Pareto principle – 20% of the most important factors have an 80% influence on the changes in the system characteristics. Therefore, the goal of using the Pareto analysis is to choose precisely these factors.

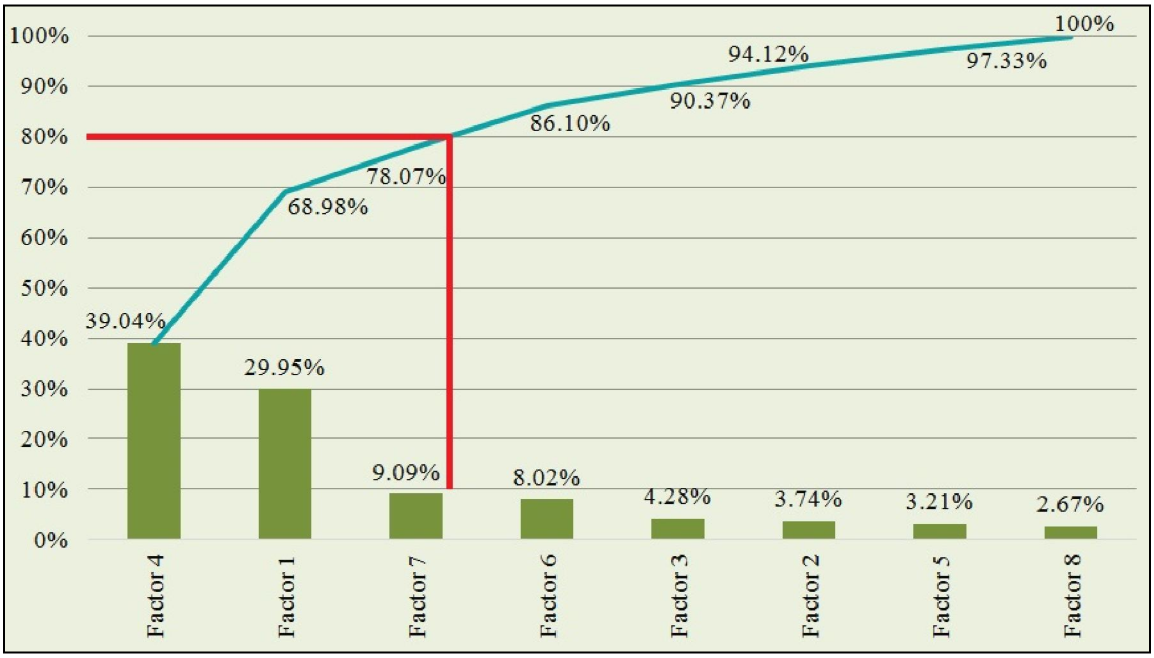


Figure 3: Pareto chart for the technological operation

Analysis of the Pareto chart makes it possible to draw the conclusion that for the considered technological operation the influential factors are 4, 1, 7. Developing the measures targeted on the elimination or reduction of these factors will make it possible to reduce the probability of emergency situations occurrence.

The Pareto analysis chooses the minimal – but sufficient for the solved task – list of factors that influence the execution of the technological operation when carrying out a specific technological process. It also provides identification of significant factors alongside with forming the array of their occurrence probability. This analysis should be carried out for each technological operation of the technological process.

4. ARTIFICIAL INTELLIGENCE METHOD

The third task can be solved using the artificial intelligence algorithms (Huang, Shiang, Jou, Chang, and Ma 2010; Jin and Sendhoff 2008; Chizhikov, Kurnasov, and Vorob'ev 2018; Chao, Peichen, and Jianping 2014) based on neural networks (Chao, Peichen, and Jianping 2014; Eroshenko and Romanov 2016).

Neural network is used to determine the probability of executing the technological operations of a specific technological process taking into account the total combination of all technological operations performed in the workshop, e.g. for any cutting technological processes used in machine-building, such as manufacturing details like plug, gearwheel, housing etc.

In order to solve this task it is expedient to use a feedforward neural network. Fig. 4 shows an example of a three-layer neural network. The network consists of five neurons, has two entries f_1 and f_2 and one exit p. The neurons are represented by circles, the synapses are represented by directed arrows. y_1, y_2, …, y_5 are the functions of non-linear transformation, generically different for each neuron. w_1, w_2, …, w_{10} are the synapses' weight. The black dots represent the branching points for entries.

The network is considered trained if all weights w_1, w_2, …, w_{10} have specific values selected in such a manner that if the input data is directed to entries f_1 and f_2, at the exit p we will receive some result that complies with our task.

Application of neural networks makes it possible to use the analysis of already existing data to forecast the possible variant of introducing some additional factors of influence, such as e.g. a new workpiece supplier, when we know the percentage of defective items but don't know the percentage of failures in the machine that uses the new workpieces in the technological process.

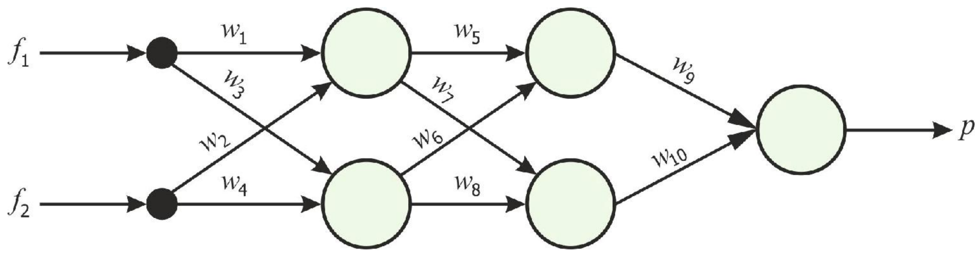


Figure 4: Three-layer neural network

Considering the neural network as a set of elements that perform some calculations on the data provided to them, we can study the solution of the main and the secondary tasks. The main task includes the functions of neural network generation, training and testing. The secondary tasks are: identifying the importance of the input signals, neural network contrasting, example training, identifying the minimal list of input parameters, obtaining a logically transparent neural network and, finally, data mining.
The neural network calculates the probability of failure for each operation of the technological process.

5. SIMULATION MODELING OF THE TECHNOLOGICAL PROCESS FLOW

Constructing a simulation model of the technological process flow and calculating the completion probability for each technological process in the workshop (fourth task) while taking into account the completion probability of separate technological operations can require the methods of system analysis. As an example of this, let us consider a system S in which a random process occurs with a finite number of discrete process states (Kashirskaya, Kholopov, Shmeleva, and Kurnasov 2017): $s_1, s_2, \ldots, s_l, \ldots,$.
States $s_1, s_2, \ldots, s_l, \ldots$ can be qualitative, i.e. described by words, for example the sequences of the technological operations, and also can be characterized by random values (random vector).
During the calculations we evaluated the capacity of system S to change from state s_i to the required state s_l directly or through other states. Directed graphs were used for the visualization of discrete states; the vertices of which corresponded with the system states. A state graph is shown in Fig. 5.

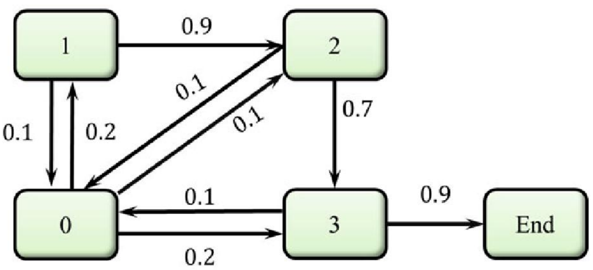


Figure 5: System state graph

The arrow pointing from vertex s_i to vertex s_l represents the possibility of the system S to change from state s_i to state s_l directly, bypassing other states. This representation makes it possible to demonstrate the technological process as consisting of successive technological operations.
The main task of the research is to find the unconditional probabilities of the system S being at any step (k) in the state s_i; i.e. the probability $p_i(k)$:

$$p_i(k) = P\{S(k) = s_i\} \quad (i = 1, 2, \ldots, n; \; k = 0, 1, \ldots,). \qquad (2)$$

In order to find these probabilities it is required to know the conditional probabilities of system S changing on the step k into the state s_j, if it is known that at the previous step (k-1) it was in the state s_i. Let us denote this probability

$$p_{ij}(k) = P\{S(k) = s_j \mid S(k\text{-}1) = s_i\} \quad (i, j = 1, 2, \ldots, n). \qquad (3)$$

The probabilities $p_{ij}(k)$ are called transition probabilities at step k.
Transition probabilities can be conveniently presented in the form of a matrix. The matrix corresponding with the Fig. 5 is written as:

$$p_{ij} = \begin{pmatrix} 0.5 & 0.2 & 0.1 & 0.2 & 0 \\ 0.1 & 0 & 0.9 & 0 & 0 \\ 0.1 & 0 & 0.2 & 0.7 & 0 \\ 0.1 & 0 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \qquad (4)$$

At the initial moment of time the system is in the state s_1, therefore: $p_0(0) = 0$, $p_1(0) = 1$, $p_2(0) = 0$, $p_3(0) = 0$, $p_{4\_end}(0) = 0$.
After making calculations we will receive the characteristics of the technological process (Kashirskaya, Kholopov, Shmeleva, and Kurnasov 2017). For example, at k=3 we will get: $p_0(3) = 0.2$, $p_1(3) = 0.03$, $p_2(3) = 0.07$, $p_3(3) = 0.1$, $p_{4\_end}(3) = 0.6$.
The model of calculating the probability of changing from k-1 technological operation to k forms the probability of implementing the technological process. The modeling procedure is iteratively repeated in real time taking into account the already completed technological operations.

6. OUTLOOK OF THE DEVELOPMENT OF MONITORING SYSTEMS CONSTRUCTION METHODOLOGY

The total combination of the considered methods comprises a practical methodology for constructing systems of monitoring the execution of technological processes. The developed system should consider the requirements to the designed database that contains possible causes of emergency situations. In order to create this database it is necessary to apply the methods of structuring and accumulating information during the expert evaluation analysis, Pareto analysis, neural network training, and basing on their results input the data into the information system (Kashirskaya, Kurnasov, Kholopov, and Shmeleva 2017). The list of causes of emergency situations that are input into the information system should comply with the standards of qualitative information characteristics.
The considered methods of analyzing the technological processes that meet the main requirements to the interaction of the elements in the information system

with regard to their interconnection can be demonstrated by the flow chart (Fig. 6).

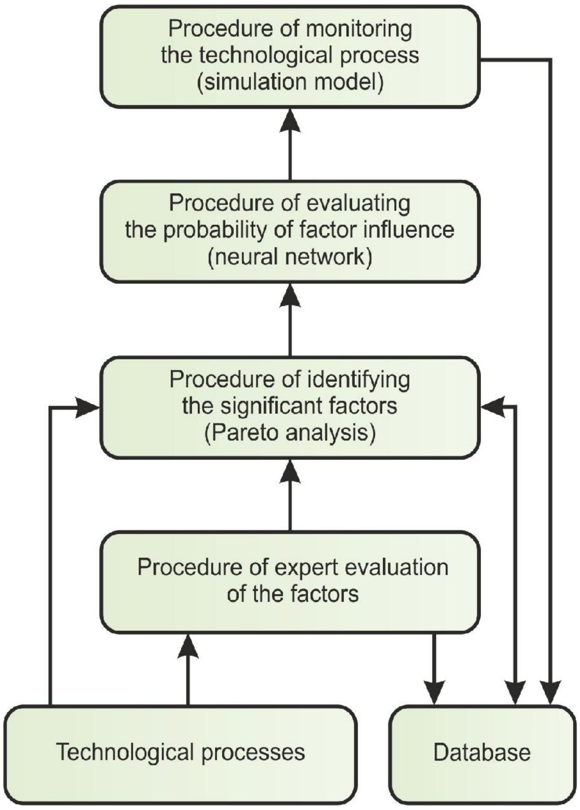


Figure 6: Flow chart of the system of monitoring the technological process implementation

The procedure of identifying the significant factors on the basis of Pareto analysis is aimed at forming a training sample for the neural network. It also enables ranging the factors, and, if needed, makes it possible to adjust the weight coefficients. Pareto analysis is designed for the identification and statistical accumulation of data that can be efficiently processed using neural networks. In our case the main task of the neural network is to forecast the probability of the disruptions in the intended execution of the technological operation. Basing on the probability analysis of the occurring factors and their inter-influence it is possible to evaluate the necessity to stop the technological operation or to take the decision of eliminating the factors upon the completion of the technological process.

Expert evaluations and statistical analysis should be used to construct a training sample. Basing on the study of factors that cause system failure: $f_1, f_2, ..., f_i, ...$ it is possible to compile the probability table of occurring factors $p_1 = p(f_1)$, $p_2 = p(f_2)$, ..., $p_i = p(f_i)$, ... that influence failures in the normal machines operation s_1, s_2, s_3, …, s_n, and the corresponding probability of failures of the normal execution of the technological operation p % under the influence of these factors (Table 3) (Kashirskaya, Kurnasov, Kholopov, and Shmeleva 2017).

Table 3: Training sample for the neural network

	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	…	p_m
s_1	0	1	1	1	2	3	4	2	…	3
s_2	0	1	2	2	0	0	4	3	…	0
s_3	0	1	1	3	3	2	2	0	…	2
…	…	…	…	…	…	…	…	…	…	…
p	0	3	4	3	1	2	8	1	…	2

The main attention should be devoted to the construction of the training sample for the creation of the database. After this it is possible to evaluate the importance of the input signals, the minimal significant list of the input parameters, conduct the neural network testing, evaluate the calculation results and also to compare the received results with the available data.

The neural network is trained using the input array and allows analysis of the factor occurrence probability, real-time tracing and evaluation of the technological operation execution and the system characteristics.

The probability of each technological process completion is calculated in order to evaluate their satisfiability by forecasting and is essentially a simulation model.

The sequence of performing the mentioned procedures (Fig. 6) is a model of a technological process monitoring system in digital machinery production.

7. CONCLUSION

The monitoring system can be developed using the expert evaluation method, Pareto analysis, neural network and simulation modeling. The production process is analyzed to form a training sample for the neural network that is designed to evaluate the occurrence probability of factors that cause failures in the technological operations. An array of possible failure variants is input into the neural network and it evaluates the percentage of favorable outcomes for the execution of all production processes. The results of simulation modeling make it possible to deduce the probability of completing all technological processes during the work shift and also to identify the possible causes of failures before the production cycle is complete. The findings of this research can be developed into a fully valid intellectual system of monitoring the completion of machine-building technological processes.

The developed models described in the article were integrated into the software complex for automation of high-tech equipment control in digital machinery production that is being developed in scope of applied research and experimental development commissioned by the Ministry of Education and Science of the Russian Federation.

ACKNOWLEDGMENTS

The work was carried out in scope of applied research and experimental development (PNIER) commissioned by the Ministry of Education and Science of the Russian Federation (PNIER unique ID is RFMEFI58016X0008).

REFERENCES

Romanov A., Romanov M., Kharchenko A., Kholopov V., 2016. Unified architecture of execution level hardware and software for discrete machinery manufacturing control systems. Proceedings – 14th IEEE Student Conference on Research and Development: Advancing Technology for Humanity, SCOReD 2016, p. 7810088. December 13-14, Kuala Lumpur (Malaysia). DOI: 10.1109/SCORED.2016.7810088

Kashirskaya E.N., Kurnasov E.V., Kholopov V.A., Shmeleva A.G., 2017. Methodology for assessing the implementation of the production process. Proceedings of 2017 IEEE 2nd International Conference on Control in Technical Systems, CTS 2017, pp. 232–235. October 25-27, St. Petersburg (Russia). DOI: 10.1109/CTSYS.2017.8109533

Kashirskaya E.N., Kholopov V.A., Shmeleva A.G., Kurnasov E.V., 2017. Simulation model for monitoring the execution of technological processes. Proceedings of 2017 IEEE 2nd International Conference on Control in Technical Systems, CTS 2017, pp. 307–310. October 25-27, St. Petersburg (Russia). DOI: 10.1109/CTSYS.2017.8109553

Desineni R., Berndlmaier Z., Winslow J., Blauberg A., Chu B.R., 2007. The Grand Pareto: A methodology for identifying and quantifying yield detractors in volume semiconductor manufacturing. IEEE Transactions on Semiconductor Manufacturing, 20(2), pp. 87-100. May 07. DOI: 10.1109/TSM.2007.896641

Kalach G.G., Romanov A.M., Tripolskiy P.E., 2016. Loosely coupled navigation system based on expert system using fuzzy logic. Proceedings of the 19th International Conference on Soft Computing and Measurements, SCM 2016, pp. 167-169. May 25-27, St. Petersburg (Russia). DOI: 10.1109/SCM.2016.7519717

Dmitriev S.A., Khalyasmaa A.I., Doroshenko V.O., Romanov A.M., 2016. Fuzzy dynamic model of power equipment state assessment. Proceedings of the 2016 IEEE North West Russia Section Young Researchers in Electrical and Electronic Engineering Conference, EIConRusNW 2016, pp. 547-552. February 2-3, St. Petersburg (Russia). DOI: 10.1109/EIConRusNW.2016.7448242

Huang P.B., Shiang W.-J., Jou Y.-T., Chang C., Ma C., 2010. An in-process adaptive control of surface roughness in end milling operations. 2010 International Conference on Machine Learning and Cybernetics, ICMLC 2010, vol. 3, pp. 1191-1194. July 11-14, Qingdao (China). DOI: 10.1109/ICMLC.2010.5580906

Jin Y., Sendhoff B., 2008. Pareto-based multiobjective machine learning: An overview and case studies. IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, 38(3), pp. 397-415. DOI: 10.1109/TSMCC.2008.919172

Chizhikov V.I., Kurnasov E.V., Vorob'ev E.I., 2018. Capture of an object on the basis of tactile surface recognition. Russian Engineering Research, 38(4), pp. 251–255. DOI: 10.3103/S1068798X18040044

Chao L., Peichen H., Jianping W., 2014. A study of semiconductor industry accidents: Making predictions based on BP artificial neural networks. IEEE International Conference on Industrial Engineering and Engineering Management, pp. 492-496. December 10-13, Bangkok (Thailand). DOI: 10.1109/IEEM.2013.6962460

Eroshenko S.A., Romanov A.M., 2016. Decision support system for science-technical solutions efficiency assessment based on hybrid neural networks. Proceedings of the 19th International Conference on Soft Computing and Measurements, SCM 2016, pp. 462-465. May 25-27, St. Petersburg (Russia). DOI: 10.1109/SCM.2016.7519814

LINEAR REGRESSION AND AGENT-BASED MODELING APPROACH FOR EQUIPMENT MARKET VALUE PREDICTION

Lingzi Wu [(a)], Mostafa Ali [(b)], Estacio Pereira [(c)], Simaan M. AbouRizk [(d)]

[(a),(b),(c),(d)]Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada T6G 2W2

[(a)]lingzi1@ualberta.ca, [(b)]mostafaali@ualberta.ca, [(c)]estacio@ualberta.ca, [(d)]abourizk@ualberta.ca

ABSTRACT

The acquisition, ownership, and operation costs of construction equipment represent a considerable portion of a general contractor's budget. The ability to accurately estimate the current market value of equipment is vital for successful fleet management. Although market valuation models have been built, they do not consider the impact of human behavior on selling price. Understanding basic auction mechanisms and the relationships between human behavior and selling price would improve the credibility and accuracy of equipment market value estimation—particularly for construction equipment, which is often sold at auction. This research proposes a hybrid linear regression and simulation modeling approach to uncover patterns in historical equipment resale data and to use these data, in consideration of human behavior, to simulate a typical North American equipment auction. Feasibility of the proposed approach is demonstrated through a case study.

Keywords: agent-based modeling, linear regression, equipment market value, decision support system

1. INTRODUCTION

Industrial construction projects depend heavily on the use of equipment, which can vary in price from tens to hundreds of thousands of dollars. In addition to acquisition costs, ownership and operating costs, which can range from 40% to 90% of the depreciation value depending on equipment type, must also be covered over the course of the equipment's life (Peurifoy, Schexnayder, and Shapira 2006). Although considerable, 35% of the American army's equipment maintenance cost is associated with the oldest 10% of their equipment (Peurifoy, Schexnayder, and Shapira 2006), indicating that replacement of existing, inferior equipment with newer models could considerably reduce age-associated maintenance costs. For this strategy to be effective, however, companies must ensure that they are able to accurately estimate and predict market prices of equipment.

Approximately half of certain construction equipment transactions (60%; Section 4: Illustrative Example) are completed through auction. Typical North American auctions follow an English-style auction, where auctioneers open the auction with a suggested opening bid (sometimes a reserve bid). This is followed by an iterative process, where bidders make incremental bids. When the bid is no longer raised, the last bidder wins the purchase of the item at the final bid price (Cliff 2003, Lucking-Reiley 1999, Vickrey 1961). Auction resale offers several advantages, including a quick cash return for sellers and potentially lower-than-market selling prices for bidders. These advantages are dependent, however, on the behavioral characteristics of the bidders. For example, aggressive bidders may result in an over-escalation of equipment sale prices.

Although a variety of market value estimation methods have been developed and reported in literature (Lucko 2011, Zong 2017), these methods are unable to consider the influence of bidder behavior on auction sale prices. Understanding basic auction mechanisms and the relationship between human behavior and selling price would improve the overall reliability and accuracy of equipment resale prices. Methods designed to incorporate human behavior into the valuation of construction equipment, however, have yet to be developed.

An approach capable of enhancing market value estimation of used construction equipment by incorporating behavioral bidder attributes is proposed. Linear regression is used to discover potential patterns in the construction equipment resale market, and an agent-based model is used to implement the results of this stage to simulate a typical North American equipment auction. The approach was applied to a case study, where multiple simulation experiments were performed to examine the relationship between various human behaviors and selling prices.

2. LITERATURE REVIEW

2.1. Equipment Valuation Methods

Various equipment valuation methods have been reported in literature. Zong (2017) explored equipment residual values using data mining techniques. Models were generated using the K nearest neighbor and random forest algorithms based on historical data to predict residual values. Lucko (2003) proposed a linear regression model to predict the residual value of various types of heavy equipment. Later, Lucko (2011) included economic indices in the statistical model of residual market values to reflect the impact of the economy on equipment selling prices at auction. In spite of these contributions, neither of the previous research studies examined or accounted for the influence of human behavioral factors during equipment auction.

2.2. Agent-Based Modeling

Agent-based modeling offers a conceptual framework to simulate various components of the auction process. Within an auction simulation model, resources are allocated between multiple agents based on a set of rules (Vidal 2010). Agent-based modeling has been used to successfully simulate a variety auction processes and market behaviors. Zhang, Bi, and Shen (2017) employed agent-based modeling to simulate the continuous double action market, where the price-volume relationship was examined with different investor structures. Huang, Liu, and Shi (2016) constructed an agent-based model to analyze and optimize an auction strategy for the grain market in China. Farnia, Frayret, Lebel, and Beaudry (2017) used the agent-based technique to simulate a multiple-round Timber auction, which was a sealed first-price auction. Anatolitis and Welisch (2017) employed an agent-based modeling approach to examine the performance of the two most commonly applied auction pricing rules (pay-as-bid and uniform pricing) in Germany with a certain degree of learning algorithm.

3. RESEARCH METHODOLOGY

In this research study, linear regression was used to identify patterns and derive inputs (i.e., equipment attributes and expected market values of equipment) for use by the agent-based model. An overview of the research methodology is illustrated in Figure 1.

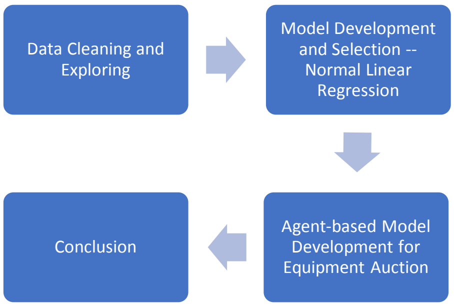


Figure 1: Research Methodology Diagram

3.1. Data Collection and Preparation

Historical sales data of previously-owned Chevrolet Silverado K2500 HD trucks were used to demonstrate the application the proposed approach. These trucks were chosen because (1) ¾ ton trucks are a common component of construction equipment fleets regardless of project scope and size and (2) ample sales data and (3) transaction conditions are available.

Data were collected from EquipmentWatch (https://equipmentwatch.com/) and consisted of information summarized in Table 1. Basic data cleaning methods were employed to ensure the data were valid and consistent.

The statistical computing software R (R Core Team 2017) was used to discover patterns between various covariates and market price (i.e., "Price_USD"). The relationship between market price and mileage is illustrated in Figure 2. At this stage, one outlier, indicated by the red circle in Figure 2, was identified.

Table 1: Equipment Data Labels and Descriptions Before and After Cleaning

Original Label	Description	Cleaned Label
Transaction	Auction or resale	Transaction
Marketplace	Company that handled transaction	Marketplace
Year	Model year of truck	ModelYear
Manufacturer/ Model	Manufacturer and model of truck	N/A
SerialNumber	Unique truck ID	SerialNumber
Meter Reads	Mileage of truck	Mileage _KMs
Location	Sate (US) or province (CAN) where equipment was sold	Location
Date	Date of sale	Date
Price	Sale price in USD	Price_USD

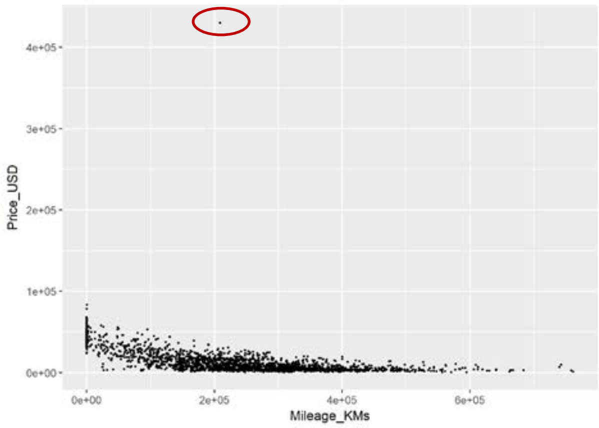


Figure 2: Relationship between market price ("Price_USD") and mileage "Mileage_KMs"

The final dataset (excluding the outlier) contained 3895 instances. Various types of plots, such as boxplots, point plots, and violin plots (a combination of a rotated kernel density plot and a box plot), were used to visualize the relationship as well as the statistical distribution of the covariates versus the market price for this dataset. Additional covariates were generated for further understanding of the data. Country (classified based on location), year of sale (year obtained from date of sale), age (difference between year of sale and model year), and condition (classified as defined in Table 2), were generated. Notably, nominal attributes were converted into numerical attributes for the artificial learning process using the command as.numeric in R. Relationships between these covariates and market price are plotted in Figure 3.

Table 2: Condition Groups and Corresponding Mileage Ranges

Condition	Mileage Range (km)
Low Mileage	< 10,000
Used, With Warranty	10,000 – 40,000
Used, Without Warranty	40,000 – 100,000
Old	100,000 – 150,000
Scrap	> 150,000

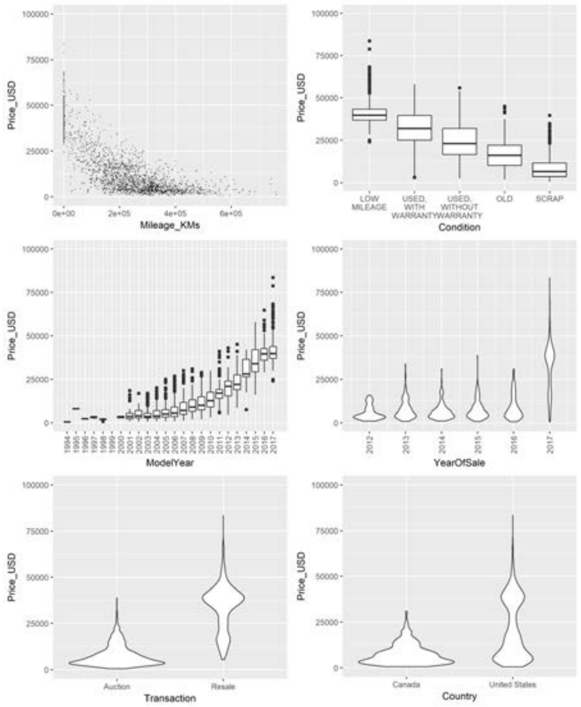


Figure 3: Relationships between market price ("Price_USD") and Covariates, Excluding Outlier

3.2. Linear Regression

Due to its ease of use and simplicity, linear regression was chosen as the means of predicting expected market values for input into the simulation model. The entire dataset was randomly divided: 75% of the 3895 instances were assigned as the training dataset, and 25% were assigned as the testing dataset. Based on the training set, nine normal linear regression models were developed using lm command in R.

Fitness of the linear regression model was examined using error- and criterion-based indicators. Specifically, sum square error, mean square error, root mean square error, log likelihood, adjusted R^2, Akaike Information Criterion (AIC), and the Bayes Information Criterion (BIC) were used in this research to select the best fitting linear regression model. Formulae for AIC and BIC are given as Equations 1 and 2, respectively.

$$AIC = -2(\text{log-likelihood}) + 2(\text{parameter number}) \quad (1)$$

$$BIC = -2(\text{log-likelihood}) + 2[\text{number of parameters} \times \log(\text{number of observations})] \quad (2)$$

Parameter values of the errors, R^2, log likelihood, AIC, and BIC for the nine models are summarized in Appendix A. Model 2 was associated with the lowest error rate, highest log-likelihood, and smallest AIC and BIC.

3.3. Agent-Based Model

An agent-based model was constructed to simulate a multiple-round sequential English auction. The proposed agent-based model, developed using Repast Simphony (North, Collier, Ozik, Tatara, Macal, Bragen, and Sydelko 2013)—a Java-based modeling system, is comprised of three elements: two agents (BIDDERS and EQUIPMENT) and one environment (AUCTION) in which the two agents interact and behave.

3.3.1. Element 1: BIDDERS

Attributes and behaviors of the agent BIDDERS are defined in Table 3. Based on these, two primary behaviors are defined, namely *calculate* and *update.* Initial values for BIDDERS are imported into the simulation through a comma-separated values file (.csv), allowing attributes to be easily modified for each simulation experiment.

Table 3: Attributes of Agent BIDDERS

Attributes	Description
Name	Unique ID of BIDDER
Size	Scale of BIDDER company
MyBudget	Initial equipment budget of BIDDER
MyNeed	Initial equipment number required by BIDDER
MyMaxRate	Aggressiveness of BIDDER: parameter in "MaxBid" calculation
MyIncRate	Aggressiveness of BIDDER: parameter in "MyBid" calculation
MyIncrement	Function of "MyIncRate" and EQUIPMENT "Increment"
MaxBid	"MyBudget"/"MyNeed" * "MyMaxRate"
Informed	Whether bidder is informed of the expected market value (EQUIPMENT "AppraisalValue")
MaxInformed Value	Threshold value for *calculate* decision

The *calculate* behavior represents every simulation click, where each BIDDER conducts a calculation and a logic check to determine if and how much to bid. The behavioral flowchart for the *calculate* behavior is illustrated in Figure 4.

The behavioral characteristic of aggressiveness was defined intuitively: the more aggressive the BIDDER, the greater the maximum bid the BIDDER is willing to offer and the larger the increment the BIDDER will adopt each round. Aggressiveness, therefore, is represented by the attributes "MyMaxRate" and "MyIncRate".

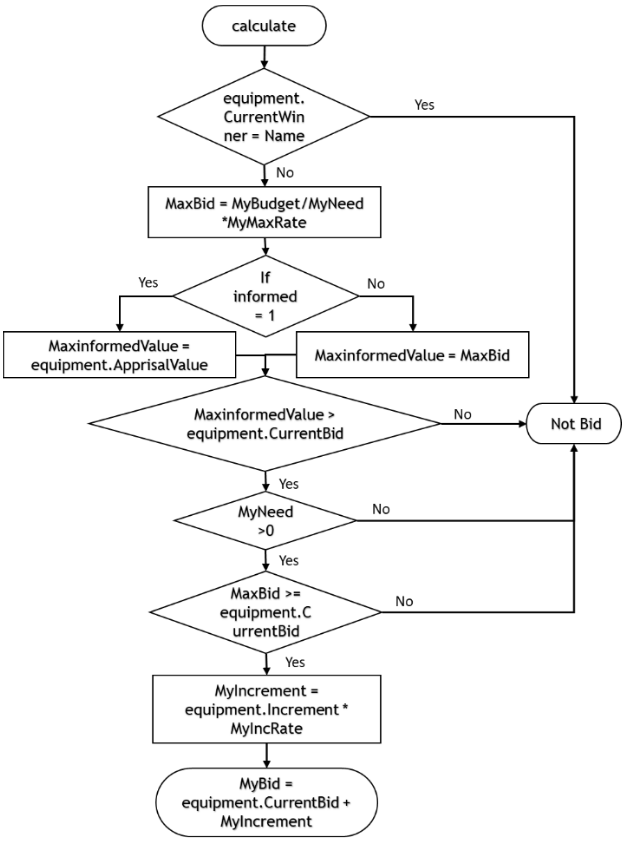


Figure 4: Flowchart for *Calculate* Behavior of BIDDERS

The *update* behavior represents the risk associated with the individual BIDDER who won the auction (i.e., highest bid for two consecutive rounds). It is included as a means of incorporating uncertainty and constraints into future simulations and is only called when equipment is sold.

When the *update* behavior is called, a random Boolean value, named "prof," is generated. A true "prof" value indicates that the BIDDER who won the current EQUIPMENT will make a profit operating this piece of equipment, and a certain percentage of the generated profit is then reallocated to this BIDDER's equipment budget. The increase is determined by drawing a machine-generated random number between 1 and 100 that represents the percentage of the EQUIPMENT's selling price by which the budget will increase. If the random number is greater than 80, the BIDDER not only increases the equipment budget but also the equipment need by 1 to represent business expansion.

In contrast, a false "prof" value indicates that the BIDDER loses money operating the equipment. Similarly, equipment budget and need are updated as described. A behavioral flowchart for the *update* behavior is illustrated in Figure 5.

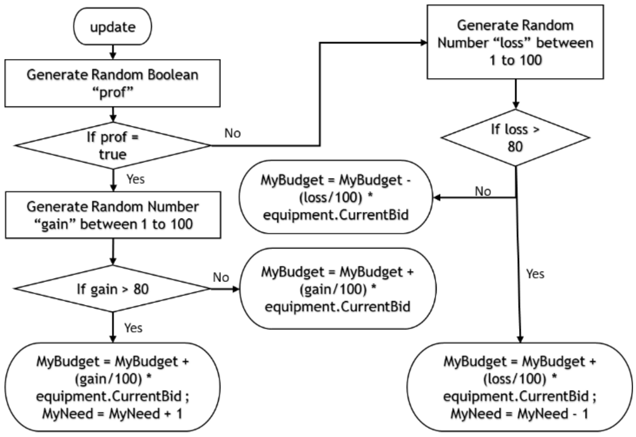


Figure 5: Flowchart for *Update* Behavior of BIDDERS

3.3.2. Element 2: EQUIPMENT

The agent EQUIPMENT represents both the equipment being auctioned and the auctioneer. Attributes and behaviors, indicated in Table 4, are defined to incorporate characteristics of both roles. Behaviors for EQUIPMENT are *decide*, *generate*, and *result*. The initial value for EQUIPMENT is generated using a random seed.

Table 4: Attributes and Description of EQUIPMENT

Attributes	Description
EquipmentNo	Unique ID of EQUIPMENT
MileageKM	Covariate of the linear regression model
Location	Covariate of the linear regression model
ModelYear	Covariate of the linear regression model
Condition	Covariate of the linear regression model
Country	Covariate of the linear regression model
YearofSale	Covariate of the linear regression model
AppraisalValue	Response to linear regression model representing expected market value
Increment	Minimum amount by which a new bid must exceed the previous bid
MinimumBid	Opening bid
CurrentBid	Captures bid amount for current round
CurrentWinner	Captures BIDDER "Name" whose bid was selected as "CurrentBid"

The *decide* behavior is a scheduled method that is initiated at simulation step 1 with an interval of 1. It is responsible for determining which BIDDER has the highest bid in each round and whether the EQUIPMENT is sold or not. To achieve this, bids are collected from each BIDDER. Then, bids are evaluated, and the maximum bid and bidder name are posted as the "CurrentBid" and the "CurrentWinner," respectively. If two or more bidders have the same highest bid, the tie is

resolved using random selection. BIDDERS access this information in the following simulation time click.
The *decide* behavior then compares the “CurrentBid” of the new time step with the “CurrentBid” of the previous time step. If the “CurrentBid” value has changed, the auction continues. If the “CurrentBid” value has not changed, the EQUIPMENT is sold. In the latter case, the two conditional behaviors, *result* and *generate,* are called. The *result* behavior logs the attribute and selling price of the equipment along with the bidder’s name in a .csv file. Using randomly generated attributes, the *generate* behavior then creates a new auction item.

3.3.3. Element 3: AUCTION

The AUCTION element is the environment where the BIDDER and EQUIPMENT agents are created prior to simulation and where they interact after the simulation is initiated.

4. ILLUSTRATIVE EXAMPLE

The objectives of this example were to (1) demonstrate the functionality of the proposed approach and (2) unveil potential associations between ultimate selling prices of Chevrolet Silverado K2500 HD truck and the percentage of informed BIDDERS. An input table of 45 bidders with nine different combinations of attribute values (i.e., each combination was assigned to five BIDDERS) was created. Note that, in this example, the budget was set so that available budgets of BIDDERS did not limit a BIDDER’s ability to bid. Attributes of the nine combinations are listed in Table 5.

Table 5: Attribute Combinations of BIDDERS

Name	MyMaxRate	MyIncRate	Informed	MyBudget	MyNeed	Size
AA	1	1	1	40000	5	Small
AF	1.5	1	1	40000	5	Small
AK	2	1	1	40000	5	Small
AP	1	1	1	120000	15	Medium
AY	1.5	1	1	120000	15	Medium
AZ	2	1	1	120000	15	Medium
BI	1	1	1	300000	30	Large
BJ	1.5	1	1	300000	30	Large
BT	2	1	1	300000	30	Large

Eleven simulation scenarios were constructed. The percentage of informed BIDDERS and the ratio of aggressive, non-informed BIDDERS were systematically varied between scenarios. Each scenario was run for 30,000 simulation clicks. Notably, because the auctions did not have a defined number of rounds, the total number of auction items (i.e., equipment) varied between simulation scenarios. Differences between the simulated selling price and the expected market price (hereafter referred to as “the differential”), are plotted in Figure 6 and summarized in Figure 7. Because BIDDERS’ budgets always exceeded expected market prices in this particular example, differentials were always positive (i.e., selling price > expected market price).

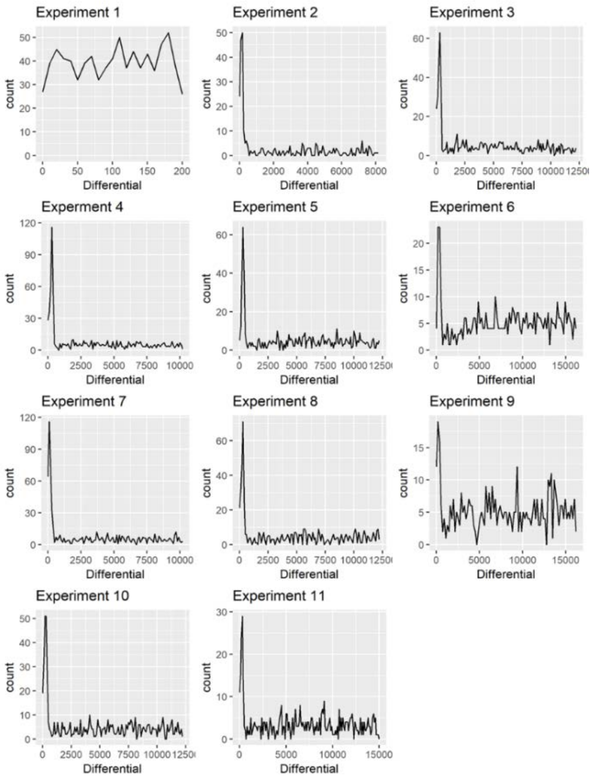


Figure 6: Frequency Polygon Plot for Each Scenario

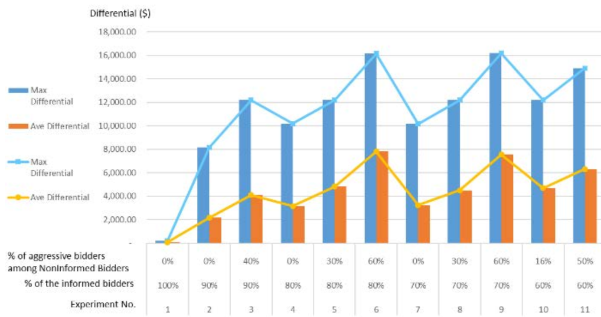


Figure 7: Maximum and Average Differentials for Each Scenario

As anticipated, when every bidder was informed of expected market values (Scenario 1), selling prices of equipment were similar to expected market values. Fluctuation of differentials between selling and expected market price was increased when the percentage of informed bidders was reduced (e.g., Scenario 1 vs. 2) or (2) the percentage of non-informed bidders that were aggressive was increased (e.g., Scenario 2 vs. 3), indicating that bidder knowledge and aggressiveness influence selling price.
Notably, the influence of informed bidders appeared to supersede the influence of aggressive bidders on selling price: differentials were increased despite the increase in non-informed aggressive bidders (e.g., Scenario 3 vs. 4). These results suggest that being informed of and

bidding based on historical market values, as opposed to bidding based on bids put forth by other bidders, is essential for avoiding overbidding. Additionally, if the bidder is familiar with other bidders' bidding strategies or habits, they may react accordingly to avoid being drawn into the irrational bidding game.

5. CONTRIBUTIONS AND FUTURE WORK

The aim of the present study was to introduce a research approach, which combined linear regression and simulation modeling methods, to more reliably estimate selling prices of equipment at auction by considering the effect of human behavior on auction results. Patterns in the resale market and resultant expectant market prices were determined using a linear regression approach. An agent-based model implemented the results of the data mining stage to simulate a typical North American multi-round equipment auction. Multiple simulation scenarios were evaluated to examine the relationship between various human bidding behaviors and selling price.

The presented case study demonstrated that overbidding (particularly in aggressive bidding situations) can be dampened by the presence of informed buyers. Practitioners, therefore, are suggested to be aware of expected market prices prior to equipment auctions.

While the proposed model has successfully demonstrated the ability to consider human behavior in auction-mediated equipment resale, generalizability and applicability of the model remain limited by several factors. First, the model developed here represents a multiple-round sequential English auction, where an auction item is only generated after the current item is sold. Thus, the bidders are unaware of equipment available in subsequent auction rounds. Having knowledge of upcoming bid items prior to auction can factor into the aggressiveness of a bidder. A future model could consider generating an equipment inventory at the beginning of the simulation and assigning various behaviors to bidders depending on their need for the equipment available in subsequent auction rounds. Second, the BIDDER information (i.e., input data) were arbitrarily generated for the purposes of demonstrating the functionality of the simulation model. The meaningfulness of the simulation model's outputs depends upon the specificity of the input data. Arbitrary data, such as those used in the present example, can be used to conduct various scenario analyses, while particular data corresponding to specific bidders can be used to obtain a more representative selling price. Third, the case study only considered two aspects of human auction behavior: knowledgeability and aggressiveness. Further research should focus on other potential human factors, as well as national and regional economic and market conditions, that can impact auction results.

Building upon the current approach by addressing the aforementioned points in future work is expected to result in an enhanced version of this model that is capable of predicting equipment market prices accurately and with a high level credibility. Achievement of such a model is expected to lead to the development of a practical decision support tool for improved equipment management in practice.

ACKNOWLEDGEMENTS

This project was supported by a Collaborative Research and Development Grant (CRDPJ 492657) from the Natural Sciences and Engineering Council of Canada.

APPENDIX A

Model #		1		2		3		4		5		6		7		8		9	
Model	Coefficients	Estimate	Pr(>|t|)	Estimate	Pr(>|t|)	Estimate	Pr(>|t|)	Estimate	Pr(>|t|)	Estimate	Pr(>|t|)	Estimate	Pr(>|t|)	Estimate	Pr(>|t|)	Estimate	Pr(>|t|)	Estimate	Pr(>|t|)
Covariate	intercept	-1953000	<2e-16	-2019000	<2e-16	-2041000	<2e-16	-3037000	<2e-16	-1996000	<2e-16	-2192000	<2e-16	-2297000	<2e-16	-3063000	<2e-16	-3232000	<2e-16
	Transaction	10280	<2e-16	10530	<2e-16	10570	<2e-16	9364	<2e-16	10420	<2e-16	10610	<2e-16	10840	<2e-16	9451	<2e-16	9615	<2e-16
	ModelYear	1524	<2e-16	1516	<2e-16	1512	<2e-16	1515	<2e-16	998	<2e-16	1097	<2e-16	1148	<2e-16	1530	<2e-16	1613	<2e-16
	Mileage_KMs	-0.02	<2e-16	-0.02	<2e-16	-0.02	<2e-16	-0.02	<2e-16	-0.02	<2e-16	0	<2e-16	-0.02	<2e-16	-0.02	<2e-16	-0.02	<2e-16
	Location	14.41	0.02	7.05	0.23	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	YearOfSale	-547.10	0.00	-506.90	0.00	-492.00	0.00	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Age	NA	NA	NA	NA	NA	NA	NA	NA	-557.60	0.00	-425.10	0.00	-448.60	0.00	NA	NA	NA	NA
	Marketplace	-59.03	0.42	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Condition	-711.40	0.00	-695.50	0.00	-667.60	0.00	-673.80	0.00	-743.90	0.00	-753.70	0.00	NA	NA	-844.40	0.00	NA	NA
	Country	1620.00	0.00	1721.00	0.00	1865.00	0.00	1404.00	0.00	1709.00	0.00	NA	NA	NA	NA	NA	NA	NA	NA
Result	Sum square error	82743635001		81785000000		85279348491		84133381046		85571036117		84666544987		87421132424		85759324521		87190392375	
	Mean square error	28336861		28008562		29205256		28812802		29305149		28995392		29938744		29369632		29859723	
	Root mean square error	5323.24		5292.31		5404.19		5367.76		5413.42		5384.74		5471.63		5419.38		5464.41	
	log likelihood	-29196.42		-29179.41		-29240.5		-29220.74		-29245.48		-29229.97		-29276.71		-29248.69		-29272.85	
	degree of freedom	2911		2912		2913		2914		2913		2914		2915		2915		2916	
	R squared (Adjusted)	0.889		0.889		0.886		0.887		0.886		0.885		0.881		0.886		0.883	
	AIC	58412.85		58376.82		58496.99		58455.49		58506.96		58473.93		58565.42		58509.38		58555.70	
	BIC	58472.64		58430.64		58544.82		58497.34		58554.80		58515.79		58601.30		58545.25		58585.60	

REFERENCES

Anatolitis V., and Welisch M., 2017. Putting renewable energy auctions into action – an agent-based model of onshore wind power auctions in Germany. Energy Policy, 110, 394-402.

Cliff D., 2003. Explorations in evolutionary design of online auction market mechanisms. Electronic Commerce Research and Applications, 2, 162-175.

Farnia F., Frayret, J., Lebel, L., & Beaudry, C. 2017. Agent-based simulation of multiple-round timber combinatorial auction. Canadian Journal of Forest Research, 47(1), 1-9.

Huang J., Liu L., Shi L. 2016. Auction policy analysis: an agent-based simulation optimization model of grain market. Proceeding of Winter Simulation Conference, pp. 3417-3428. December 11-14, Arlington (Virginia, USA).

Lucking-Reiley, D., 1999. Using field experiments to test equivalence between auction formats: Magic on the internet. The American Economic Review, 89 (5), 1063-1080.

Lucko G., 2003. A statistical analysis and model of the residual value of different types of heavy construction equipment. Thesis (PhD). Virginia Polytechnic Institute and State University.

Lucko G., 2011. Modeling the residual market value of construction equipment under changed economic conditions. Journal of Construction Engineering and Management, 137 10, 806-816.

North M. J., Collier N. T., Ozik, J., Tatara, E. R., Macal, C. M., Bragen, M., & Sydelko, P., 2013. Complex adaptive systems modeling with repast simphony. Complex Adaptive Systems Modeling, 1(1).

Peurifoy R. L., Schexnayder C. J., and Shapira, A., 2006. Construction planning, equipment, and methods. 7th ed. Boston, MA: McGraw-Hill Higher Education.

R Core Team. 2017. R: A language and environment for statistical computing. Available from: https://www.R-project.org [accessed Jan. 19, 2018].

Vickrey W., 1961. Counterspeculation, auctions, and competitive sealed tenders. The Journal of Finance, 16, 8-37

Vidal, J. M. 2006. Fundamentals of multiagent systems. Available from : http://jmvidal.cse.sc.edu/papers /mas.pdf [accessed April 14, 2018].

Zhang W., Bi Z., and Shen D., 2017. Investor structure and the price–volume relationship in a continuous double auction market: An agent-based modeling perspective. Physica A: Statistical Mechanics and its Applications, 467, 345-355.

Zong Y., 2017. Maintenance cost and residual value prediction of heavy construction equipment. Thesis (Master). University of Alberta.

STUDY ON FUEL ECONOMY AND REGENERATIVE BRAKING OF 48V MILD HYBRID VEHICLE WITH TWO MOTORS

Seongmin Ha[(a)], Wonbin Na[(b)], Hyeongcheol Lee[*(c)]

[(a) (b)]Department of Electric Engineering, Hanyang University, 222, Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea
[(c)] Division of Electrical and Biomedical Engineering, Hanyang University, 222, Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea

[(a)] haha4100@hanyang.ac.kr, [(b)] nao6114@hanyang.ac.kr, [(c)] hclee@hanyang.ac.kr

ABSTRACT

In this paper, we model 48V MHEV with P0 + P4 structure and develop HCU control algorithm. This study analyzes the characteristics of fuel consumption and regenerative braking according to the ratio of two motor powers. The HCU control algorithm consists of a mode decision algorithm and a power distribution algorithm. The vehicle model was developed using the Autonomy program. The HCU control algorithm was developed using Matlab/Simulink. Fuel economy simulation and regenerative braking energy analysis are performed in the FTP75 driving cycle environment. In this paper, the operating characteristics of 48V MHEV are compared according to the power ratio of two motors, and the regenerative braking energy is analyzed.

Keywords: Mild hybrid electric vehicle(MHEV), 48V System, Fuel economy, Regenerative braking

1. INTRODUCTION

Within the EU, road transport accounts for about 21% of total CO2 emissions and is the only major sector in which greenhouse gas emissions are still increasing. The EU is regulating the automotive industry to reduce CO2 emissions to 95g/km from 2021 with the development of innovative technologies for CO2 reduction and long-term investment commitments. Accordingly, we have developed a mild hybrid system that improves fuel efficiency by applying a power system consisting of a 48V battery, which is lower than the human body harmonic limit voltage of 60V, and a 12V battery, And it is responding to exhaust gas regulations to be strengthened in the future.

The 48V system is capable of reducing CO2 emissions and improving fuel efficiency through the start-stop function and regenerative braking of the engine. In addition, as shown in Figure 1, the dual power system that can supply power to high load devices such as air conditioner compressor, steering drive, pump and electric heater need (Kuypers, M. 2014).

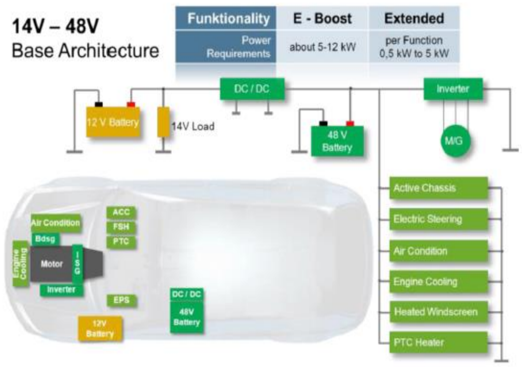


Figure 1: Multi-voltage architecture

The mild hybrid system structure based on the existing engine-based powertrain and BSG(Belt-driven Starter Generator) is under study to improve the fuel efficiency of the mild hybrid vehicle according to the BSG control because it can secure the price competitiveness while lowering the complexity of the entire system. The fuel economy improvement effect by the 48V mild hybrid system is generally as good as the structure capable of electric driving operation and mounting a high output electric motor. However, since the high-output electric motor becomes a factor of price and weight increase, an electric motor having an appropriate output according to the kind of vehicle must be selected.

Unlike high-power full hybrids, 48V-based mild hybrid vehicles are difficult to significantly improve engine efficiency due to operating point control, so it is appropriate to improve fuel economy through regenerative braking maximization and torque assist optimization.

As shown in Figure 2, there are five kinds of structures according to the motor mounting position of 48V MHEV. Research on vehicle characteristics and power distribution for a structure with a single motor has been actively pursued (Bao, R. 2017).

In the previous study, fuel economy was analysed for vehicles equipped with two motors with a maximum output of 11 kW BSG and a maximum output power of 10 kW Rear - Axle motor according to how to use regenerative braking energy. When only EV mode is used, it has been confirmed that optimum fuel economy can be achieved (S, HA. 2017).

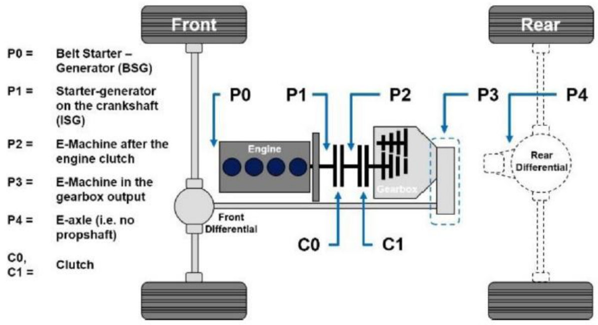


Figure 2: Configuration according to motor position

In this paper, we utilize the vehicle model(48V MHEV P0+P4 structure) and supervisory control algorithm developed in the previous research (S, HA. 2017). The sum of the maximum power of both motors equals 20kW, and analyzes the fuel consumption, vehicle operation and regenerative braking characteristics according to the power ratio of two motors.

2. 48V MHEV SYSTEM MODEL

In this paper, 48V MHEV with P0 + P4 structure is modeled as shown in Figure 3.

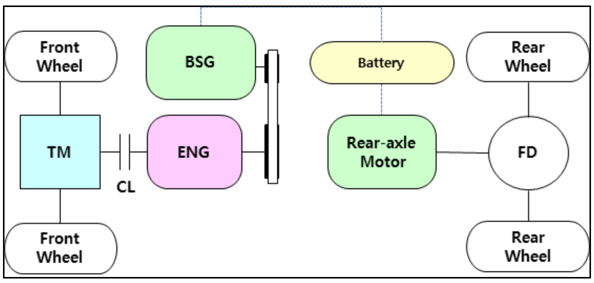


Figure 3: Target Vehicle

TM is the transmission, CL is the clutch, ENG is the engine, and FD is the final drive. The front wheel can be driven by the engine and the BSG, and the engine and the BSG are connected by a belt, which allows the torque assist and engine start via the BSG. The rear wheel is driven by a rear-axle motor.

2.1. Engine model

Table 1 shows the specification of the engine of the target vehicle.

Table 1: Engine specification

Engine parameter	
Type	Inline 4-cylinder
Fuel type	Gasoline
Maximum torque (Nm)	168.7
Maximum power (kW)	99.54

Figure 4 represents maximum torque and BSFC(Brake Specific Fuel Consumption) at engine torque and speed conditions. As shown in Equation (1), the engine model outputs the engine output torque determined by the host controller (HCU) to the torque value between the maximum torque in the throttle maximum open state and the engine friction torque in the closed state.

$$T_{eng} = \min(T_{wott}(\omega_{eng}), \max(T_{HCU,eng} + T_{eng,idle}, T_{ctt}(\omega_{eng}))) \quad (1)$$

Where,

$$T_{eng,idle} = \max(k_p(\omega_{idle} - \omega_{eng}), 0)$$

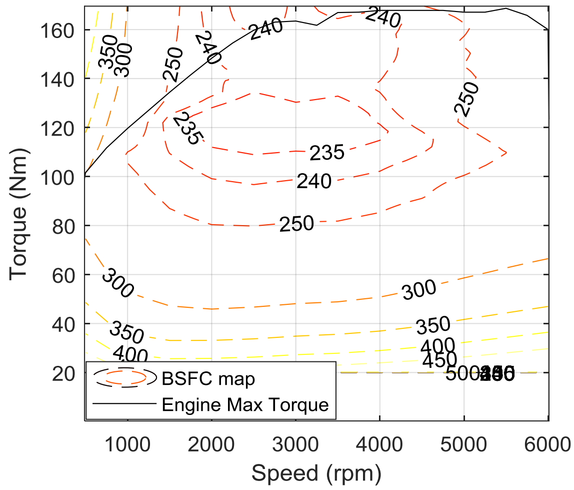


Figure 4: Engine maximum torque and BSFC map

2.2. Motor model

The motors are classified into two types according to their mounting positions. The BSG is connected to the engine by a belt. The rear-axle motor is connected to the rear reduction gear. According to the simulation case, each motor power is defined differently.

Table 2: Electric motor specification

Electric motor parameter		
Type	BSG	Rear-axle motor
Max power (kW)	15/10/5	15/10/5
Max torque (Nm)	95.4/63.6/31.8	47.7/31.8/15.9
Base Speed (RPM)	1500	3000
Max Speed (RPM)	16000	12000

The drive motors mounted on the front and rear wheels have the same model structure, and the corresponding motor model reflects the motor torque command value determined by the HCU in the output. The output torque of the motor model is limited to the maximum drive or braking torque possible at the current motor speed. The torque of the motor is expressed by Equation (2), and the power of the motor has a relationship as shown in Equation (3).

$$T_{mot} = \min(T_{\max}(\omega_{mot}), \max(T_{HCU.mot} - T_{\max}(\omega_{mot}))) \quad (2)$$

$$P_{mot} = T_{mot} \cdot \omega_{mot} \cdot \eta^{k}(T_{mot}, \omega_{mot}) \quad (3)$$

Where,

$$k = \begin{cases} 1 & (T_{mot} \cdot \omega_{mot} < 0) \\ -1 & (T_{mot} \cdot \omega_{mot} \geq 0) \end{cases}$$

2.3. LDC model

LDC (Low Voltage DC-DC Converter) supplies power to a 12V electric field load. The 12V total field load power is 217W and the LDC efficiency is 0.95, both of which are fixed values. The LDC output power is given by Equation (4).

$$P_{LDC} = \frac{P_{Load}}{\eta_{LDC}} \quad (4)$$

2.4. Battery model

The 48V battery cell voltage is calculated from the open circuit voltage of the battery cell, the cell internal resistance, and the current. The temperature is assumed to be maintained at 40 ° C, and the internal resistance is determined according to the direction of the current and the SOC as shown in equation (5).

$$V_{bat.cell} = V_{OCV}(SOC) - R_{\text{int}}(SOC, \text{sgn}(I_{bat}) \cdot I_{bat} \quad (5)$$

Where,

$$R_{\text{int}} = \begin{cases} R_{discharge} & (\text{sgn}(I_{bat}) \geq 0) \\ R_{charge} & (\text{sgn}(I_{bat}) < 0) \end{cases}$$

The current of the 48V battery is calculated by the equation (6) with the sum of BSG power, rear-axle motor power and LDC power divided by the battery voltage.

$$I_{bat} = \frac{P_{mot.bsg} + P_{mot.rear} + P_{LDC}}{N_{cell} \cdot V_{bat.cell}} \quad (6)$$

Nominal voltage is 3.7V, and capacity is 23Ah. 13 battery cells are connected in series.

3. SUPERVISORY CONTROL ALGORITHM

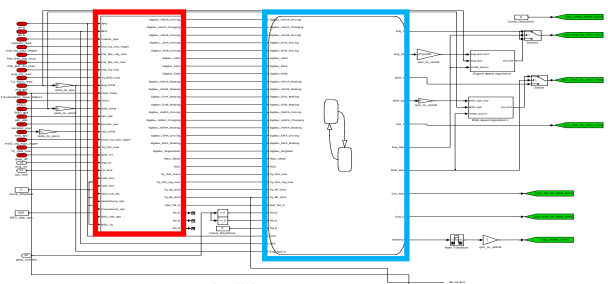


Figure 5:Supervisory control algorithm(Power distribution(red)/Mode decision(blue) algorithm)

As shown in the figure 5, the upper control algorithm was developed using Simulink, and consists of the mode decision algorithm and the power distribution algorithm of each mode. Details of driving mode and power distribution can be found in previous paper(S, Ha. 2017).

4. SIMULATION

4.1. Simulation case

The sum of each motor maximum power is 20kW, and the power ratio of each motor is configured differently for each simulation case. The simulation cases are shown in the table 3.

Table 3: Electric motors power by simulation case

Simulation case	Motor Power (Rear-Axle motor / BSG)	
A	15kW	5kW
B	10kW	10kW
C	5kW	15kW

4.2. Simulation result

In all cases, the speed of the vehicle follows the speed of the FTP-75 cycle without significant error, as shown in Figure 6.

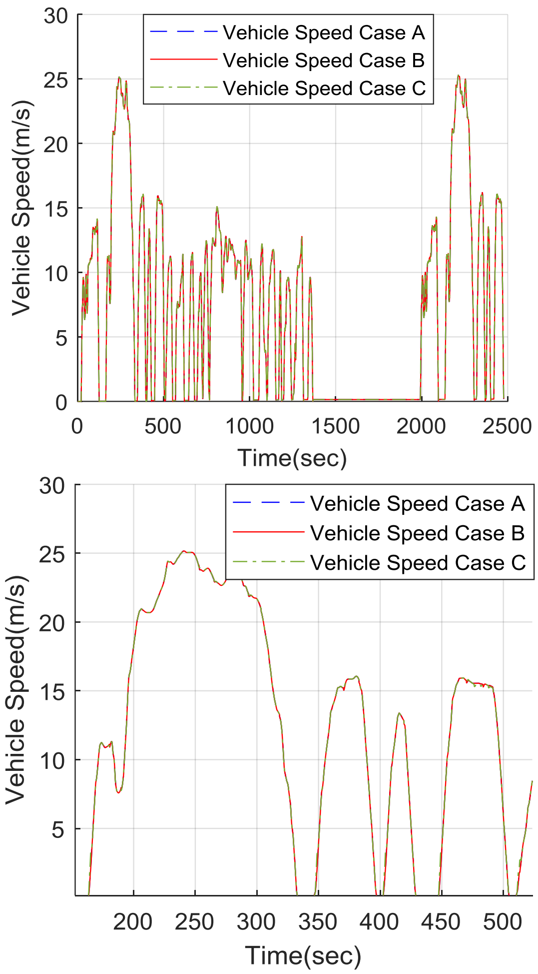


Figure 6:Vehicle Speed at simulation time(Top) / at 160sec to 520sec(Bottom)

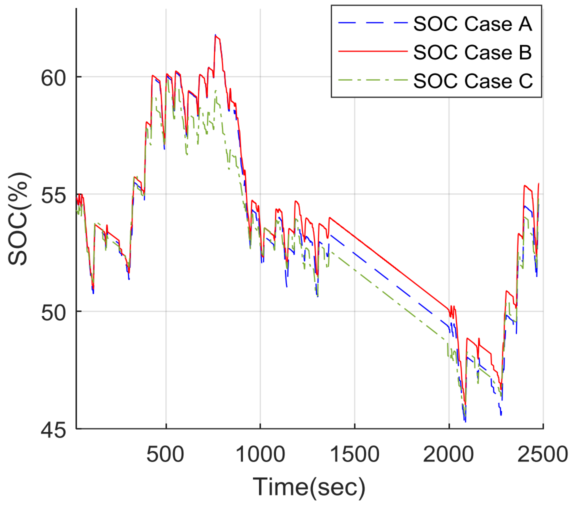


Figure 7: SOC of 48V battery at simulation time

As can be seen in figure 7, the final SOC of a 48V battery does not exceed 0.5% at the initial SOC 55%. Therefore, in order to compare the fuel economy of each case, the fuel economy effect due to the SOC difference is excluded.

Table 4: Fuel economy results

Simulation case	Motor Power (Rear-Axle motor / BSG)	Fuel economy
A	15kW/5kW	19.48 km/l
B	10kW/10kW	19.33 km/l
C	5kW/15kW	18.59 km/l

The fuel efficiency of each case is shown in Table 4. Based on Case B, the fuel economy difference of Case A and B is 0.15km/l (0.78%). The fuel economy difference of Case B and C is 0.74km/l (3.98%). The cause of the difference in fuel efficiency is confirmed through analysis of simulation results.

4.3. Analysis of simulation result

From the simulation results, the regenerative braking energy for each motor is calculated as shown in equation (7).

$$E_{reg} = \int_0^T V_{mot}(t) \times |I_{mot}(t)| \times k_{reg}(BPS)\,dt \quad (7)$$

Where,

$$k_{reg} = \begin{cases} 1 & (BPS \neq 0) \\ 0 & (BPS = 0) \end{cases}$$

Table 5: Regenerative braking energy (kJ)

Simulation case	Rear-Axle motor	BSG	Total
A	2066.44	246.46	2312.90
B	1567.34	675.57	2242.91
C	889.52	1244.48	2133.99

E_{reg} is the regenerative braking energy, V_{mot} is the motor input voltage, I_{mot} is the motor input current, T is simulation time and BPS is the brake pedal signal. The calculated regenerative braking energy is shown in Table 5.

Based on Case B, the total regenerative braking energy difference is 69.99kJ(3.12%) for A and B, and 108.92kJ (5.1%) for B and C. This is very small compared to the fuel economy difference. Even with the same motor total power, the rear-axle motor has better mechanical efficiency than BSG, resulting in a difference in regenerative braking energy. In other words, Case C uses regenerative braking energy more inefficiently than other cases.

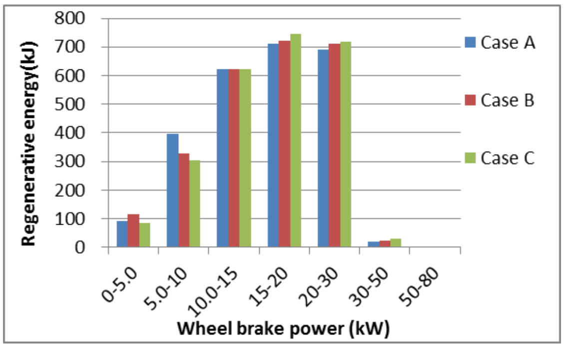


Figure 8: Regenerative braking energy according to braking power range of wheel

Figure 8 shows the regenerative braking energy charged for each braking power range on the wheel. The regenerative braking is mainly performed with a wheel braking power of 10 to 30 kW. The higher the BSG maximum power, the more regenerative braking energy is obtained in the higher power range. Although there is a difference according to the case, there is a similar tendency because the total power of the motor is the same.

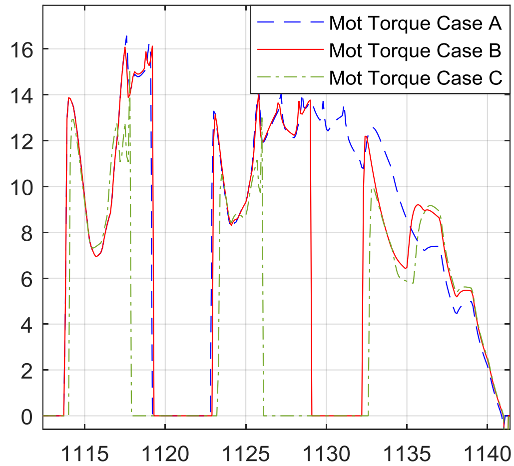


Figure 9: Rear-axle motor torque in the EV mode driving section (1112sec - 1142sec)

Figure 9 shows the torque of the rear-axle motor in the EV mode section. The time to drive the EV mode is the longest in Case A, and the shortest in Case C. This is because the maximum torque that can be output depends on the motor power. The larger the maximum torque, the wider the area that can be driven in the EV mode. To quantitatively confirm the energy used in the EV mode for each case, the electric energy used in the EV mode was calculated as in Equation (8).

$$E_{EV} = \int_0^T V_{mot}(t) \times |I_{mot}(t)| \times k_{EV}\, dt \quad (8)$$

Where,

$$k_{EV} = \begin{cases} 1 & (\text{EV driving mode}) \\ 0 & (\text{not EV driving mode}) \end{cases}$$

Table 6: Electric energy used in EV driving mode

Simulation case	Electric energy(kJ)
A	1600.76
B	1503.04
C	946.65

Table 6 shows the calculated electric energy during EV mode driving. The difference between the total regenerative braking energy of cases A and B and the energy difference used in EV mode are similar. This is because cases A and B use electric energy in EV mode driving, except for electric load and loss. In Case C, there is much more EV energy difference than regenerative braking energy difference with B. Since the maximum power is low, the energy obtained from regenerative braking can't be used as EV mode at all. So the remaining energy is consumed via torque assist in HEV mode.

In summary, it is a way to obtain the optimum fuel efficiency by consuming all the electric energy obtained from the regenerative braking in the EV mode that runs with the engine turned off.

5. CONCLUSION

In this paper, fuel efficiency and regenerative braking energy are analyzed when the power sum of two motors is the same and the ratio of power is different in 48V MHEV of P0 + P4 structure. The difference of regenerative braking energy showed a slight difference according to the ratio of motor power. This occurs because the rear-axle motor is more mechanically efficient than the BSG, which is belt-coupled to the engine. However, the effect of fuel economy is not significant. A more important factor is the use of energy from regenerative braking. Even if the same amount of electric energy is consumed, EV drive and HEV assist differ much in fuel efficiency. In the EV drive mode, energy can be used more efficiently by turning off the engine and releasing the clutch to exclude engine friction from the drive system.

The maximum power of the rear-axle motor must be selected to at least consume the regenerative braking energy in the EV mode. The maximum power of the BSG must be selected to at least perform engine starting. In 48V systems, increasing the power sum of both motors is not efficient because the battery output power is limited. The BSG has the minimum power for starting the engine and the higher the power ratio of the rear-axle motor, the better the fuel economy of the 48V MHEV system with two motors.

ACKNOWLEDGMENTS

This work was supported by the Industrial Strategic technology development program, 20000578, A development of optimal eco driving system in HEV/PHEV based on connected vehicle environment funded By the Ministry of Trade, industry & Energy(MI, Korea).

REFERENCES

S, Ha., T, Park., W, Na,. and H, Lee,."Power distribution control algorithm for fuel economy optimization of 48V mild hybrid vehicle," Proc. of the Int. Conference on Modeling and Applied Simulation, 2017, ISBN 978-88-97999-91-1, 185-190.

A. Rick and B. Sisk, "A Simulation Based Analysis of 12V and 48V Microhybrid System Across Vehicle Segments and Drive Cycles," SAE Technical Paper, 2015-01-1151.

Bao, R., Avila, V., and Baxter, J., "Effect of 48V Mild Hybrid System Layout on Powertrain System Efficiency and Its Potential of Fuel Economy Improvement," SAE Technical Paper 2017-01-1175, 2017.

German, J. "Hybrid vehicles: Trends in technology development and cost reduction," ICCT, http://www.theicct.org/hybrid-vehicles-trends-technology-development-and-cost-reduction, 2015.

Awadallah, M., Tawadros, P., Walker, P., and Zhang, N., Comparative fuel economy, cost and emissions analysis of a novel mild hybrid and conventional vehicles. Proc IMechE Part D: J Automobile Eng 2017; 1-17.

Kuypers, M., "Application of 48 Volt for Mild Hybrid Vehicles and High Power Loads," SAE Technical Paper2014-01-1790, 2014, doi:10.4271/2014-01-1790.

MODEL-BASED FAULT DETECTION AND ISOLATION ALGORITHM FOR VEHICLE SEMI-ACTIVE SUSPENSION SYSTEM

Wonbin Na[(a)], Juin Lee[(b)], Hyeongcheol Lee[(c)]

[(a),(b)] Department of Electric Engineering, Hanyang University, 222, Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea
[(c)] Division of Electrical and Biomedical Engineering, Hanyang University, 222, Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea

[(a)]nao6114@hanyang.ac.kr, [(b)] galaxyeagle@hanyang.ac.kr, [(c)] hclee@hanyang.ac.kr

ABSTRACT

Recently, mechatronics systems are increasingly being applied to vehicles. Among them, ECS (Electronically Controlled Suspension) system is applied to improve the stability of the vehicle and the comfort of the driver. There are Three kinds of ECS; semi-active and full active systems with hydraulic or pneumatic actuators and electronic controls added and support to the existing passive suspension consist of spring and damper. Especially, the semi-active system is less costly than the full-active system, and because of the difference in control method, energy usage is relatively low, so it is widely applied to advanced vehicles. The application of this system is increasing the importance of fault diagnosis of these systems in order to ensure the stability of the intended vehicle and the comfort of the driver. In this paper, we propose a model-based fault detection and isolation algorithm based on parity equation and incident matrix method.

Keywords: model-based fault diagnosis, electronically controlled suspension(ECS), semi-active suspension

1. INTRODUCTION

The suspension system of the vehicle is mounted between the axle and the vehicle body to support the weight of the vehicle and to mitigate vibration caused by obstacles on the road. These suspension systems play a decisive role in improving the driving performance and ride comfort (Kim and Lee, 2011). For example, suspension determines wheel alignment (toe, camber, and caster) to create maximum traction force and to reduce the roll of the vehicle. The suspension has a trade-off relationship between driving performance and rides comfort according to spring coefficient and damping coefficient. The characteristics of suspension are determined according to the vehicle type. In the past, suspension characteristics were fixed once the vehicle was designed. Recently, the Electronically Controlled Suspension (ECS) system is developed using the Electronic Control Unit (ECU). The ECS system can control automatically the characteristic of the suspension according to the road condition and driving situation.

The ECS system is largely divided into two types, Active type ECS (Active ECS) and Semi-Active type ECS (Semi-Active ECS). The Active ECS controls the spring coefficient using air spring suspension and the damping coefficient using a solenoid valve at the damper. This system can maintain a constant height of vehicle and control vehicle attitude like roll and pitch (Börner, Strakey, Weispfenning and Isermann, 2002). But ECS system has an expensive price, complex structure, and high energy usage because of the solenoid valve and the air spring. The Semi-Active ECS control only damping coefficient using solenoid valve based on Sky Hook control method. Semi-Active ECS system applies to many vehicles because this is cheap and have the simple structure. So, this paper focus on Semi-Active ECS system.

As mentioned above, the ECS system is an important factor in determining ride comfort and driving performance. The fault detection and isolation of the ECS system are essential for driver's safety and convenience.

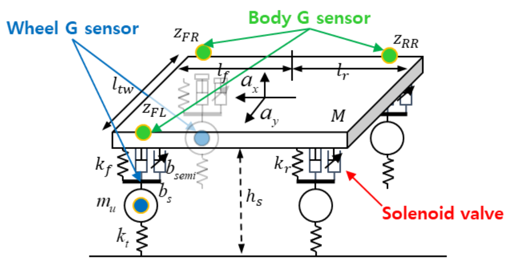


Figure 1: ECS System Scheme

ECS system to be considered in this paper constitutes fig.1. The ECS, which improves driver ride comfort and improves driving stability, has three G sensors and two G sensors on the vehicle body for control. In this paper, we constitute the fault detection and isolation algorithm model based parity equation.

Proc. of the Int. Conference on Modeling and Applied Simulation 2018,
ISBN 978-88-85741-07-2; Bruzzone, De Felice, Frydman, Longo, Massei and Solis, Eds.

2. MODEL-BASED FAULT DETECTION AND ISOLATION

2.1. Model-based residual calculation

There are output error method and polynomial method to calculate residual for fault diagnosis. In this study, a polynomial error type residual was intended to be constructed to diagnose faults. Therefore, the two estimation formulas needed (Isermann, 1995). As the first calculation, estimates (1) ~ (4) were made using longitudinal acceleration and lateral acceleration using full car model of Fig 1 (Na, Lee and Lee, 2017). M is vehicle mass, g is gravitational acceleration, l is length of vehicle, l_f is vehicle length of front from vehicle mass center, l_r is vehicle length of rear from vehicle mass center, h_s is height of vehicle mass center from ground, a_x is longitudinal acceleration of vehicle, a_y is lateral acceleration of vehicle and l_{tw} is track width of vehicle.

$$F_{z,FL,1} = \frac{Mgl_r}{2l} - \frac{m_s a_x h_s}{2l} - \frac{k_f m_s a_y h_s}{l_{tw}} \quad (1)$$

$$F_{z,FR,1} = \frac{Mgl_r}{2l} - \frac{m_s a_x h_s}{2l} + \frac{k_f m_s a_y h_s}{l_{tw}} \quad (2)$$

$$F_{z,RL,1} = \frac{Mgl_f}{2l} + \frac{m_s a_x h_s}{2l} - \frac{k_r m_s a_y h_s}{l_{tw}} \quad (3)$$

$$F_{z,RL,1} = \frac{Mgl_f}{2l} + \frac{m_s a_x h_s}{2l} + \frac{k_r m_s a_y h_s}{l_{tw}} \quad (4)$$

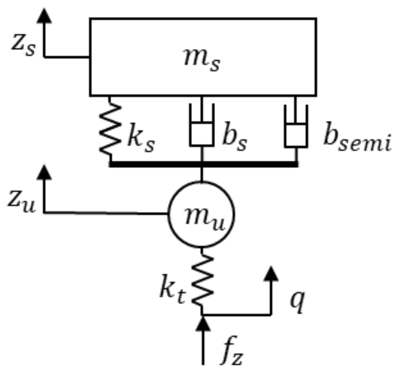


Figure 2: Quarter Car Model

In the second estimation, normal forces were calculated by considering each wheel of the vehicle as a quarter-car model as (5) ~ (7). (Rajamani, 2012)

$$m_s\ddot{z}_s + k_s(z_s - z_u) + b_s(\dot{z}_s - \dot{z}_u) = -b_{semi}(\dot{z}_s - \dot{z}_u) \quad (5)$$

$$m_u\ddot{z}_u + k_s(z_u - z_s) + b_s(\dot{z}_u - \dot{z}_s) + k_t(z_u - q) = b_{semi}(\dot{z}_s - \dot{z}_u) \quad (6)$$

$$f_z + k_t(q - z_u) \quad (7)$$

m_s is sprung mass of quarter car model, m_u is unsprung mass of quarter car model, k_s is spring coefficient of semi-active suspension, k_t is spring coefficient of tire, b_s is damping coefficient of semi-active suspension, f_z is normal force of each wheel. Therefore, normal force can be calculated by (8).

$$F_{z,i,2} = (m_s + m_u)g - m_s\ddot{z}_s - m_u\ddot{z}_u, \quad (i = FL, FR, RL, RR) \quad (8)$$

Here, when considering the configuration of sensors in the actual system, the output of non-existent sensors should be estimated as (9) ~ (16) by the existing sensors. v_x is longitudinal velocity of vehicle.

$$\dot{z}_{s,fl}(t) = \int_{t_0}^{t} \ddot{z}_{s,fl}(t)\,dt \quad (9)$$

$$\dot{z}_{s,fr}(t) = \int_{t_0}^{t} \ddot{z}_{s,fr}(t)\,dt \quad (10)$$

$$\dot{z}_{s,rl}(t) = \int_{t_0}^{t} \ddot{z}_{s,fl}(t)\,dt - \int_{t_0}^{t} \ddot{z}_{s,fr}(t)\,dt + \int_{t_0}^{t} \ddot{z}_{s,rr}(t)\,dt \quad (11)$$

$$\dot{z}_{s,rr}(t) = \int_{t_0}^{t} \ddot{z}_{s,rr}(t)\,dt \quad (12)$$

$$\dot{z}_{u,fl}(t) = \int_{t_0}^{t} \ddot{z}_{u,fl}(t)\,dt \quad (13)$$

$$\dot{z}_{u,fr}(t) = \int_{t_0}^{t} \ddot{z}_{u,fr}(t)\,dt \quad (14)$$

$$\dot{z}_{u,rl}(t) = \dot{z}_{u,fl}(t + (l_f + l_r)/v_x) \quad (15)$$

$$\dot{z}_{u,rr}(t) = \dot{z}_{u,fr}(t + (l_f + l_r)/v_x) \quad (16)$$

Thus, the result of subtracting the two estimates are (17) ~ (20) in the form of a polynomic error.

$$r_1(s) = F_{z,FL,1}(a_x, a_y) - F_{z,FL,2}(\ddot{z}_{s,FL}, \ddot{z}_{u,FL}) \quad (17)$$

$$r_2(s) = F_{z,FR,1}(a_x, a_y) - F_{z,FR,2}(\ddot{z}_{s,FR}, \ddot{z}_{u,FR}) \quad (18)$$

$$r_3(s) = F_{z,RL,1}(a_x, a_y) - F_{z,RL,2}(\ddot{z}_{s,FL}, \ddot{z}_{s,FR}, \ddot{z}_{s,RR}, \ddot{z}_{u,FL}) \quad (19)$$

$$r_4(s) = F_{z,RR,1}(a_x, a_y) - F_{z,RR,2}(\ddot{z}_{s,RR}, \ddot{z}_{u,FR}) \quad (20)$$

Additionally, the roll angle of the vehicle was estimated and added to the FDI algorithm. (21) is equation for estimate angle of roll of vehicles. The suspension deflection used in (21) is calculated as (22). Since this study assumes driving on flat roads, the suspension deflation is assumed to be (23). The second estimation equation for roll angle is (24). We can calculate residual of (25) using (21) and (24). (Ryu and Gerdes, 2004)

$$\theta_1 = \frac{\Delta z_{fl} - \Delta z_{fr} + \Delta z_{rl} - \Delta z_{rr}}{2l_{tw}} \quad (21)$$

$$\Delta z_i = z_{s,i} - z_{u,i}, (i = fl, fr, rl, rr) \quad (22)$$

$$\Delta z_i \approx z_{s,i}, \ (i = fl, fr, rl, rr) \quad (23)$$

$$\theta_2 = -\left(\frac{Mh_s}{k_{roll}}\right) a_y \quad (24)$$

$$r_5: \theta_1 - \theta_2 = c_{10}(\, a_y, \ddot{z_{s,fl}}, \ddot{z_{s,fr}}) \quad (25)$$

θ is vehicle roll angle and k_{roll} is coefficient of vehicle roll.

2.2. Residual analysis and simulation

We analyzed the relationship between faults and residuals using the Incident matrix method (Blanke, Kinnaert, Lunze, and Stroswiecki, 2006). This allows you to see which each fault affect residuals. Table 1 below shows the incident matrix for the residuals proposed and used in this paper. The result can be expected to affect the residuals 1, 3, and 5 when the body G sensor has a fault. Likewise, other fault cases can be predicted, and residuals results can be independent to separate each failure.

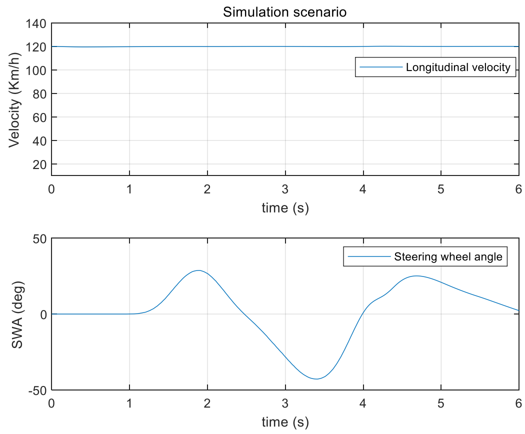


Figure 3: Simulation Scenario

Table 1: Incident Matrix

Residuals	Sensors						
	a_x	a_y	Body G sensor			Wheel G sensor	
			FL	FR	RR	FL	FR
r1	X	X	X			X	
r2	X	X		X			X
r3	X	X	X	X	X	X	
r4	X	X			X		X
r5		X	X	X			

Carsim, a vehicle dynamics simulator, was used to verify estimation. And proposed FDI algorithm is verified using the DLC (Double Lane Change) scenario in Fig 3. In case of normal situation, we confirmed that the actual values and the estimated signals used for residuals are very consistent as shown in Fig 4 ~ 8.

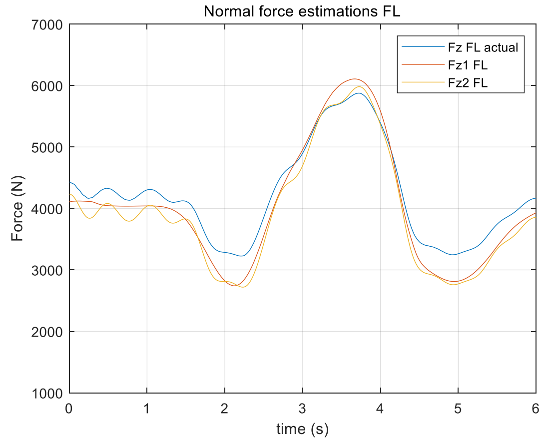


Figure 4: Normal Force FL Estimations

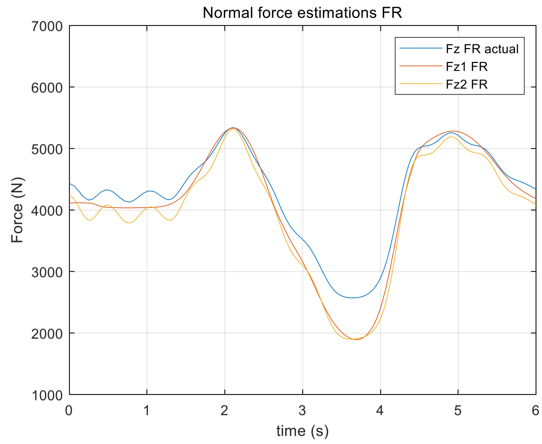


Figure 5: Normal Force FR Estimations

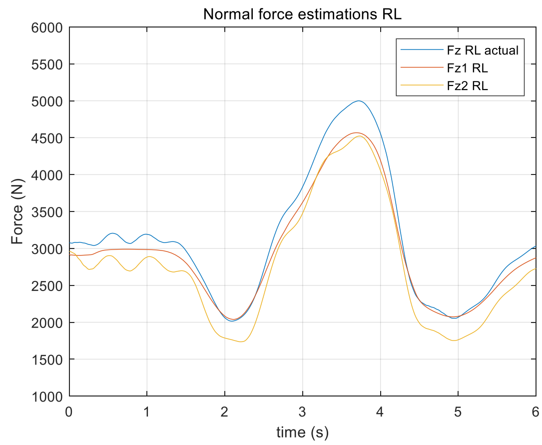


Figure 6: Normal Force RL Estimations

Proc. of the Int. Conference on Modeling and Applied Simulation 2018,
ISBN 978-88-85741-07-2; Bruzzone, De Felice, Frydman, Longo, Massei and Solis, Eds.

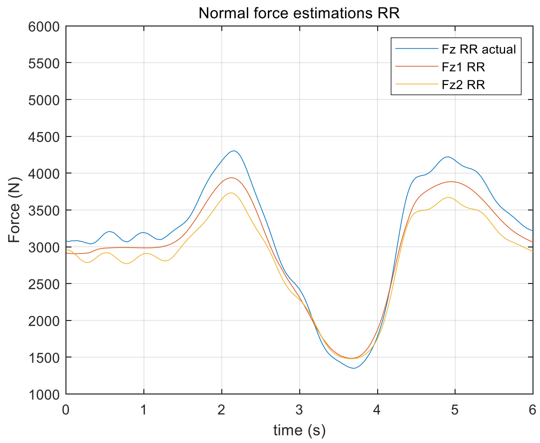


Figure 7: Normal Force RR Estimations

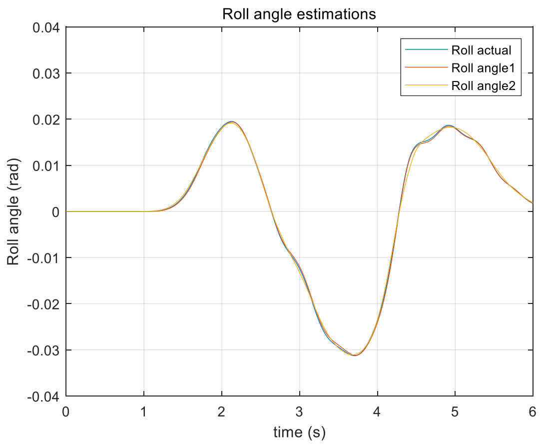


Figure 8: Roll Angle Estimations

2.3. FDI Simulation result

Finally, the proposed FDI algorithm was verified by injecting 0.1 g of fault at 3s. As shown in section 2.3, residuals are very low when there is no failure (Fig 10). Likewise, as analyzed in the Incident matrix of section 2.3, we can see that the residuals are large when the Body G sensor has fault. Therefore this residuals deviate normal boundary; Threshold. Fig 9 shows scheme of fault detection using threshold in polynomial error method residual calculation. For fault detection, we fixed threshold values properly (Blanke, Kinnaert, Lunze, and Stroswiecki, 2006).

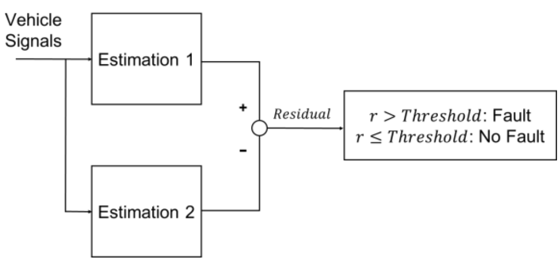


Figure 9: Fault Detection scheme

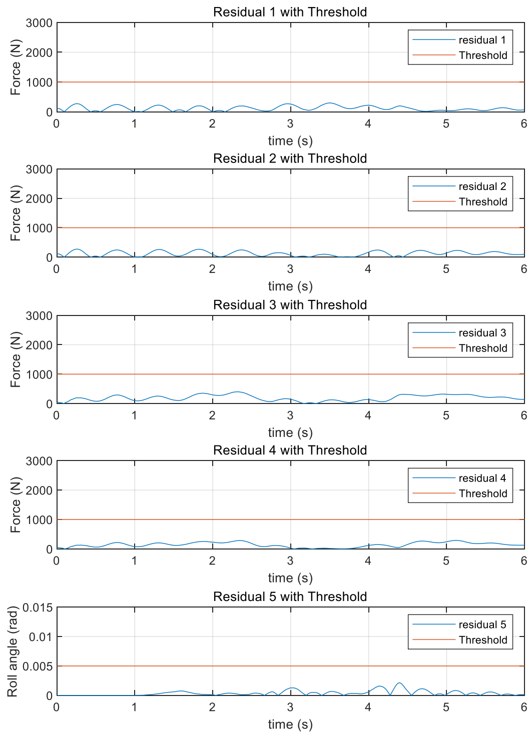


Figure 10: Residuals 1~ 5 with No Fault

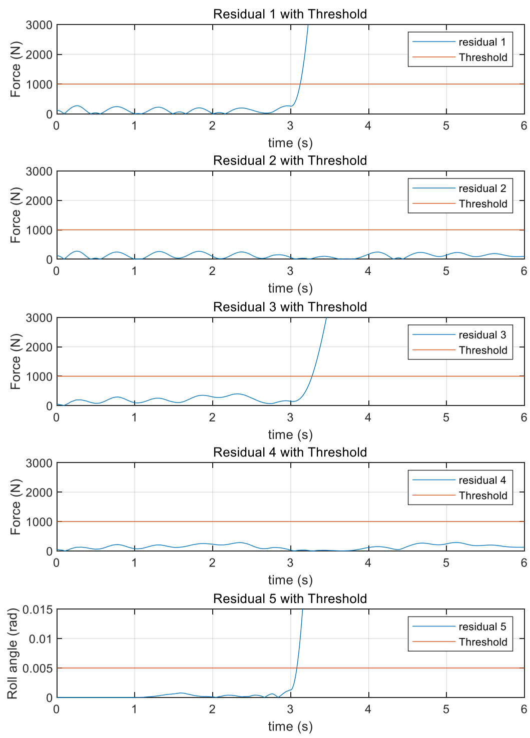


Figure 11: Residuals 1~5 with Body G sensor FL Fault

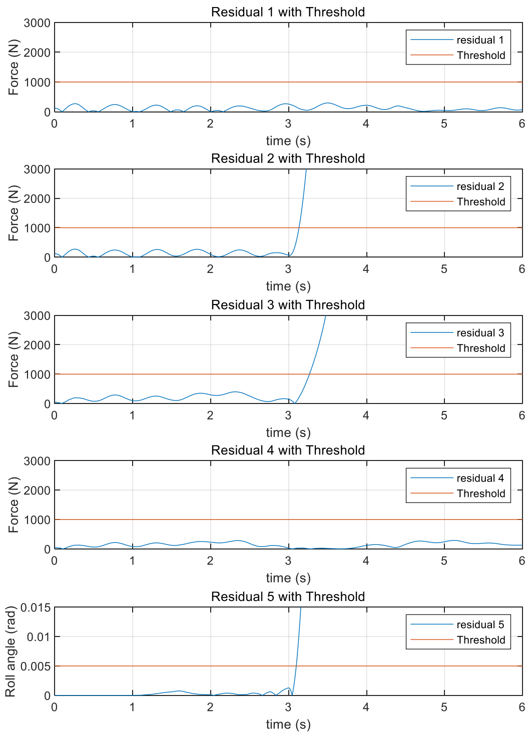


Figure 12: Residuals 1~5 with Body G sensor FR Fault

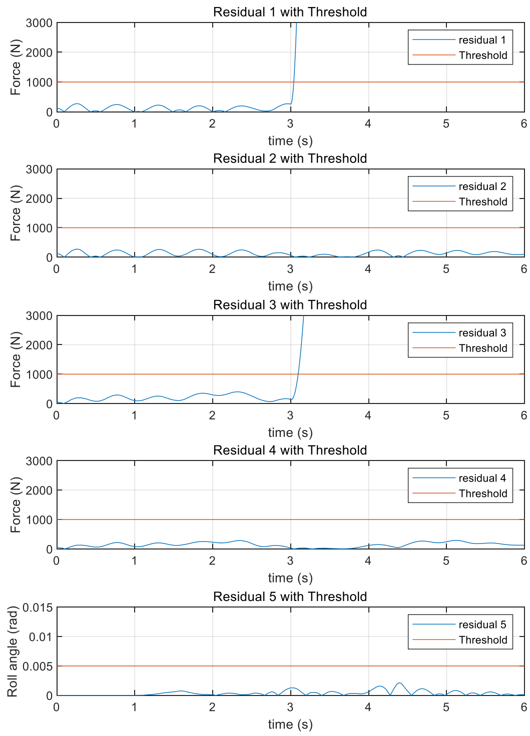


Figure 14: Residuals 1~5 with Wheel G sensor FL Fault

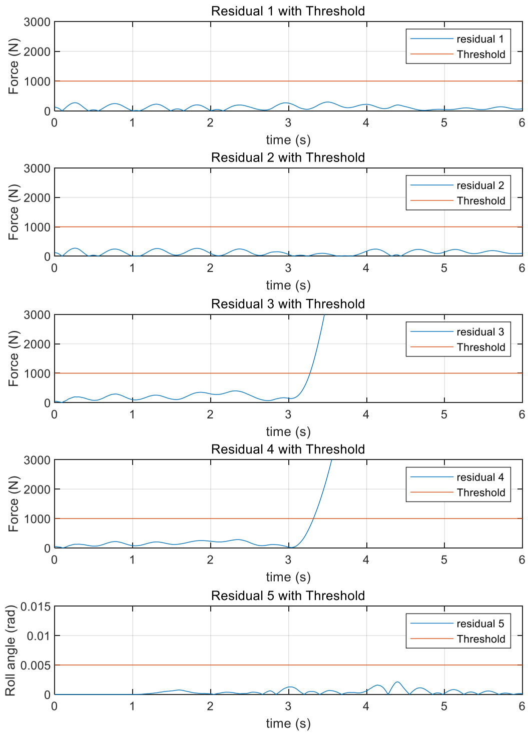


Figure 13: Residuals 1~5 with Body G sensor RR Fault

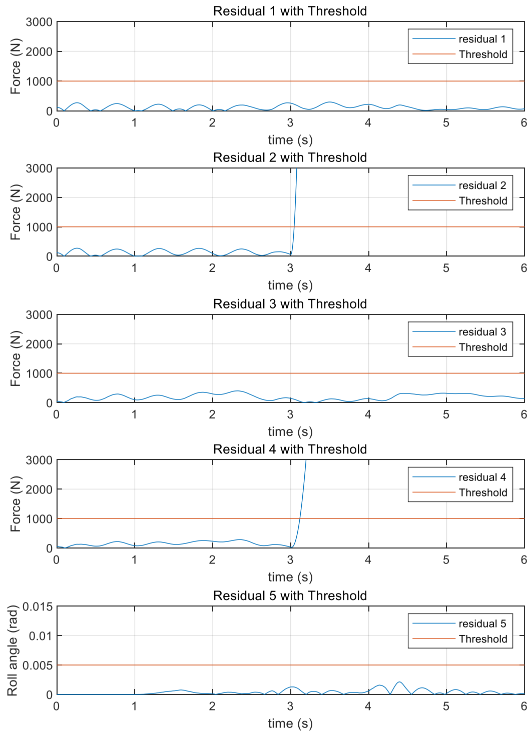


Figure 15: Residuals 1~5 with Wheel G sensor FR Fault

3. CONCLUSION

Fault detection and isolation algorithms for sensors in ECS systems have been developed using parity equations and incidence matrix methods. We also compare the actual vertical force and roll angle outputs of the Carsim vehicle dynamics simulator with the FDI algorithm's outputs. And we have simulated and analyzed the results of each faulty case. As a result, all five sensors were able to determine and isolate faults.

In this paper, we have assumed that there are three G sensors and two G sensors in the system considering the specifications of the actual ECS system. We expect the FDI algorithm to be applicable to the actual ECS system through the simulation results of the study. When applied to real-world systems, fault diagnosis is possible with the chassis control controller, and integrated controllers will provide robust control over fault tolerance. This research will be useful for researchers studying the fault diagnosis of an autonomous navigation system and ITS vehicle, where sensor measurement and fault diagnosis are important.

ACKNOWLEDGMENTS

This work is supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No. R7117-16-0132, The development of future technology for mixed traffic intersection integrated environmental information regarding the autonomous at the crossroads of the city)

REFERENCES

Kim H., Lee H., 2011. Fault-tolerant control algorithm for a four-corner closed-Loop air suspension system. Transactions on industrial electronics, Volume 58: 2866–4879.

Börner M., Strakey H., Weispfenning T., and Isermann R., 2002. Model based fault detection of vehicle suspension and hydraulic brake systems. Mechatronics, Volume 12, Issue 8: 999-1010.

Savaresi S., Poussot-Vassal C., Spelta C., Sename O., Dugard L., 2010. Semi-active suspension control design for vehicles.

Miller L., 1988. Tuning passive, semi-active, and fully active suspension systems, 2047-2053.

Gao H., Sun W., Shi P., 2010. Robust sampled-data H-inf control for vehicle active suspension systems. Transactions on control systems technology, Volume 18, Issue 1: 238-245.

Isermann R., 1995. Model based fault detection and diagnosis methods. American control conference.

Na W., Lee J., Lee H., 2017. Fault detection and isolation of semi-active system using integrated method. International Conference on control, automation and Systems, 1313-1317.

Rajamani R., 2012. Vehicle dynamics and control.: Modeling and Simulation.

Ryu J., Gerdes J., 2004. Estimation of vehicle roll and road bank angle, american control conference, 2110-2115.

Blanke M., Kinnaert M., Lunze K., and Stroswiecki M., 2006. Diagnosis and fault-tolerant control. Springer.

FUEL OPTIMAL SPEED PROFILE USING TRAFFIC INFORMATION AT INTERSECTION

Jooin Lee[a], Seongmin Ha[b], Hyeongcheol Lee[c]

[a],[b] Department of Electric Engineering, Hanyang University, 222, Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea
[c] Division of Electrical and Biomedical Engineering, Hanyang University, 222, Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea

[a]galaxyeagle@hanyang.ac.kr, [b]haha4100@hanyang.ac.kr, [c]hclee@hanyang.ac.kr

ABSTRACT

A connected vehicle can get data from the Intelligent Transportation System (ITS). This allows the car to estimate accurate vehicle states (speed, acceleration, instantaneous fuel consumption) and to use traffic information (signal, queue length, shockwave). Traffic information is very important for vehicle driving at urban intersections because traffic information (shock wave, queue length) means the effect of the driver. The fuel optimal speed profile algorithm must contain vehicle and traffic information. This paper proposes fuel optimal speed profile algorithm for the connected vehicle at the intersection. The algorithm uses the model predictive control method (MPC) to reflect traffic information constraints. The constraints change according to the traffic situation(lead vehicle, following vehicle). The algorithm is verified using a microscopic simulation tool (AIMSUN).

Keywords: fuel consumption, traffic information, intelligent transportation systems(ITS), model predictive control

1. INTRODUCTION

Environmental regulations for carbon dioxide emissions (CO2) and vehicle fuel consumption are being strengthened around the world. Carbon dioxide pollution is highly related to vehicle characteristics and driving patterns. Developing technology to improve the fuel efficiency of vehicles has reached the peak, which shows a limit to the improvement of CO2 emissions. As a result, many people have tried to improve the driving pattern or traffic system. The intelligent traffic system (ITS) and connected vehicles have received attention.

ITS collects accurate traffic information (signal, queue etc.) using various sensors (Camera, Radar etc.) and connected vehicles use the information by communicating with ITS server. The connected vehicles can reduce CO2 emissions by applying fuel optimal speed control method using traffic information. Fuel optimal speed control method previews future driving situation, and reduce idling, excessive decelerating or accelerating. This control algorithm can be very effective in driving in a city with high traffic volume.

In this paper, we propose fuel optimal speed control algorithm using traffic prediction information and estimated vehicle states. Model predictive control method (MPC) is applied to the algorithm. MPC method can determine the input by reflecting the estimated future information. The future traffic information (shock wave, queue length etc.) can be predicted through the microscopic traffic model using information collected at ITS. The future vehicle states (speed, acceleration, instantaneous fuel consumption) can be estimated through the longitudinal vehicle dynamics model using the sensor values measured in the vehicle. Connected vehicles obtain all the information via V2X communication and CAN communication. Vehicle sensor data such as speed and acceleration is obtained via CAN communication. WAVE is used as the V2X communication method. In addition, high-precision vehicle position information can be received via UWB and GPS sensor.

The fuel optimal speed control algorithm is verified using the AIMSUN microscopic traffic simulation tool and Python. We modeled two consecutive intersection models reflecting the actual traffic environments (signal, road, traffic flow). The CO2 emissions are used as an evaluation index.

2. VEHICLE STATE ESTIMATION MODEL

Vehicle states are required to determine the fuel optimal speed profile. Vehicle conditions include speed, acceleration, and instantaneous fuel consumption. A longitudinal vehicle dynamics model and simplified fuel consumption model are used to estimate the states of the vehicle.

2.1. Longitudinal Vehicle Dynamics Model

The longitudinal dynamics model of the vehicle consists of the traction force and the resistance force (Thomas D. Gillespie 1992). The traction force includes the engine driving force considering powertrain efficiency. And the resistance force includes the air resistance force and the gradient resistance force. Actual vehicle dynamics are complex, but we use a simplified model because of the computational complexity. The equations for the constructed model are (1) ~ (3).

$$m\frac{dv(t)}{dt} = F_{tract}(t) - F_{aero}(t) - F_{grade}(t) \quad (1)$$

$$F_{aero} = \frac{1}{2}\cdot \rho \cdot A_f \cdot C_d \cdot v^2(t) \quad (2)$$

$$F_{grade} = m \cdot g \cdot \sin(\alpha) \quad (3)$$

2.2. Track-to-Track Sensor Fusion

The track-to-track sensor fusion method is used to estimate the exact speed, acceleration (Hernsoo Hahn 2008). Accelerometer, GPS, and wheel speed sensors are fused using sensor fusion method. The Bar-Shalom-Campo formula, one of the sensor fusion methods, is used. The equations for the constructed model are (4).

$$\begin{aligned}\hat{x}^+[k] = {} & (P^+[2,k] - P^+[21,k])(P^+[1,k] + P^+[2,k] - P^+[12,k] - P^+[21,k])^{-1}\hat{x}^+[1,k] \\ & + (P^+[1,k] - P^+[12,k])(P^+[1,k] + P^+[2,k] - P^+[12,k] - P^+[21,k])^{-1}\hat{x}^+[2,k]\end{aligned} \quad (4)$$

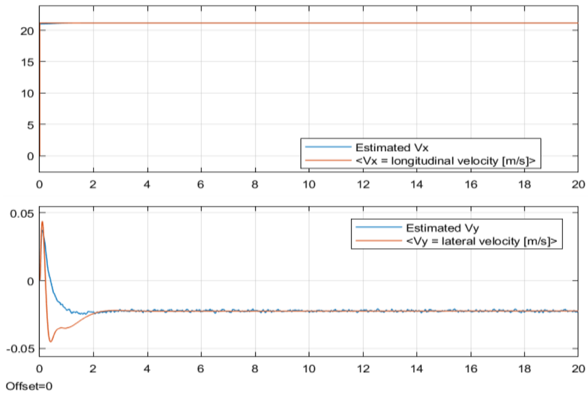


Figure 1: Longitudinal Speed Vx and Lateral Speed Vy estimation using sensor fusion method with GPS and Accelerometer

2.3. Simplified Fuel Consumption Model

The simplified fuel consumption model is applied to the algorithm (Engin Ozatay 2013) (Hesham Rakha 2011). Simplified fuel consumption model uses the traction force and speed as variables. The actual fuel consumption model has nonlinear behavior. The simplified model is applied because it is easier to reduce computation time and to interpret optimization problems. The equations for the constructed model are (5).

Figure 2 and Figure 3 shows the estimated speed and fuel consumption data in the FTP-75 city driving cycle. As shown in Figure 2 and Figure 3, the longitudinal vehicle dynamics model and the simplified fuel consumption model are compared with the actual data, and it is confirmed that the error is very small. Therefore, it can be verified that the model is valid.

$$\dot{m}_f = \frac{m}{eH_{L^\bullet drive}} \times a_{input}(t) \times v(t) + \frac{P_{Loss} \times V_d}{4\aleph \times eH_L}\left(\frac{\geq(t)}{R_{wh}}\right) \times v(t) + \dot{m}_{idle} \quad (5)$$

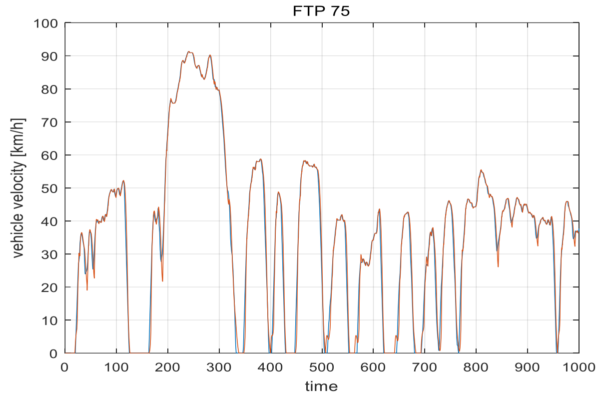


Figure 2: Vehicle velocity estimation result for FTP-75 city cycle

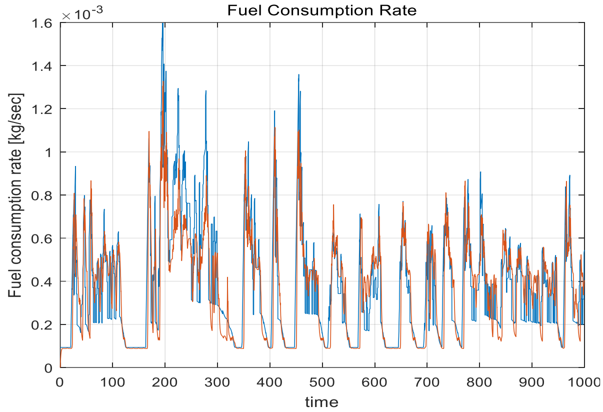


Figure 3: Instantaneous fuel consumption estimation result for FTP-75 city cycle

3. TRAFFIC STATE ESTIMATION MODEL

The traffic state estimation model estimates the queue length and the queue latency time. The urban intersection has interrupted traffic flow characteristics in which the traffic flow is cut off by a traffic light. This traffic flows characteristic creates a queue. Therefore, the queue can be used as an indicator for determining the traffic situation ahead. We estimate the queue and queue latency time using the shockwave theory of traffic engineering.

3.1. Queue Length Estimation Model

Queue length is calculated for vehicles that cannot cross the intersection within the remaining traffic light time. The queue length is estimated using the number of vehicles, vehicle length, and the average headway distance. The estimation value is calculated and updated every unit time. The equations for the constructed model are (6).

$$Q_{Length} = (d_{clearance} \times N_{stop}) + \sum_{n=0}^{N_{stop}} L_{veh} \quad (6)$$

3.2. Queue Latency Time Estimation Model

The queue latency time is the time at which the last vehicle in the queue begins to move (Luc Binette 1985) (Tom V.Mathew 2007). The queue latency time is used to predict the movement of the queue ahead. The shock wave theory of traffic engineering is used to estimate the queue latency time. The interrupted traffic flow by the traffic lights changes the traffic flow characteristics, and the change produces a shock wave. The shock wave propagation speed is the speed of the shock wave.

It is assumed that the average speed and the traffic density have a linear relationship using the Greenshield linear model. We use the Greenshield model to calculate the maximum traffic density and maximum traffic volume. These data are used to estimate the shock wave speed. The queue latency time is estimated using the shock wave speed and the queue length. The equations for shock wave speed are (7). The equations for the constructed model are (8).

$$W = \frac{q_m - 0}{k_m - k_j} \tag{7}$$

$$T = \frac{Q_{Length}}{W} \tag{8}$$

4. FUEL OPTIMAL SPEED PROFILE

The fuel consumption model is composed of a cost function to design a controller that can drive a vehicle with minimum fuel consumption. Fuel optimal speed profile uses a model predictive control method that can easily reflect future preview information about the vehicle and traffic. By applying constraints on vehicle and traffic conditions, the controller calculates optimal speed so that it does not violate the laws of physics and the driver does not feel any sense of heterogeneity (Junbo Jing 2014).

4.1. Leading Vehicle Identifying Model

Vehicles driving at intersections can be classified into two types as the leading vehicle and the following vehicle. Leading vehicle and following vehicle have different control purposes. The lead vehicle is determined by two processes. First, it is determined whether it is possible to pass the intersection during the remaining displayed time. After that, the headway vehicle is determined by analyzing the interval between the vehicles that cannot pass. If the distance between vehicles is wider than the safety headway distance, it is decided to be the leading vehicle.

$$d_{gap} \geq d_{safety} \tag{9}$$

In order to determine the vehicle safe headway distance to the front vehicle, several simulations have been carried out to obtain safe headway distance data that the vehicle normally tries to maintain with the preceding vehicle depending on the speed. And the acquired safe headway distance data is linearized and used.

4.2. Model Predictive Control

A fuel consumption model is constructed as a cost function to design a controller that minimizes fuel consumption. The model predictive control method is applied by using the corresponding cost function. The model predictive control method calculates the optimal solution using the preview information in the discrete time domain. Simplified fuel consumption model is derived in quadratic form and applied to cost function.

$$\underset{x[k],\,u[k]}{Minimize}\; J(x[k],u[k]) = \underset{x[k],\,u[k]}{Minimize} \sum (\dot{m}_f) \cdot h \tag{10}$$

The model predictive controller applies three constraints so that the driver does not feel a sense of heterogeneity without violating the laws of physics. The first constraint applies to longitudinal vehicle dynamics. Longitudinal vehicle dynamics are constraints on the physical movement of the vehicle.

$$\therefore\; x[k+1] = \widetilde{A}_d \cdot x[k] + \widetilde{B}_d \cdot u[k] + \widetilde{B}_{dd} \tag{11}$$

Second, the maximum acceleration that the engine can generate and the maximum deceleration that can be applied by the brake are applied as constraints. This is a constraint on the characteristics of the vehicle.

$$-\frac{1}{m_{eq}} \times F_{brake\_\max} \leq a_{traction} \leq T_{\max} \times \frac{\eta \times \gamma}{m_{eq} \times R_{wh}} \tag{12}$$

Thirdly, jerk which is the change of acceleration is applied as a constraint. Jerk is a constraint on the sense of heterogeneity that the driver feels.

$$-1.5 \leq u[k] - u[k-1] \leq 1 \tag{13}$$

The above three constraints are applied to the model predictive controller.

4.3. Leading Vehicle Control Strategy

Since the leading vehicle does not have a subject that can follow in front, it controls traffic conditions such as traffic lights and queues. Therefore, the traffic condition is controlled by adding constraints. The control strategy of the lead vehicle is divided into deceleration strategy and acceleration strategy.

First, the deceleration strategy controls the speed by considering the remaining signal time and the queue latency time. The remaining signal time and the queue latency time are used to calculate the time required to reach the stop line. Then, a position controller using the distance to the stop line and the estimated queue length is applied.

$$\underset{x[k],\,u[k]}{Minimize}\left\{\left(\frac{1}{2}U^{T}QU+RU+N\right)+\left(\frac{1}{2}E_{leading}^{T}PE_{leading}\right)\right\}$$
(14)

$$E_{leading}=D_{stop}-x_{ego}$$
(15)

Second, the acceleration strategy controls the speed to reach the target speed. A target speed is required for the vehicle to reach a constant speed. In this paper, we set the speed as the target speed when the individual vehicles enter the intersection service section. This is because it is usually the speed at which drivers usually want to drive

$$\underset{x[k],\,u[k]}{Minimize}\left\{\left(\frac{1}{2}U^{T}QU+RU+N\right)+\left(\frac{1}{2}E_{v}^{T}PE_{v}\right)\right\}$$
(16)

$$E_{v}=\dot{x}_{target}-\dot{x}\ \ (\therefore \dot{x}=velocity)$$
(17)

4.4. Following Vehicle Control Strategy

The following vehicle maintains a safe distance from the preceding vehicle. The following vehicle control strategy controls the position and the speed to follow. The controller was designed by adding a term that can follow the preceding vehicle while maintaining the safe distance to the cost function. The safe distance is determined through several simulations.

$$\underset{x[k],\,u[k]}{Minimize}\left\{\left(\frac{1}{2}U^{T}QU+RU+N\right)+\left(\frac{1}{2}E_{x}^{T}PE_{x}\right)\right\}$$
(18)

$$E_{x}=d_{gap}-d_{safety}=(x_{pre}-x_{ego})-d_{safety}$$
(19)

5. RESULT & CONCLUSION

5.1. Simulation Environment

The algorithm is verified in a simulation environment that reflects actual urban intersection environment. We recorded the video at the actual intersection and measured traffic signal time and traffic volume. The measured data (speed limit, road shape, etc.) was used for modeling using a microscopic traffic simulator AIMSUN. Figure4 shows the measured signal time information.

Figure 4: two consecutive intersection environments for modelling (Korea, Seoul, Gangnam)

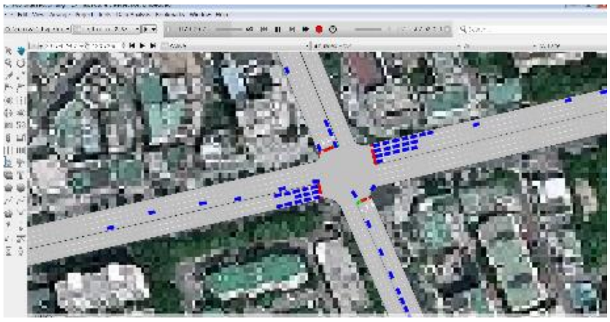


Figure 5: two consecutive intersection environments for modelling (Korea, Seoul, Gangnam)

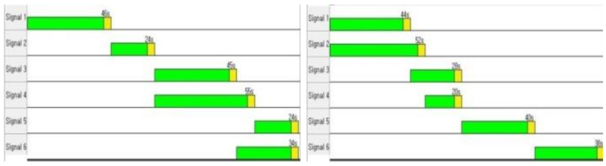


Figure 6: two consecutive intersection environments for modelling (Korea, Seoul, Gangnam)

To verify this algorithm, we compare the simulation result with the general driver model provided by AIMSUN. Fuel consumption is used as a comparison target as a result of the simulation.

AIMSUN worked with Python 2.7.13, a programming tool, to control the vehicles in the intersection environment modeled by AIMSUN. We used the CVXPY solver provided by the Python library for optimization.

5.2. Simulation Result

The results of the comparison between the general driver model (Gipp's model) and the fuel consumption optimum speed profile algorithm are show in the figure7 and figure8. Figure7 shows the comparison of the speed profile of the vehicle stopped at two intersections. Figure8 shows the comparison of the velocity profile of the decelerating vehicle considering the queue at one intersection. As a result, it has been confirmed that the acceleration or deceleration pattern of a vehicle passing through intersections can be reduced.

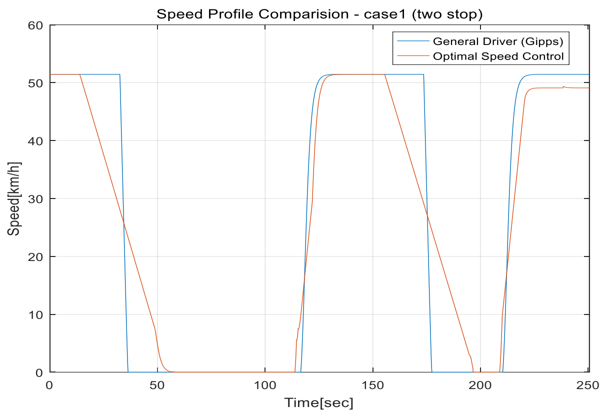


Figure 7: Simulation Results: speed profile through two consecutive intersections.

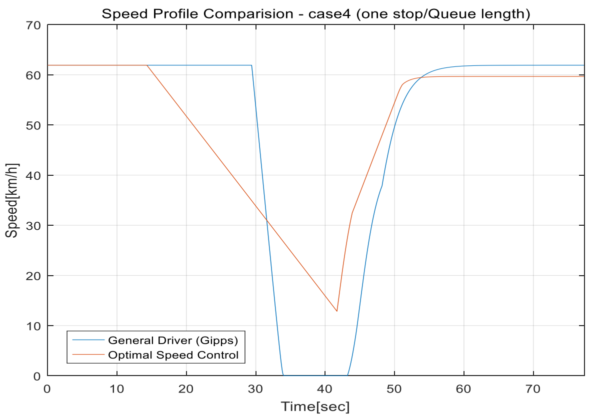


Figure 8: Simulation Results: speed profile through one intersection.

Table 1 shows the results of comparing fuel consumption with changing traffic volume. The fuel-optimal speed profile algorithm shows an average 9% improvement in fuel consumption over the general driver model.

Table 1: Simulation Results: fuel efficiency improvement

Number of Vehicle [vehicle]	Fuel consumption (General Driver Model) [kg]	Fuel consumption (Fuel-optimal speed profile algorithm) [kg]	Fuel efficiency improvement rate [%]
76	2.92	2.69	7.87
90	3.18	2.88	9.43
108	3.87	3.51	9.3
450	15.8	14.3	9.49

As can be seen from the results of the above speed profile, it is possible to improve fuel efficiency by reducing the rapid acceleration or deceleration by predicting traffic situation information.

5.3. Conclusion

In this paper, we propose a fuel-optimal speed profile algorithm to reduce CO2 emissions from intersections. It was confirmed that the fuel efficiency at intersections increased by about 9% in the environment where various traffic was considered. Using individual vehicle information and traffic information obtained from the intelligent transportation system, we predicted the queue length and queue latency time ahead and set up a driving strategy suitable for traffic simulation. Therefore, it is possible to plan a driving strategy in which the fuel consumption and the driver do not feel a sense of heterogeneity.

However, the problem of designing the cut-in situation which causes the error in the simulation and a large amount of computation load according to the MPC characteristic are seen as problems to be solved in the future.

ACKNOWLEDGMENTS

This work is supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No. R7117-16-0132, The development of future technology for mixed traffic intersection integrated environmental information regarding the autonomous at the crossroads of the city)

REFERENCES

Hesham Rakha., and Raj kishore Kamalanathsharma., 2011. Eco-driving at signalized intersections using V2I communication. Intelligent Transportation Systems , Washington.

Matthew Barth., Sindhura Mandava., Kanok Boriboonsomsin., and Haitao Xia., 2011. Dynamic ECO-driving for arterial corridors. Integrated and Sustainable Transportation System (FISTS), Vienna..

Tien-Wen Sung., Liang-Cheng Shiu., Fu-Tian Lin., and Chu-Sing Yang., 2015. A Speed Control Scheme of Eco-Driving at Road Intersections. Third International Conference on Robot, Vision and Signal Processing (RVSP), Kaohsiung.

Hao Yang., Hesham Rakha., and Mani Venkat Ala., 2016. Eco-Cooperative Adaptive Cruise Control at Signalized Intersections Considering Queue Effects. IEEE Transactions on Intelligent Transportation Systems, China.

Raj kishore Kamalanathsharma., and Hesham Rakha., 2014. Leveraging connected vehicle technology and telematics to enhance vehicle fuel efficiency in the vicinity of signalized intersections. Journal of Intelligent Transportation Systems, London.

Md. Abdus Samad Kamal., Masakazu Mukai., Junichi Murata., and Taketoshi Kawabe., 2013. Model predictive control of vehicles on urban roads for

improved fuel economy. IEEE Control Systems Society, India.
M.A.S. Kamal., M. Mukai., J. Murata., and T. Kawabe., 2010. On board eco-driving system for varying road-traffic environments using model predictive control. IEEE International Conference on Control Applications , Yokohama, Japan.
M.A.S. Kamal., M. Mukai., J. Murata., and T. Kawabe., 2011. Ecological driving based on preceding vehicle prediction using MPC. IFAC Proceedings Volumes 44.1 , Milano, Italy.
TD Gillespie., 1992. Fundamentals of Vehicle Dynamics. Society of Automotive Engineers (SAE).
Engin Ozatay., Umit Ozguner., Dimitar Filev., and John Michelini., 2013. Analytical and numerical solutions for energy minimization of road vehicles with the existence of multiple traffic lights. IEEE 52nd Annual Conference on Decision and Control (CDC), Florence , Italy.
Nam Heo., Heesu Youn ., Jooin Lee,. and Hyeongcheol Lee., 2016. Development of Optimal Traffic Signal Control Algorithm Using each Vehicle Characteristics at the Downtown Intersection. The Korean Society Of Automotive Engineers Annual Autumn Conference, Washington.
P.G.Gipps., 1981. A behavioural car-following model for computer simulation. Transportation Research Part B: Methodological.
P.G.Gipps., 1986. A model for the structure of lane-changing decisions. Transportation Research Part B: Methodological.
J Barceló., and J Casas., 2005. Dynamic network simulation with AIMSUN. Simulation approaches in transportation .
Hesham Rakha., Kyoungho Ahn., Kevin Moranand., BartSaerens., and Eric Van denBulck., 2011. Virginia tech comprehensive power-based fuel consumption model: Model development and testing. Transportation Research Part D: Transport and Environment 16.7, USA.
Hiraoka Toshihiro., Taketoshi Kunimatsu., Osamu Nishihara., and Hiromitsu Kumamoto., 2005. Modeling of driver following behavior based on minimum-jerk theory. 12th World Congress ITS, San Francisco.
Hernsoo Hahn., Hanseok Ko., Osamu Nishihara., and Sukhan Lee., 2008. Multisensor Fusion and Integration for Intelligent Systems. the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems.
Luc Binette., Michael A. Dopita., and Ian R. Tuohy., 1985. Radiative shock-wave theory. II - High-velocity shocks and thermal instabilities. Astrophysical Journal.USA.
Tom V. Mathew., and K V Krishna Rao., 2007. Traffic stream models. National Programme on Technology Enhanced Learning (NPTEL). India.
Junbo Jing., 2014. Vehicle Fuel Consumption Optimization using Model Predictive Control based on V2V communication, The Ohio State University, Ohio.

A STUDY ON OPERATING CHARACTERISTICS OF A 48V ELECTRIC SUPERCHARGER AND P0 CONFIGURATION MOTOR FOR FUEL ECONOMY IMPROVEMENT

Yonghyeok Ji[(a)], Junyong Park[(b)], Hyeongcheol Lee[(c)]

[(a),(b)]Department of Electrical Engineering, Hanyang University, 222, Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea
[(c)]Department of Electrical and Biomedical Engineering, Hanyang University, 222, Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea

[(a)]youg0839@hanyang.ac.kr, [(b)]jyp0802@hanyang.ac.kr, [(c)]hclee@hanyang.ac.kr

ABSTRACT

Recently, various researches and developments of 48V mild hybrid system are being actively studied due to advantages in terms of cost. A 48V hybrid system with various configuration has been proposed, and some companies have proposed the system that downsizes engine by using a P0 configuration motor and electric supercharger. The P0 configuration motor and electric supercharger all support the power of an engine, but the research on how each system operates in order to improve fuel economy is still lacking. Therefore, in this paper, the operating characteristics for improving the fuel economy of each system used as the power assistant of the engine are compared. In order to compare the most optimized behavior for fuel economy, we used *Dynamic Programming (DP)*, a global optimization technique, and Dynamic programming simulation was performed in Matlab environment.

Keywords: 48V mild hybrid system, P0 configuration motor, BSG, Electric supercharger, Dynamic programming, fuel economy

1. INTRODUCTION

As fuel economy and emission gas regulations are strengthened globally, environmentally friendly vehicles of various types (BEV, FCEV, (P)HEV, etc.) are being introduced to the market. However, due to the problem of the high-cost electric drive system or charging infrastructure, it is still difficult for vehicles to replace the demand for existing internal combustion engine vehicles. According to this trend, a mild hybrid system based on a 48V power supply has recently attracted attention as a new alternative and many kinds of research about 48V mild hybrid system are being conducted. (Malte Kuypers 2014, Mark Schudeleit and Christian Sieg 2015, Anthony Rick and Brain Sisk 2015, Andreas Baumgardt and Dieter Gerling 2015a,b, Anthony Rick and Brain Sisk 2015, Zifan Liu and Andrej Ivanco 2016, Junyong Park and Taeho Park 2017)

Hybrid drive systems are generally divided into P0~P4 according to the location of the motor (Figure 1). Many studies have been done on the CO2 reduction effect and cost efficiency for each configuration. (Dr. Ing. Olivier COPPIN 2016, Ran Bao and Victor Avila 2017, Thomas Eckenfels and Florian Kolb 2018)

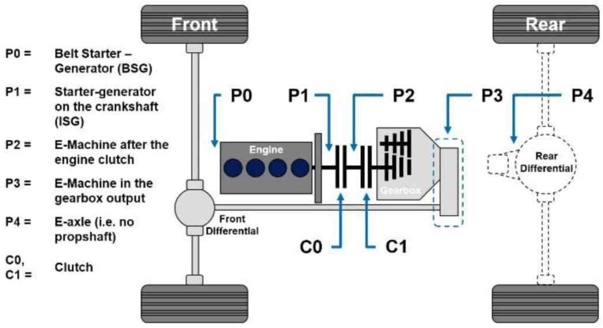


Figure 1: P0~P4 Configuration

And recently, research and prototypes of 48V electric superchargers have been actively developed to maximize the downsizing effect of an engine with low electric motor power. Ricardo Martinez-Botas and Apostolos Pesiridis (2011) introduced an electric supercharger as the boosting option for downsized engines. Aaron Isenstadt and John German (2016) provided an analysis of turbocharged, downsized gasoline engine technology developments and trends including E-boosting and 48V hybrid systems. Bo Hu and Chunlin Chen (2017) introduced the trends about a 48V mild hybrid system and 48V e-boosting, and referred that using both systems together can provide a synergy effect. DR.-ING. Richard Aymanns and DR.-ING. Tolga Uhlmann (2018) investigated the benefits that a 48V system can offer for electric supercharging when compared with a standard 12V system. Prasad Sajjan Divekar and Beshah Ayalew (2010) proposed the system that configured with an electric supercharger and turbo-generator for a diesel engine, and control strategy for the system. Bryn Richards and Tran Hoang Hiep (2015) and Ivan Filidoro (2015) showed that downsizing the engine with electric supercharger can provide enough responsiveness and desirability. Hiep Hoan Tran and Bryn Richard (2016) developed a performance specification for a 48V electric supercharger to satisfy a range of downsized gasoline engine application. Stephan Tavernier and Samuel Equoy (2013) optimized 48V PMSM motor design for an electric supercharger. All these studies show that electric superchargers contribute to the acceleration performance and fuel efficiency of the vehicle.

In order to take advantage of these electric superchargers, some companies have recently proposed a system in which an electric supercharger is applied to a 48V mild hybrid system of a P0 configuration. The P0 configuration motor and electric supercharger all support the power of an engine, but the research on how each system operates in order to improve fuel economy is still lacking. Therefore, in this paper, the operating characteristics for improving the fuel economy of the electric supercharger and the P0 configuration motor are compared. To do this, we apply Dynamic programming to the gasoline engine with P0 configuration motor, the gasoline engine with electric supercharger, and the gasoline engine with an electric supercharger and P0 configuration motor. After that, we investigate and compare the result of Dynamic programming.

This paper is organized as follows. Section 2 describes the target vehicle in this paper, and Section 3 describes the program setting and vehicle model for performing Dynamic programming. In Section 4, the results of Dynamic programming using vehicle models of Section 3 were investigated. The simulation was performed in the linear acceleration situation in the fixed gear to compare the operating characteristics of each device according to the operating point of the engine.

2. TARGET VEHICLE

In this paper, we cover the following three vehicle configuration

1. Gasoline Engine + Electric supercharger
2. Gasoline Engine + P0 configuration motor
3. Gasoline Engine + Electric supercharger + P0 configuration motor

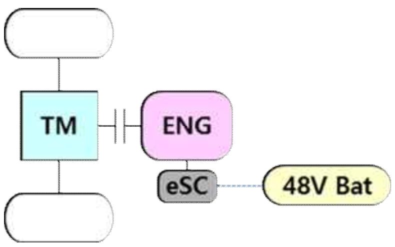


Figure 2: Engine + Electric Supercharger Configuration

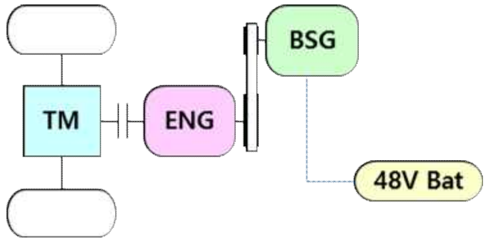


Figure 3: Engine + P0 Motor Configuration

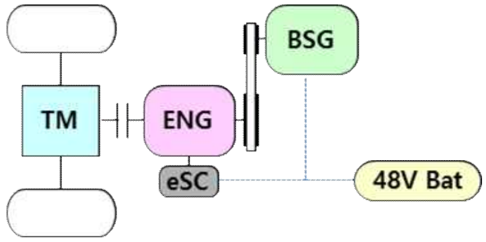


Figure 4: Engine + Electric Supercharger + P0 Motor Configuration

As shown in the figure, the motor with P0 configuration is a belt-driven starter and generator (BSG), which is connected to the engine crankshaft by a belt. In the figure, black solid lines represent mechanical connection structures and blue dotted lines represent electrical connection structures. Table 1 shows the specifications of the main components of the above vehicle configurations.

Table 1: Specification of vehicle component

Component	Specification
Engine	1500 [cc] Gasoline Engine
P0 Motor	10 [kW] PMSM
Electric Supercharger	5 [kW] PMSM
Battery	48 [V] / 20 [Ah] Lithium-Ion

3. MODEL FOR DYNAMIC PROGAMMING

We used Dynamic programming to compare the behavior that represents optimal fuel economy for each vehicle configuration shown in Figures 2~4. Dynamic programming is a method of global optimization strategies that find optimal solutions of optimal control problems based on Bellman's principle of optimality. For the hybrid drive system, Dynamic programming is mainly used to analyze the optimal fuel economy results by investigating all possible paths of the vehicle system in a given driving cycle in advance. Therefore, in this paper, Dynamic programming is used to investigate the operating characteristics that represent the optimal fuel efficiency of the targeted vehicle configuration.

3.1. Dynamic programming with dpm.m function

In this paper, we use Matlab function dpm.m for Dynamic programming. Dpm.m is a Dynamic programming algorithm introduced by Olle Sundström and Lino Guzzella (2009) to solve the general optimization problem. To solve the optimal control problem using the Dpm.m function, the user must define the Main.m function and the Model function.m function as shown in Figure 5. The Main.m function sets the range and grid of input, state for Dynamic programming and executes dpm.m. Model function.m uses the given inputs to calculate the cost and the state of the next step and returns these values to dpm.m. In this paper, the model function is the vehicle model that calculates the fuel consumption and the SOC variation according to the power distribution ratio and the electric supercharger speed.

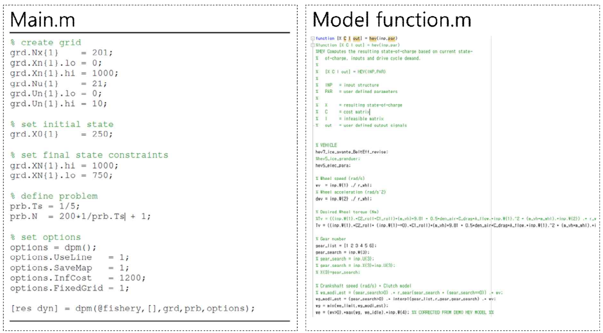


Figure 5: Main.m & Model function.m for Dynamic programming

3.2. Main.m function setting

As mentioned in Section 3.1, we need the Main.m function and the Model function.m to use dpm.m. This section introduces the settings for the Main.m function.

First, we set the input as power distribution ratio (u_1) between the engine and the motor, and the electric supercharger speed (u_2). The power distribution ratio of the engine and the motor is defined as Equation (1).

$$u_1 = T_{m,Crank}/T_{tot} \tag{1}$$

Where $T_{m,Crank}$ is the motor torque at crankshaft, i.e., the pulley ratio is considered, and T_{tot} is the total desired torque at the crankshaft for the vehicle to meet the driving condition.

The state is set to SOC, and the range and grid of SOC are shown in Table 2.

State	Range	Grid
SOC	0.3 ~ 0.9	21 points

In the Main.m function, we can also set the speed and acceleration profile for the simulation. Details of speed and acceleration profile which is used in this paper are introduced in Section 4.1.

3.3. Model function.m setting (Quasi-static vehicle model)

This section introduces the Model function.m for Dynamic programming. As mentioned above, the model function in this paper is a vehicle model, and a quasi-static vehicle model that reduces the dynamics characteristics and reflects nonlinear characteristics as a map is used because the long computation time is consumed when the model becomes complicated due to the characteristics of Dynamic programming. The quasi-static vehicle model calculates fuel consumption and battery SOC variation according to given input values (u_1, u_2), given vehicle speed and acceleration profile, and the conceptual diagram of the process is shown in Figure 6.

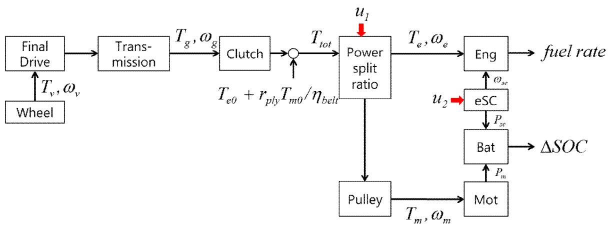


Figure 6: Conceptual diagram of quasi-static vehicle model for Dynamic programming

As shown in the figure, after the desired torque (T_v) at the wheel is converted to the torque at the crankshaft via the final drive gear and the transmission. And then it is summed with the driving resistance torque of the motor and engine (T_{m0}, T_{e0}) to calculate the total desired torque (T_{tot}) at the engine crankshaft. The calculated total desired torque is divided by the torque of the motor and the engine by u_1. And it determines the power of the motor and engine together with speed of the motor and engine. The determined power of the motor (P_m) becomes the output power of the battery together with the power consumed by the electric supercharger (P_{eSC}), and the SOC variation amount is calculated by the output power of the battery. The fuel consumption is determined by the determined engine torque and speed. The equations for the above process is shown step by step as follows.

3.3.1. Desired torque and speed at wheel

The desired torque at the wheel is determined by the longitudinal dynamics of the vehicle as shown in Equation (2). In the equation below, r_{whl} is the wheel radius, $C_{roll,1}$, $C_{roll,2}$ is the rolling resistance coefficient, m_{veh} is the weight of the vehicle, g is the gravitational acceleration, C_{aero} is the air resistance coefficient and

ω_v is the wheel speed. Where ω_v and $\frac{d\omega_v}{dt}$ are the values defined in the Main.m function as the condition of the problem.

$$T_v = r_{whl}\left\{(C_{roll,1} + C_{roll,2}\omega_v)m_{veh}g + (C_{aero}\omega_v^{\ 2} + m_{veh}\frac{d\omega_v}{dt})\right\} \tag{2}$$

3.3.2. Desired torque and speed at transmission input shaft

The desired torque and the desired speed in the transmission input shaft are determined in consideration of the gear ratio and the desired torque and speed at the wheel. In the Equations (3) and (4), r_{FD}, r_{TM} are respectively the final drive gear ratio and the transmission gear ratio, η_{TM} is the efficiency of the transmission, and ω_{e_limit} is the maximum engine speed.

$$T_g = \begin{cases} \dfrac{T_v}{r_{FD}r_{TM}\eta_{TM}} & (T_v > 0) \\ \dfrac{T_v}{r_{FD}r_{TM}}\eta_{TM} & (T_v \le 0) \end{cases} \tag{3}$$

$$\omega_g = min(\frac{\omega_v}{r_{FD}r_{TM}}, \omega_{e\_limit}) \tag{4}$$

3.3.3. Desired torque and speed at the engine crankshaft

The total desired torque of the powertrain for power distribution is calculated at the crankshaft of the engine. The total desired torque is calculated by adding the desired torque of the transmission input shaft and the driving resistance torque of the engine and the motor as shown in Equation (5). In the equation below, T_{tot} is the total desired torque at the engine crankshaft, r_{ply} is the gear ratio of the pulley, T_{e0}, T_{m0} are the drive resistance torque of the engine and motor respectively, and η_{belt} is the belt efficiency.

$$T_{tot} = T_{e0} + r_{ply}T_{m0}/\eta_{belt} + T_g \tag{5}$$

The driving resistance torque of the engine and the motor is calculated as shown in Equations (6) and (7) respectively. In the equations below, J_{eng}, J_m are the rotational inertia of the engine and the motor, $T_{e,fric}$ is the engine friction torque, ω_e, ω_{eSC}, ω_m are the speed of the engine, electric supercharger, and motor. And $T_{e,fric}$ is the value determined according to the speed of the engine and electric supercharger as shown in Equation (6) and Figure 7.

$$T_{e0} = J_{eng}\frac{d\omega_e}{dt} - T_{e,fric}(\omega_{eSC}, \omega_g) \tag{6}$$

$$T_{m0} = J_m\frac{d\omega_m}{dt}, \quad \omega_m = r_{ply}\omega_g \tag{7}$$

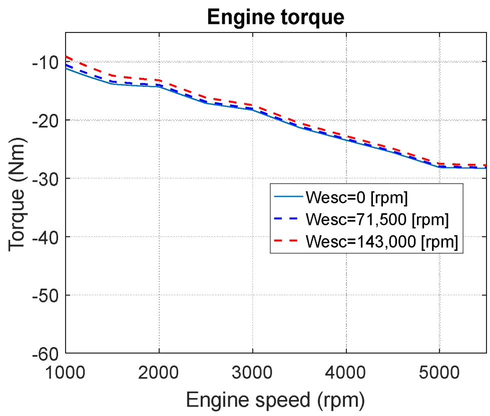


Figure 7: Engine friction torque by the engine and electric supercharger speed

3.3.4. Motor torque and speed determination

u_1 is defined as Equation (1). And the motor is connected to the crankshaft by a pulley, so the torque and speed of the motor are determined as shown in Equation (8) and (9).

$$T_m = \begin{cases} \dfrac{u_1 T_{tot}}{r_{ply}\eta_{belt}} & (T_{tot} > 0) \\ \dfrac{u_1 T_{tot}}{r_{ply}}\eta_{belt} & (T_{tot} \le 0) \end{cases} \tag{8}$$

$$\omega_m = r_{ply}\omega_e \tag{9}$$

At this point, the maximum and minimum torques of the motor are represented by Equation (10). Where T_{m_min}, T_{m_max}, T_{e_max}, and T_{regen_max} are the minimum and maximum torque of the motor, the maximum torque of the engine, and the maximum regenerative torque during braking.

$$\begin{aligned} &max(T_{m\_min}, T_{tot} - T_{e\_max}) \le T_m \le T_{m\_max} \quad (T_{tot} > 0) \\ &T_{regen\_max} \le T_m \le T_{m\_max} \quad (T_{tot} \le 0) \end{aligned} \tag{10}$$

3.3.5. Engine torque and speed determination

The torque of the engine is determined as the difference between the total desired torque and the motor torque determined by the power distribution, like Equation (11). In Equation (12), T_b is the friction brake torque, which serves to compensate the desired torque that the motor cannot cover during braking.

$$T_e = T_{tot} - T_m r_{ply} \eta_{belt} \quad (T_{tot} > 0) \tag{11}$$

$$T_b = T_{tot} - T_m r_{ply} \eta_{belt} \quad (T_{tot} \le 0) \tag{12}$$

The maximum and minimum torques of the engine are as follows.

$$0 \le T_e \le T_{e\_max} \quad (T_{tot} > 0) \tag{13}$$

The engine speed is basically the same as the transmission input shaft speed, but the minimum speed is limited to the idle speed. In this paper, the idle speed is about 90 [rad /s].

$$\omega_e = min(\omega_g, \omega_{e,idle}) \tag{14}$$

3.3.6. Electric supercharger torque and speed determination

As shown in Equation (15), the torque of the electric supercharger is determined by the load torque map which has the operating point of the engine as variables. The speed of the electric supercharger, ω_{eSC} determined by the input variable u_2

$$T_{eSC} = T_{eSC\_Load}(T_e, \omega_e, \omega_{eSC}) \tag{15}$$

$$\omega_{eSC} = u_2 \tag{16}$$

3.3.7. Fuel consumption determination

Since the torque and speed of the engine are calculated in Section 3.3.5., the fuel consumption can be obtained. The fuel consumption is determined by the map which has torque and speed of the engine, the speed of the electric supercharger as variables.

$$\Delta m_{fuel} = \Delta m_{fuel\_map}(T_e, \omega_e \omega_{eSC}) \tag{17}$$

3.3.8. Battery model and SOC determination

The power consumed by motor and electric supercharger, P_m, is given by Equation (18). In that equation, P_m is the power consumed by the motor and the electric supercharger, and $P_{m,P0}, P_{m,eSC}$, η_m, η_{eSC} are the power consumption and efficiency of the motor and the electric supercharger respectively.

$$\begin{aligned} &P_m = P_{m,P0} + P_{m,eSC} \\ &P_{m,P0} = \begin{cases} T_m \omega_m / \eta_m & (T_m \ge 0) \\ T_m \omega_m \eta_m & (T_m < 0) \end{cases} \\ &P_{m,eSC} = T_{eSC} \omega_{eSC} / \eta_{eSC} \end{aligned} \tag{18}$$

We use simplified internal resistance battery model (Rint model), and SOC variation by this model is as Equation (19). In the equation (19), I_{bat} is the output current of the battery, η_c is the Columbic efficiency, V is the voltage of the battery, R_{int} is the internal resistance of the battery, C_{bat} is the capacity of the battery, T_s is the unit step time, and x_1 represents SOC.

$$\begin{aligned} &P_m = T_m \omega_m + T_{eSC} \omega_{eSC} \\ &I_{bat} = \eta_c \frac{V - \sqrt{V^2 - 4R_{int}P_m}}{2R_{int}} \\ &\Delta SOC = -\frac{1}{C_{bat}} I_{bat} T_s \\ &x_1(k+1) = x_1(k) + \Delta SOC \end{aligned} \tag{19}$$

3.3.9. Cost

The cost is the fuel consumption of the engine. The Dpm.m function synthesizes the cost per step to determine the optimal path which has the minimum fuel consumption. Therefore, if we set the cost as shown in equation (20), the result of Dynamic programming will be the path with the lowest fuel consumption for the configuration.

$$J = \Delta m_{fuel} \tag{20}$$

4. DYNAMIC PROGRAMMING RESULT

4.1. Test scenario

In order to compare the operating characteristics of the motor and the electric supercharger for improving the fuel economy according to the operating point of the engine, the linear acceleration test of the engine is set as a test scenario. A graphical representation of the test scenario is shown in Figure 8, and the scenario is defined in the Main.m function.

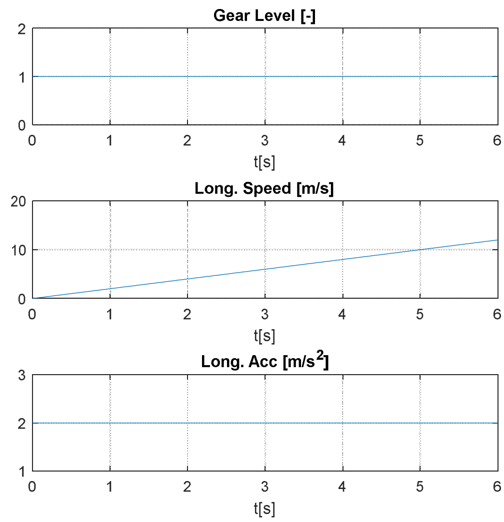


Figure 8: Test scenario

4.2. Simulation result

For the configuration of Figures 2~4, we simulate test scenario in Section 4.1. And the electric energy used, that is, the SOC variation amount is limited to be the same. The simulation results for the configuration with only the electric supercharger (Figure 2) are shown in Figure 9~11. As can be seen in Figure 10, when using only the electric supercharger, it can be seen that boosting the engine in the low-speed range helps improve fuel economy. However, as shown in Figure 11, when using only the electric supercharger, it shows a disadvantage that high desired torque is not satisfied.

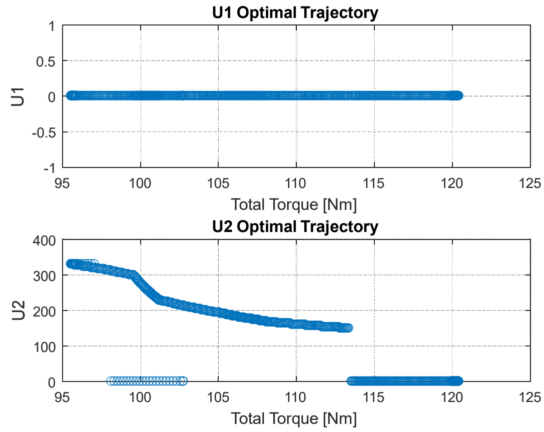


Figure 9: u_1, u_2 vs total torque (using electric supercharger only)

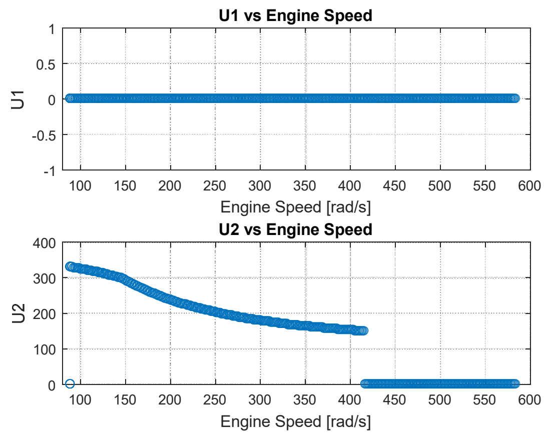


Figure 10: u_1, u_2 vs engine speed (using electric supercharger only)

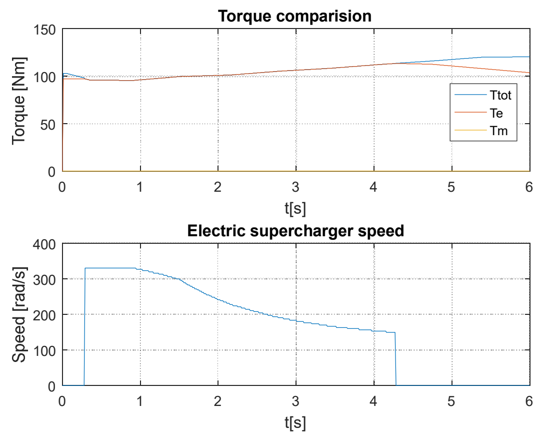


Figure 11: Torque and electric supercharger speed vs time (using electric supercharger only)

The following is the simulation result for the case where only the motor of P0 configuration is used (Figure 3). In this case, we can see that it does not operate in the low-speed range of the engine, unlike the case where only the electric supercharger is used, but operate to satisfy the high desired torque

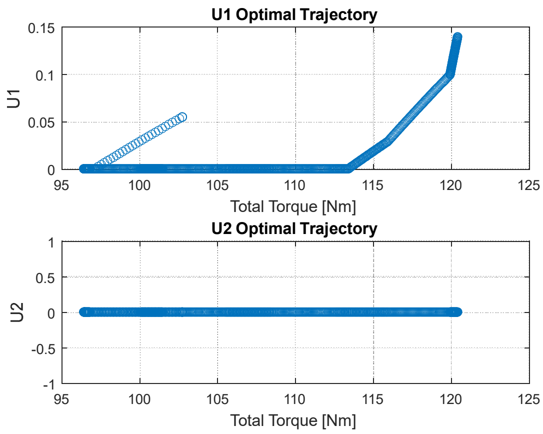


Figure 12: u_1, u_2 vs total torque (using P0 motor only)

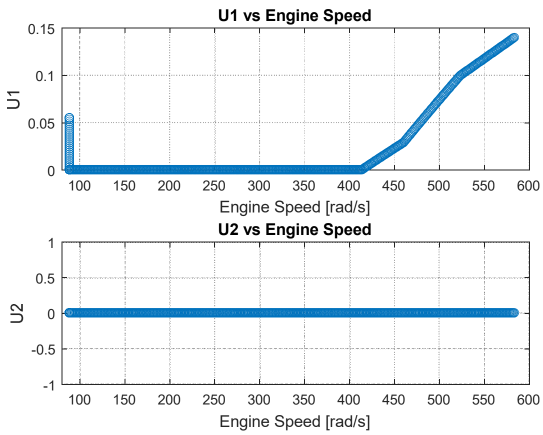


Figure 13: u_1, u_2 vs engine speed (using P0 motor only)

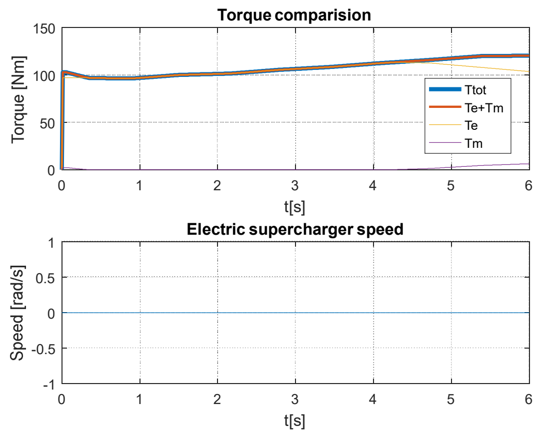


Figure 14: Torque and electric supercharger speed vs time (using P0 motor only)

The following is the simulation result when both P0 configuration motor and electric supercharger are used. In the case of a motor, it operates like that in the case of using only a motor, but in the case of an electric supercharger, it can be seen that it operates when the desired torque is high, not in engine low-speed range.

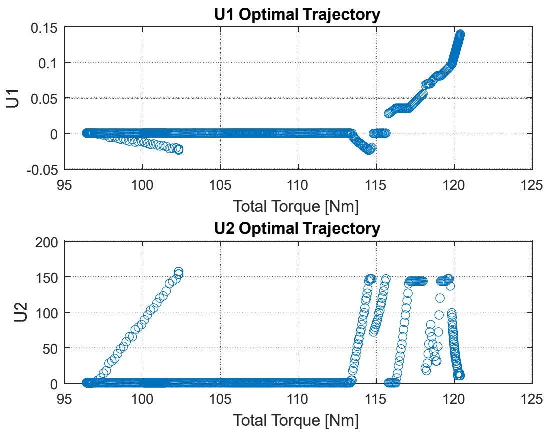


Figure 15: u_1, u_2 vs total torque (using P0 motor and electric supercharger)

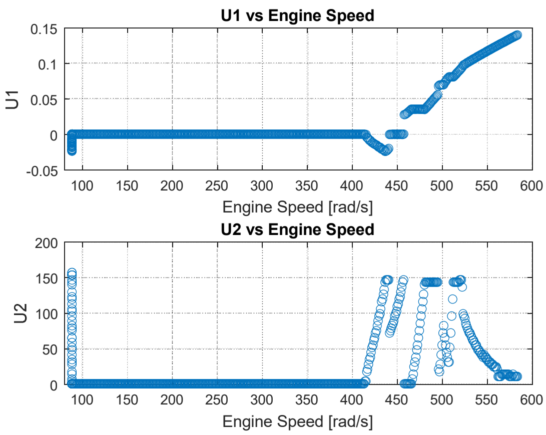


Figure 16: u_1, u_2 vs engine speed (using P0 motor and electric supercharger)

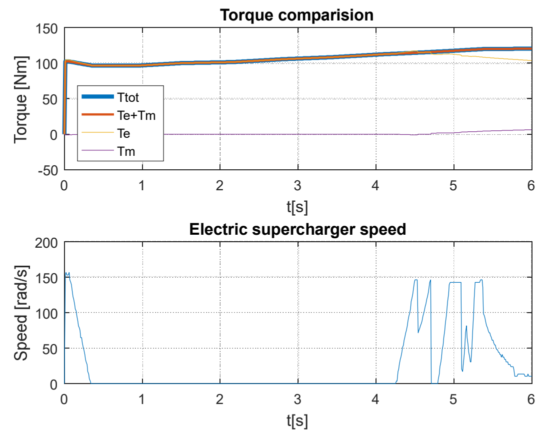


Figure 17: Torque and electric supercharger speed vs time (using P0 motor and electric supercharger)

However, simulation results for the case where the SOC variation amount is not limited are shown in Figure 18~20. As shown in the figure, when the speed of the engine is less than about 103 [rad/s], it can be seen that the motor assists the power of the engine, and thereafter it is assisted with the electric supercharger. And we can also see that the motor and the electric supercharger are used together when the desired torque is high.

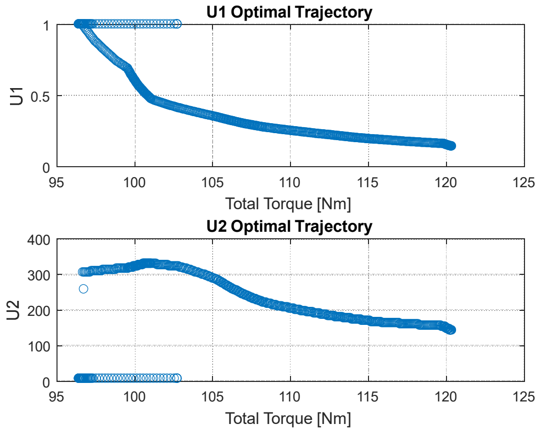


Figure 18: u_1, u_2 vs total torque (using P0 motor and electric supercharger)

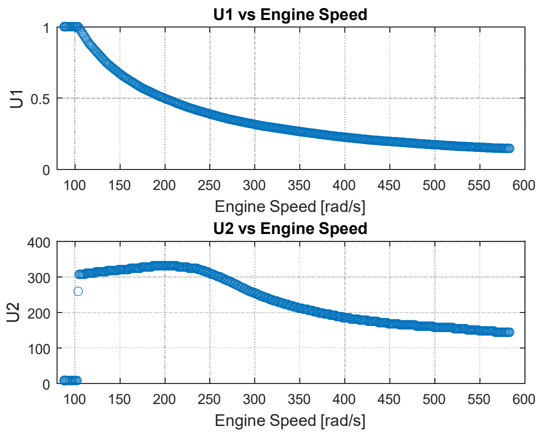


Figure 19: u_1, u_2 vs engine speed (using P0 motor and electric supercharger)

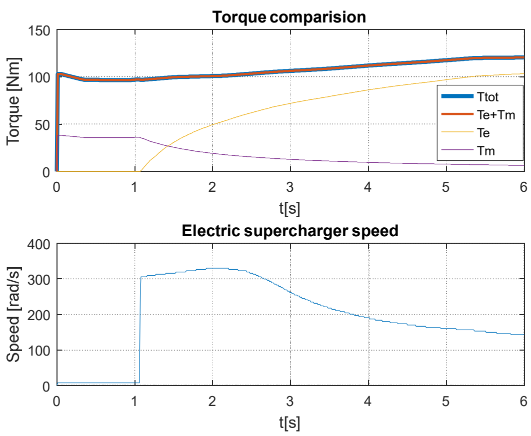


Figure 20: Torque and electric supercharger speed vs time (using P0 motor and electric supercharger)

The above simulation results are summarized as follows.

First, from the results of Figure 9~14, when the same amount of electric energy is used, we can see that it is advantageous for fuel economy that the electric supercharger operates at the low-speed region of the engine, and for the motor, we can see that it is advantageous that it operates in the region where desired torque is high. However, if the motor and the electric supercharger are used together, it is advantageous to use both devices in the region where the desired torque is high.

Second, from the results of Figure 18~20 where the battery SOC has margin, we can see that it is advantageous for fuel economy that the motor assists engine power where the engine speed is near idle speed. And where above that speed, it is advantageous that assist engine power with motor and electric supercharger together.

From these result, we can see that when P0 configuration motor and electric supercharger are used together, the electric supercharger should be controlled differently from the case of using electric supercharger only. And we can also see that using both devices has sufficient advantages in terms of fuel economy.

5. CONCLUSION

In this paper, we compare the operating characteristics of two devices, P0 configuration motor and electric supercharger, which can assist engine power of 48V mild hybrid system. We compare the operation characteristics for the case of using electric supercharger only, P0 configuration motor only, electric supercharger and P0 configuration motor together, through Dynamic programming for each case. And simulation is performed in Matlab environment. From the result of Dynamic programming, we can see that when P0 configuration motor and electric supercharger are used together, the electric supercharger should be controlled differently from the case of using electric supercharger only. And we can also see that using both devices has sufficient advantages in terms of fuel economy.

ACKNOWLEDGMENTS

This work was supported by the Industrial Strategic Technology Development Program (10076437, Development of hybrid drive topology exploration and control technology for fuel economy optimization of 48V mild HEVs) funded By the Ministry of Trade, Industry & Energy(MOTIE, Korea).

REFERENCES

Aaron Isenstadt, John German, Mihai Dorobantu, David Boggs, Tom Watson, 2016. Downsized, boosted gasoline engines, International Council on Clean Transportation, October 28.

Andreas Baumgardt, Dieter Gerling, 2015a. 48V Recuperation Storage Based on Supercaps for Automotive Application, EVS28 International Electric Vehicle Symposium and Exhibition, May 3-6, KINTEX (Korea).

Andreas Baumgardt, Dieter Gerling, 2015b. 48V Recuperation Storage Including Stabilizing 12V

Tap for HEVs, 2015 IEEE Transportation Electrification Conference and Expo (ITEC), June 14-17, Dearborn (Michigan, USA).
Anthony Rick, Brain Sisk, 2015. A Simulation Based Analysis of 12V and 48V Microhybrid Systems Across Vehicle Segments and Drive Cycle, SAE 2015 World Congress & Exhibition, April 21-23. Detroit (Michigan, USA).
Bo Hu, Chunlin Chen, Zhangsong Zhan, Xueuing Su, Tiegan Hu, Guangyong Zheng and Zhiyong Yang, 2017. Progress and recent trends in 48V hybridisation an e-boosting technology on passenger vehicle – a review, Proc IMechE PartD: J Autonobile Engineering 00(0), 1-19.
Bryn Richards, Tran Hoang Hiep, Oliver porter, Mark Parrett, Petr Matajsz, 2015. A 110kW/L, 200Nm/L Downsized Gasoline Engine With a High-speed, Permanent Magnet Electric Supercharger. Available from: http://www.aeristech.co.uk/wp-content/uploads/3_Tran_A5_2.pdf [accessed 8 July 2018]
Dr. Ing. Olivier COPPIN, 2016. Valeo Presentation – From 12 plus 12V to 48V. Available from: https://48v-vehicles.iqpc.de/downloads/valeo-presentation-from-12-plus-12v-to-48v [accessed 6 July 2018]
DR.-ING. Richard Aymanns, DR.-ING. Tolga Uhlmann, DIPL.-ing. Carolina Nebbia, DIPL.-ing. Thorsten Plum, 2014. Electric Supercharging New Opportunities with Higher System Voltage, Available from: https://link.springer.com/content/pdf/10.1007/s38313-014-0166-0.pdf [accessed 6 July 2018]
Hiep Hoan Tran, Bryn Richard and Kevin Gray, 2016. Developing a Performance Specification for an Electric Supercharger to Satisfy a Range of Downsized Gasoline Engine Application, SAE 2016 World Congress and Exhibition, April 12-14, Detroit (Michigan, USA).
Ivan Filidoro, 2015, Numerical Investigation of Steady-State and Transient Operations of a Turbocharged SI Engine with Electric Supercharger. Thesis (Master). Windsor University.
Junyong Park, Taeho Park, Hyeongcheol Lee, 2017. An optimization of power distribution considering supercharging characteristic for electric supercharger applied 48V mild hybrid vehicle, KSAE 2017 Annual Autumn Conference & Exhibition, pp.1328-1333. November 15-18, Yeosu (Jeollanam-do, South Korea).
Malte Kuypers, 2014. Application of 48 Volt for Mild Hybrid Vehicles and High Power Load, SAE 2014 World Congress & Exhibition, April 1, Detroit (Michigan, USA).
Mark Schudeleit, Christian Sieg, Ferit Kücükay, 2015. The Potential of 48V HEV in Real Driving, International Scholarly and Scientific Research & Innovation Vol: 9 No: 10, 1719-1728.
Olle Sundstrom, Lino Guzzella, 2009. A Generic Dynamic Programming Matlab Function, 18th IEEE International Conference on Control Applications Part of 2009 IEEE Multi-conference on Systems and Control Saint Petersburg, July 8-10, St. Petersburg (Russia).
Prasad Sajjan Divekar, Beshah Ayalew, and Robert Prucka, 2010. Coordinated Electric Supercharging and Turbo-Generation for a Diesel Engine, SAE 2010 World Congress & Exhibition, April 13-15.
Ran Bao, Victor Avila, James Baxter, 2017. Effect of 48V Mild Hybrid System Layout on Powertrain System Efficiency and Its Potential of Fuel Economy Improvement, WCX™ 17: SAE World Congress Experience, April 4-6, Detroit (Michigan, USA).
Ricardo Martinez-Botas, Apostolos Pesiridis, Yang Ming Yang, 2011. Overview of boosting options for future downsized engines, SCIENCE CHINA Technological Sciences, Vol. 54 No.2, 318-331.
Stephan Tavernier and Samuel Equoy, 2013. Design and Characterization of an E-booster Driven by an High Speed Brushless DC Motor, SAE International, April 8, 360-365.
Thomas Eckenfels, Florian Kolb, Steffen Lehmann, Waldemar Neugebauer, Manel Calero, 2018. 48V Hybridization, A Smart Upgrade for the Powertrain. Available from: http://schaeffler-events.com/symposium/lecture/h3/index.html [accessed 6 July 2018]
Zifan Liu, Andrej Ivanco, and Zoran S. Filipi, 2016. Impact of Real-World Driving and Driver Aggressiveness on Fuel Consumption of 48V Mild Hybrid Vehicle, SAE International, April 5, 249-258.

LEVERAGING VIRTUAL ENVIRONMENTS TO TRAIN A DEEP LEARNING ALGORITHM

Dean Reed[a], Troyle Thomas[b], Latika Eifert[c], Shane Reynolds[d], Jonathan Hurter[e], Frank Tucker[f]

[a],[b],[d],[e] Institute for Simulation and Training, University of Central Florida
[c],[f] Army Research Laboratory-Human Research and Engineering Directorate Advanced Training and Simulation Division, SFC Paul Ray Smith Center

[a]dreed@ist.ucf.edu, [b]tthomas@ist.ucf.edu, [c]latika.k.eifert.civ@mail.mil
[d]sreynold@ist.ucf.edu, [e]jhurter@ist.ucf.edu, [f]frank.m.tucker4.civ@mail.mil

ABSTRACT
Open source datasets are the typical source used to train computers to accurately detect visual objects (e.g., humans, animals, and inanimate objects) through various machine learning methods (e.g., Deep Learning (DL)). This data, however, is not feasible for use in the military domain. In this paper, a comparative analysis of real and virtual training data is provided, using the You Only Look Once (YOLO) Convolutional Neural Network (CNN) model. The main concern of this paper is to verify the process and accuracy of using a domain-specific U.S. Army Virtual Environment (VE), in contrast to a Real Environment (RE) dataset, with DL. Comparative results suggest that substituting a VE to provide training data for the DL model saves manual labour while maintaining a quality precision-recall curve.

Keywords: machine learning, deep learning, virtual environments, photogrammetry

1. INTRODUCTION

Deep Learning (DL) is a machine learning method with a focus on deep neural networks, or Convolutional Neural Networks (CNNs). With the advent of parallel processing hardware and thousands of GPU cores, it is feasible to train computers to accurately detect visual objects (e.g., humans, animals, and inanimate objects) in a real-time video feed. The capability of CNNs represent a significant area of interest to the military; this is because, given adequate training data, CNNs can learn features from any object and then identify it in new positions and environments, solving a computer's object detection problem. The U.S. Army Research Laboratory (2017) has cited machine learning as one essential research area for the future. On a similar note, computer vision and sensing, developed through machine learning, is key to the future of military robotic applications; not to mention the possibility of a tool able to analyze streaming video for specific objects (Lee, Valisetty, Breuer, Kirk, Panneton, and Brown 2018).

Yet, the difficulty of acquiring real-world training (i.e., visual object) data for the supervised DL object detection problem is a major obstacle to DL's applicability to the military domain. The difficulty of the data acquisition is due to a conjunction of the labour intensive tasks of domain design, vast amounts of photography, and image labelling. Ultimately, when custom training data is necessary, real-world objects of interest are photographed from many angles, annotated, localised, and manually provided bounding boxes, to build a ground-truth training set. Virtual Environments (VEs) can ideally provide, and extend the quantity of, training data; VEs also have the professed benefits of optimisation through different efficiency factors: automated control, increased flexibility, and ultimately a significant decrease in both labour cost and acquisition time. Although the efficiency of VEs in DL seems clear, the question that arises centres on effectiveness: "What is the difference of precision in DL, in terms of comparing virtual training data to real-world data?"

While vast public datasets exist for common sets of objects in the public sector (e.g., ImageNet), a shortcoming exists in training data for objects in the military domain. A common goal for many areas of DL is to leverage VEs for training, yet domain adaptation is a challenge (Atapour-Abarghouei and Breckon 2018). The present paper discusses overcoming the issue of this arduous process through comparing a Real Environment (RE) training method to a novel methodology employing a highly-accurate modelled VE (as the basis to simulate the ground-truth). The model used to train and detect objects used herein is the You Only Look Once (YOLO) CNN architecture (Redmon, Divvala, Girshick, and Farhadi 2016). This paper seeks to contribute to knowledge in line with research on overcoming a domain gap (i.e., there has been recent investigation on how well the virtual domain can transfer to the real domain in machine learning; e.g., see Rajpura, Bojinov, and Hegde 2017, Gaidon, Wang, Cabon, and Vig 2016). The alternative VE method has other direct benefits to the military, which are discussed in section three.

2. RELATED WORKS

The idea of augmenting or replacing real training data with computer-generated data is currently being explored with the advent of DL and the need for large quantities of novel data. Several other works investigate this with applications in a wide variety of computer vision problems. Endeavours into computer-generated training data have ranged from a simple model (Broggi, Fascioli, Grisleri, Graf, and Meinecke 2005) to an entire world (Gaidon, Wang, Cabon, and Vig 2016).

When creating a model and learning from a training domain distribution, an assumption is made that test data will be from the same or similar distribution (Gretten, Smola, Huang, Schmittfull, Borgwardt, and Scholkopf 2009). We do not need to adhere to the same distribution when training using virtual, computer-generated data. Such a domain shift can potentially skew the intended real-world performance if the shift is significant. This means that if the used training dataset is different enough from the testing set, the training model may not be able to perform competently. To use a VE effectively and limit skew, we must understand and identify the contributing factors.

Atapour-Abarghouei and Breckon (2018) attempted to avoid the domain shift issue by proposing a method to reduce the distance between the training and testing set. In their work, they adapt the real world and synthetic data to a shared intermediate domain with the goal of reducing such a discrepancy. This approach avoids the issue of artefacts within synthetic data, but seems to be at the cost of detail. The latter could provide additional features for a model to learn. The following work separates the real and virtual domains and attempts to reduce the distance through the use of high-quality virtual environments and models.

Gaidon, Wang, Cabon and Vig (2016) in particular explored the performance gap of a photorealistic VE, and its real-world counterpart, for a multi-object tracking problem. Additional parameters, such as weather and lighting conditions, were added and subsequently identified for their impact on tracking. In comparison to our work, this work is similar in that it investigates the topic of the performance gap between RE and VE models. In the following research study, we differ by exploring the performance gap of solely real and virtual models (as opposed to a pre-training scheme, which the aforementioned authors employ) and consider the tangible cost incurred by the creation of both dataset in terms of man-hours, cost estimates, and applicability to the military domain.

3. IMAGE DATA COLLECTION PROCESS

The object image data collection process comprised a group of subtasks: turntable, static, and dynamic subtasks. The turntable subtask involved collecting photos at 15-degree increments on a turntable. The static subtask involved collecting photos at precise locations within an environment. The dynamic subtask, which is unique to the VE, involved collecting photos at any position in the environment.

After the photo collection step was complete, the object within each image was annotated, localised, and given a bounding box. To achieve localisation, a masking technique was implemented to handle the majority of the RE photos. Localisation was done by taking an initial 'blanking' photo of an unaltered scene, without the object, at a given camera position. Afterwards, the object was placed in the scene, and the photo-taking process continued. The blanking shot was used to provide a mask of the object within the scene, due to the object being the only change. Then, the bounding box position was generated based on the generated mask. This subtask process was done for the majority of the turntable subtasks; any images that required further corrections were hand labelled.

Although the present effort focused on one class of objects, consider the labour involved with creating domain-specific datasets in general: in previous work, over 18,500 high-resolution (5184 x 3456) photographs for RE-based image data training material were taken by the researchers for four classes of objects. These classes included military helmets, M4-Rifle magazines, coffee cups, and common computer monitors. It is estimated that a minimum of 200 hours were spent in repetitive data collection tasks to obtain the content for these small objects.

The process employed to construct a large-scale database suitable for the YOLO model can be summarised in a few steps. Step 1 is to define the set of target object categories. Step 2 is to collect a diverse set of images to represent the object categories in Step 1. Step 3 is to annotate potentially millions of images collected in Step 2. The U.S. military domain is a challenging area to conduct a traditional RE DL implementation for a variety of reasons. Namely, military doctrine limits access of information to most of its modern physical assets. However, even upon acquiring access to the asset(s), and obtaining the volume of photos required for high-quality training in Step 2, the manual labour in annotating the photos is significant. For example, acquiring photos of high-end weapon systems, at multiple angles and at a variety of daylight conditions, would be a significant undertaking if performed manually. Also, the sheer volume of multinational military assets that could be identified in Step 1 is immense. Further, due to the intrinsic and secretive nature of military organisations, using clever crowd sourcing techniques (e.g., Deng et al. 2014) is not practical.

For our tests, we selected a single class of detection objects to satisfy Step 1: a set of military helmets. For inter-class comparison, we used a futuristic helmet developed by the Natick Soldier Center, and we used a common U.S. Army standard issue helmet. A set of 4000 staged photographs were taken, and over 62 hours of labour were accumulated for this one class.

3.1. VE Creation

The Unity game engine was selected as the rendering pipeline for the VE. Unity provides a robust Integrated Development Environment (IDE) and Applications

Programmer Interface (API) that allows for agile development and rapid prototyping. Unity supports importing numerous native file formats (.fbx, .max, .psd, etc.), allowing the results of changes to the original file to appear instantly in Unity. Also, it supports multiple scripting languages to interface with game objects, contains advanced shaders, and supports light probes; all of these aspects were applied in this study.
Scripts were created for generating images for each subtask. The images generated were taken with a one-to-one correspondence to the RE dataset creation. A bounding box is then generated on the produced image because the position of the object in space is known. For the dynamic subtask, the virtual helmet (Figure 1) was allowed to move anywhere in the virtual room. The camera was then placed facing the helmet, enabling a check to see if the helmet was in the field of view, by way of a ray cast.

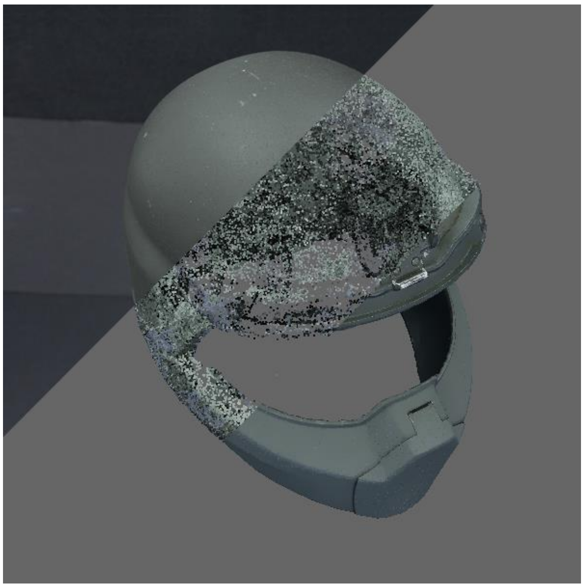


Figure 1*:* Advanced Helmet Imagery: RE photo (left) Point Cloud (centre), and VE model (right)

3.1.1. VE Creation Process

Photogrammetry techniques were utilised and combined with data from a portable Light Detection and Ranging (LIDAR) laser scanner to create a VE that replicated the RE as closely as possible. Because the RE consisted of objects (e.g., walls, desks, and monitors) that have monochromatic, flat, and featureless surfaces, photogrammetry alone could not produce the desired results. The textures were accurately created from the photographs. The caveat stems from the underlying model, which was created from the generated point cloud, and did not yield true-to-life results. This inaccuracy is due to photogrammetry's reliance on surface features to accurately reconstruct said surfaces. Some objects do not contain enough features to process accurate point data (e.g., the ceiling tiles of the room produced acceptable results because the pattern on the tiles contained many features for the reconstruction of the surface). Thus, to resolve inherit limits of photogrammetry, photographs were imported into RealityCapture (i.e., photogrammetry software), and then aligned and processed as a point cloud. Various scans produced from the laser scanner were then imported. Once both sets of point clouds were created, the two sets were aligned to produce a high-resolution mesh with correlated textures. This new mesh accurately matched the spatial dimension and texture of the RE.
Another method used to further improve the realism of the VE was creating a High Dynamic Range Imaging (HDRI) probe. The HDRI probe allowed accurate reflection and lighting data in the VE, based on the RE. A panoramic adapter on a tripod and a full-frame digital camera with an 8mm fisheye lens were used to capture images at set angles. Additionally, photographs were taken with multiple exposure angles, and saved in the RAW file format. These images were processed in PTGui (i.e., panoramic image stitching software) to create the 360-degree spherical HDRI file used in the game engine for reflections and lighting.

3.1.2. Virtual Detection Object Creation

To create the detection objects used in the training of the RE (i.e., two military helmets), corresponding realistic virtual objects were required. Photogrammetry techniques were used to generate realistic models of the two military helmets. Both helmets presented a challenge when using photogrammetry because of their lack of features. Access to a hand-held laser scanner was limited, and so a talcum powder technique was used that allowed one to introduce nonlinear features to the helmets for photogrammetry. The talcum powder technique involved lightly and uniformly covering the helmets with talcum powder in a way that created surface features. Random features were introduced to the native green solid texture by sprinkling the powder over the helmet. Using this simple method to fragment the continuous surfaces helped derive a high-quality and accurate point cloud. Each helmet was put on a tripod and set on the turntable. Three cameras were configured to take photos from different angles, which allowed simultaneous capture of the top, middle, and bottom of the helmet. This method also allowed precise angles from the camera to the helmet. Photos were captured from the same angles for each process (i.e., powder and no-powder), since the turntable had markings every 10 degrees. The reason to have the helmet in an identical position in both sets of images (i.e., powder and no powder) was to aid in the alignment of the photographs used for texturing the model and to ensure surface reconstruction. Once a point cloud was created that yielded accurate surface detail, the photographs of the helmet with the powder were swapped with the powderless helmet's UV-texture images captured in earlier steps. This procedure resulted in helmet models with excellent surface quality and realistic textures.

4. TRAINING PROCEDURE

To compare the VE and RE, the dataset was segmented into subsets. Distinct subsets were created, based upon

objects, static training, and dynamic training as described in section three. This segmentation was done to isolate collections of images, where a significant disparity between the performance of the RE and the VE sets may exist. This segmentation also allows the ability to identify characteristics of the subset that may have led to given observations.
After the collection of the VE and RE datasets, each training set was segmented into an 80:20 training-testing split. The training sets comprise all the subsets shown in Table 1. This includes the TR-ALL subset, which encompasses every other subset. A training set is denoted as TR-<Set Name>. The configurations and metadata files were structured as needed for the YOLO model. Using pre-trained ImageNet weights, training was executed for 50000 epochs for each training subset. Afterwards, both the VE- and RE-trained models were tested on the experimental validation set.

5. EXPERIMENTS AND RESULTS

To provide a comparative analysis between the VE and RE performance, both datasets were segmented into subsets and trained. The VE and RE subsets were then validated on a set of RE images that were not included in the training process. A comparison can then be made between each subset via performance measures. The performance results are given in terms of average precisions, the percentage difference between the VE and RE average precisions, and a precision-recall curve based upon the Pascal Visual Object Classes (VOC) evaluation metrics (Everingham, Eslami, Van Gool, Williams, Winn, and Zisserman 2015). The results for the TR-ALL subset, as they show a high-level view of our findings, are given in Table 2, Figure 2, and Figure 3. The variations on training time per epoch are shown in Table 3.

Table 1: Experimental Subsets

Subsets	Explanation
TR-ALL	Contains entire dataset. Validated on an ALL RE validation set (Table 2; Figures 2 and 3).
TR-H1	Contains only images of helmet 1. Validated on an H1 RE validation set.
TR-H2	Contains only images of helmet 2. Validated on an H2 RE validation set.
TR-Static	Contains only images from the static training. Validated on an ALL and Static RE validation set.
TR-Not Static	Contains images that do not include the static training. Validated on an ALL and Not Static RE validation set.
TR-Dynamic	Contains only images from the dynamic training. Validated on an ALL and Dynamic RE validation set. (Unique to VE)
TR-Not Dynamic	Contains images that do not include the dynamic training. Validated on an ALL and Not Dynamic RE validation set. (Unique to VE)

Table 2: Average Precisions of the Model Trained on the Entire Dataset

TR-ALL			
Epochs	**AP_{ve}**	**AP_{re}**	**% Difference**
400	0.0006	0	200
800	0.5788	0.916	45.1242705
10000	0.9563	1	4.463453161
20000	0.9524	1	4.879994727
30000	0.9376	1	6.440062062
50000	0.9519	1	4.932103213

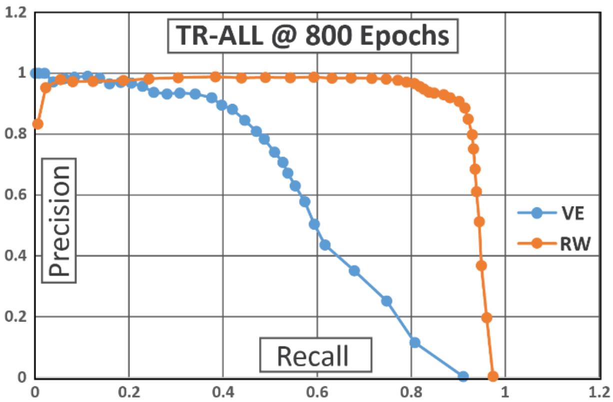


Figure 2: Entire VE vs RE Dataset After 800 Epochs of Training

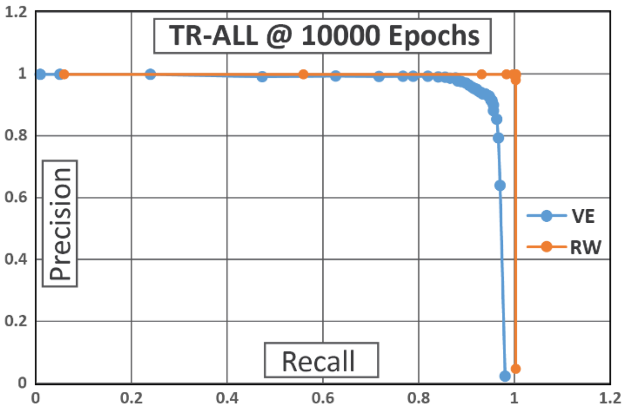


Figure 3: Entire VE vs RE Dataset After 10000 Epochs of Training

6. DISCUSSION

The results show that the RE model rapidly becomes a near-perfect classifier on the given RE validation set. The VE model lags behind the RE model initially, but reaches similar performance after 10000 epochs of training. Since the validation set is composed of RE images, the difference between the RE and VE models' training time to reach convergence is expected.

In summary, the data shows how, as the YOLO model reaches 10000 epochs, a percentage difference of less than five per cent is achieved between the VE and RE models. This is evidenced in the performance convergence seen in Figures 2 and 3 for TR-ALL, with an implication that the 40000 and 50000 epoch tests are wasted computer time for the dataset. The findings show that other training subsets perform similarly in that a less-than-five per cent difference occurs after 10000 epochs. The small performance gap between the RE and VE datasets can be mitigated further by exploring increased VE fidelity. Overall, the results support the VE approach in the absence of RE data.

Given that the present research incorporated relatively small models (i.e., helmets), a precise estimate of cost savings, across the broad spectrum of the domain of interest for the U.S. Army, cannot be made. However, a coarse savings statement can be made. In our experiment, 200 man-hours were spent on collecting RE data for a single object. Using the General Services Administration (GSA; System Studies & Simulation, Inc 2015) schedule for a single object documented, one sees that a Junior Engineer rate of $35 per hour with a conservative 45 per cent overhead leads to a per-instance cost of $10,150 for each small-to-medium RE object. Additionally, using a purely traditional 2D photograph source for training requires a match between the expected environmental conditions of detection and the photographic environments. Considering how a CNN needs to recognise objects in different conditions, such as low-light environments and environments with extreme conditions (e.g., deserts and complete snow cover), sufficient training imagery of the objects in those environments is needed to maximize recognition. Given a VE (e.g., Virtual Battlespace 3; VBS3) that contains the backdrops and 3D models of interest, data collection for the VE training material becomes a substantially automated and labor-free task; this efficiency is compounded with automation scripts that can generate datasets for thousands of distinct military objects. Ultimately, the cost of performing 2D photography in every desired environment becomes an expected, increased cost factor. A simplified cost quantification includes the number of real-world objects, multiplied by the instance cost, multiplied by the number of real-world environments. For example, to recognise 50 real-world objects, in five environments, a cost of $2,537,500 for generating the training data is found.

Noting the limits of 2D data acquisition, the analogous training data from a high-quality VE can lead to reduced cost. This reduction is based on the assumption that many target objects are included with systems like VBS3. Besides VBS3, the U.S. Army Games for Training (Baker 2018) program manages a repository that houses hundreds of unique terrains for relevant environmental backdrops, and houses over 500 freely available military objects that can be added to the already extensive library of VBS3 models. Here, VE initial costs rest on generating any 3D objects not provided by the VE, and automation scripts.

7. FUTURE WORK

7.1. Night Vision Image Generator (NVIG) Application

As a novel application area, the U.S. Army Research, Development, and Engineering Command (RDECOM), Communications-Electronics Research, Development, and Engineering Center (CERDEC) Night Vision and Electronic Sensors Directorate (NVESD) has developed a real-time, multispectral 3D image generator. The use of real imagery along with various innovative techniques and processes has helped to create a library of hundreds of vehicle models in both the visible and InfraRed (IR) spectrum. This software is currently being used extensively throughout U.S. Army Battlelabs for man-in-the-loop wargames. One future direction includes using the NVIG models as a baseline training set for DL. Based on the findings in the present effort, an assertion is made where NVIG will enable recognition of military domain objects as acquired through advanced sensors.

7.2. Experimental Considerations

To fully grasp the VE approach's benefits, future tests should identify impactful factors that contribute to the performance differences of a VE and RE. Research can explore lighting effects, model-polygonal density, ambient visual representation, and lens distortions. Further, although the present investigation considered the real and virtual as independent datasets, perhaps using real data to help boost the efficacy of a predominately virtual approach could maximise potential. The premise of this paper asserts limited access to military assets, but a next step includes using a hybrid concept as an optimisation when RE datasets for training are obtainable. This idea of a hybrid approach is mentioned in Rajpura, Bojinov, and Hegde (2017).

Table 3: Training Times for the RE and VE Subsets

Training Time (Hrs)		**Epochs**						
		400	**800**	**10000**	**20000**	**30000**	**40000**	**50000**
RE Subset	**TR-ALL**	0:26	0:52	9:09	17:39	26:13	34:48	43:11
	TR-H1	0:20	0:39	8:25	16:49	25:08	33:28	41:45
	TR-H2	0:19	0:38	8:22	17:34	25:54	34:12	42:42
	TR-Static	0:21	0:42	8:33	17:08	25:55	34:46	43:49
	TR-Not_Static	0:20	0:39	9:01	18:51	28:39	38:17	48:00

VE Subset	**TR-ALL**	0:27	0:53	10:29	20:46	31:05	41:21	51:42
	TR-H1	0:27	0:53	11:00	22:13	32:51	44:21	54:57
	TR-H2	0:25	0:50	10:18	20:36	31:21	42:12	53:12
	TR-Static	2:23	4:18	50:45	100:43	150:09	207:14	257:29
	TR-Not_Static	1:58	3:59	52:59	106:31	155:32	206:46	256:13
	TR-Dynamic	0:25	0:49	10:13	20:30	30:48	41:07	51:39
	TR-Not_Dynamic	0:26	0:50	10:08	20:28	30:48	41:10	52:13

8. CONCLUSION

This paper explored the creation, performance and potential impact of using a VE-generated training set for a DL object detection problem. We show the advantages and difficulties for creating both datasets. The work identifies the feasibility of using a VE as a cost-saving measure and quantifies the gap in performance using a systematic workflow.

REFERENCES

Atapour-Abarghouei A. and Breckon T.P., 2018. Real-time monocular depth estimation using synthetic data with domain adaptation. Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 18-22, Salt Lake City (Utah, USA).

Baker T., 2018. VBS3 Resource. Milgaming Website. Available from: https://milgaming.army.mil/VBS3/files/ResourceList.aspx [accessed 11 July 2018].

Broggi A., Fascioli A., Grisleri P., Graf T., and Meinecke M., 2005. Model-based validation approaches and matching techniques for automotive vision-based pedestrian detection. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops. September 21-23, San Diego (California, USA).

Deng J., Russakovsky O., Krause J., Bernstein M.S., Berg A., and Fei-Fei L., 2014. Scalable multi-label annotation. Proceedings of CHI 2014, One of a CHInd, pp 3099-3102. April 26 – May 1, Toronto (Ontario, Canada).

Everingham M., Eslami S.A., Van Gool L., Williams C.K., Winn J., and Zisserman A., 2015. The pascal visual object classes challenge: A retrospective. International Journal of Computer Vision, 111 (1), 98-136.

Gaidon A., Wang Q., Cabon Y., and Vig E., 2016. Virtual worlds as proxy for multi-object tracking analysis. arXiv. Available from https://arxiv.org/abs/1605.06457 [accessed 4 May 2018].

Gretton A., Smola A., Huang J., Schmittfull M., Borgwardt K., and Schölkopf B., 2009. Covariate shift in kernel mean matching. In: J Quiñonero-Candela., M. Sugiyama, A. Schwaighofer, and N.D Lawrence., eds. Dataset shift in machine learning. Cambridge, MA: MIT, 131-160.

Lee M., Valisetty R., Breuer A., Kirk K., Panneton B., and Brown S., 2018. Current and future applications of machine learning for the US Army. Aberdeen Proving Ground, MD: US Army Research Laboratory.

Rajpura P.S., Hegde R.S., and Bojinov H., 2017. Object detection using deep CNNs trained on synthetic images. arXiv. Available from: https://arxiv.org/abs/1706.06782 [accessed 4 May 2018].

Redmon J., Divvala S., Girshick R., and Farhadi A., 2016.You Only Look Once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779-788. June 27-30, Las Vegas, (Nevada, USA).

System Studies & Simulation, Inc, 2015. GSA Schedule. GSA Advantage! Available from https://www.gsaadvantage.gov/ref_text/GS00F0037P/0P5FUQ.38UNOH_GS-00F-0037P_PSSREV20151222MODS003600380039.PDF [accessed 11 July 2018].

U.S. Army Research Laboratory, 2017. U.S. Army Research Laboratory Essential Research Areas. United States Army Research Laboratory. Available from: https://www.arl.army.mil/www/default.cfm?page=2401 [accessed June 21 2018]

AUTHORS BIOGRAPHY

Dean Reed (B.S. in Computer Science, UCF 2000) is a Senior Associate for Simulation with the Institute for Simulation and Training (IST) of the University of Central Florida. Mr Reed is a veteran of the U.S. Army, and leads a team of developers at IST. He has worked on a vast array of projects under the auspices of the University, including NASA Vision Spaceport. He is currently managing team efforts directed at evolving future training ranges on behalf of the U.S. Army.

Troyle Thomas is an M.S. student in Computer Science at the University of Central Florida. His research interests includes computer vision, natural language processing and machine learning applications, with particular attention to unsupervised learning methods. Mr Thomas has been a Graduate Research Assistant at the Institute for Simulation and Training since 2017. His primary responsibilities involve this paper, the research of virtual environments, and their application to deep learning training.

Latika (Bonnie) Eifert (M.S. in Computer Engineering, UCF 2003) is a Science and Technology Manager at the U.S. Army Research, Development and Engineering

Command (RDECOM), Army Research Laboratory, Human Research Engineering Directorate, Simulation and Training Technology Center (ARL-HRED ATSD) located in Orlando, Florida. Ms Eifert manages several projects associated with simulation and training. She is also supporting the Defense Advanced Research Project Agency (DARPA) by managing research program efforts.

Shane Reynolds is a graduate of UCF in Digital Media. He has specialized in compelling 3D content development and mobile game engine development for over 10 years. Mr Reynolds is a veteran of the U.S. Air Force. Shane is a Research Associate at the Institute for Simulation and Training where he has been a faculty member since 2008. His primary activities involve research and integration of modern technologies to train dismounted Soldiers at the squad level. Currently, his focal areas are technologies involving virtual reality, augmented reality, and photogrammetry.

Jonathan Hurter is a Research Technician at the Institute for Simulation & Training (IST), the latter a part of the University of Central Florida (UCF). Jonathan earned his Master's degree in Modeling & Simulation, as well as a Certificate in Instructional Design for Simulations, from UCF. His background covers digital media, instructional design, and technical writing. He has contributed to multiple simulation projects, including simulation-based training, avatar science, and malodor studies.

Frank Tucker is a Research Program Manager at the U.S. Army Research Laboratory in Orlando, Florida, managing research projects not only internally but also for the Defense Advanced Projects Agency (DARPA) in a variety of technologies, which include semiconductor lasers, quantum-based sensors, computer vision, and inertial sensors. Mr. Tucker has a BSc. in Electrical Engineering from the New Mexico Institute of Technology and has eighteen years of experience leading and managing several research efforts.

NEW APPROACH TO PMSM PARAMETERS IDENTIFICATION

Jean Claude Rakotoarisoa[(a)], Gino Hadjee[(b)], Jean N. Razafinjaka[(c)]

[(a),(c)] Electrical Engineering Department, Automation Laboratory, University of Antsiranana, Madagascar.
[(b)] Electrical Engineering Department, Electricity Laboratory, University of Antsiranana, Madagascar.

[(a)]rajeanclaude@gmail.com, [(b)]ginhadj@hotmail.com, [(c)]razafinjaka@yahoo.fr

ABSTRACT
This paper aims to the identification of a Permanent Magnet Synchronous Machine parameters by using Diffedge approach. In this case, the identification problem is considered as an optimization one. Parameters sensibility, the exact calculation of gradient and Jacobean in respect with the parameters are done by using symbolic derivative expressed by diagram. Two optimization process are here proposed: the first method uses the formal Jacobean found by finite differences and the second one uses the formal Jacobean calculated by Diffedge. The proposal validation consists of the convergence comparison and obtained results. It is here highlighted that the Diffedge approach leads to better results as convergence, robustness

Keywords: PMSM, optimization, Diffedge, Jacobean, identification, formal derivation.

1. INTRODUCTION

Permanent Magnet Synchronous Machine (PMSM) have now become increasingly used in various fields of use because of their superior performance over other types of machines. Among the many types of commands proposed to control the PMSM in torque and speed, the vector control is the most efficient and especially to control the behavior of the engine in transient regime.
However, effective vector control requires a good knowledge of the electrical and mechanical parameters of the engine, hence the importance of carrying out identification techniques to estimate the value of these parameters. For optimization, control, identification, calibration of parameters, monitoring, diagnosis, parametric sensitivity analysis, sizing of a process, exact calculation of the gradient, Jacobian with respect to parameters, is often essential. It increases the speed and accuracy of the convergence to the optimum when the all-digital fails.
For dynamic systems with nonlinearities such as PMSM, using finite differences, the gradient calculation, Jacobian, Hessian requires the judicious choice of no derivation. In many cases, the numerical derivative is unstable and inaccurate because of numerical approximation problems. In these situations, it is better to rely on Diffedge, which is an exact differentiator of models described as a Matlab / Simulink block diagram. The calculation of the derivative, with respect to the parameters, is carried out by the application of the rules JM (Gilbert R.J.C, Le Vey G. and John M. 1991). This corresponds to the need of the user to stay in the same working environment instead of converting his model to C, Fortran, etc.
In this work it is proposed to apply an optimization technique using this derivation method to identify the electrical and mechanical parameters of a PMSM used in a wind energy conversion system.
The identification of electrical parameters is done assuming that the speed of the machine is constant. That of the mechanical parameters is in transient mode of the speed.

2. DIFFEDGE METHODOLOGY

Diffedge is a differentiation tool designed for systems represented by block diagrams in Matlab Simulink. Its goal is to provide a result that remains in the same diagram environment. It was developped by Appedge in 2005 (John M. and Cambois T. 2004; Bastogne T., Thomassin M. and John M. 2007) and is the result of a long familiarity with the suject (Gilbert R.J.C, Le Vey G. and John M. 1991). Optimization often require to get an analytical expression of the gradient function, which is a non trivial task in this setting, but then real time optimization tools may be tested and translated in C/C++ code ready for use on embedded processors, using e.g. Embedded Coder®. Several ways may be considered, but they do not all allow to handle models with discontinuities (switch, saturation, etc.) or heterogenous mathematical representation mixing continuous and discrete time models. We will illustrate the possibility on some simple examples.

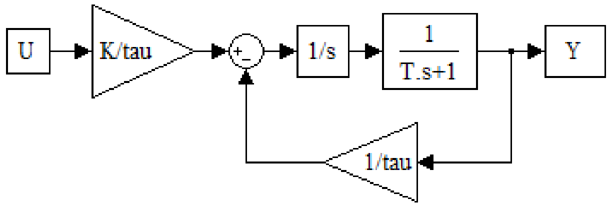


Figure 1: Academic example, second order transfer function

2.1. Graphic differentiation methodology

In the diagram of the Figure1, we want to compute the derivative of Y with respect to the parameters tau, T and K

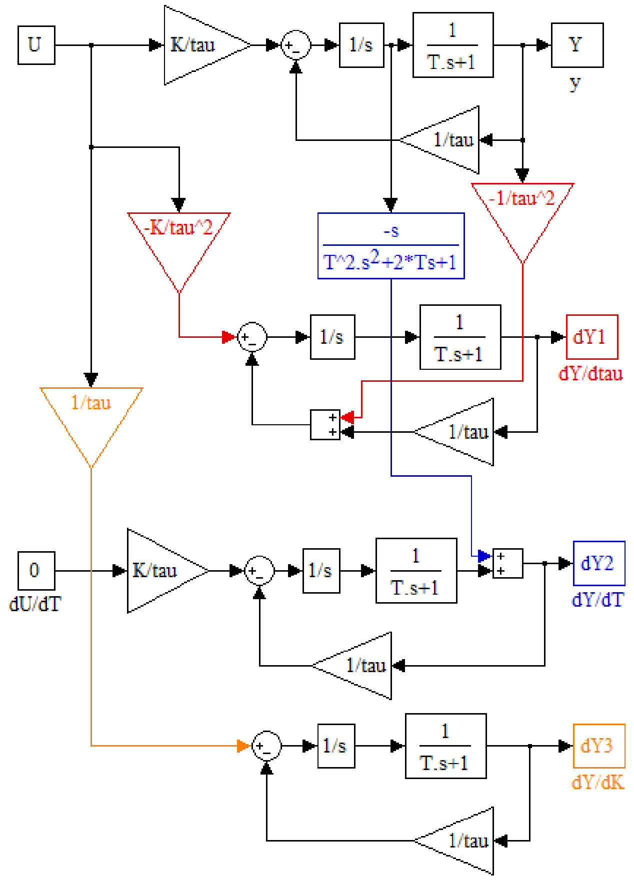


Figure 2: Derivatives with respect to the model parameters

This diagram was obtained by applying rules based on the technical field of "Automatic Differentiation" to derive the conditional structures as well as the structural properties of block diagram and by applying of the formula for derivatives of composite functions.

Seven rules of differentiation, including two rules on the links and five on the blocks allow to build the derivative model.

2.2. Links rules

- Whatever the block through by the differentiated flow, all its outputs will be affected by it. This rule can be applied for scalar links, vector and matrix
- All blocks unaffected by the derivative flow and not depending on derivative parameter can be considered as a constant source equal to zero like constant block, source from workspace , etc In other words, when the inputs do not carry the flow derivative with respect to derivative parameter, the block does not appear in the derivative block diagram.This rule is useful for simplifying the derivative model.

2.3. Blocks rules

- All the inputs/outputs of each of block and the sub block of the original scheme should be accessible at every step time of the simulation.
- For all linear blocks H(k) in U through by differentiated flow, one can built the following scheme that contains the original block diagram increased of derivative with respect to a parameter k. In fact, when the blocks are linear and do not depend on the differentiation parameter, we just need to duplicate the model into the derivative flow. The original model is increased as many times as there are parameters with respect to which differentiation is performed.
- For a nonlinear bloc, it may be necessary to use a computer algebra system to compute the derivatives, according to the mathematical definition of the block.
- For a conditional block (Switch, hysteresis, max, min, trigger subsystem, logic(and, or, etc), state flow, saturated integrator ...) which is defined by a piecewise or event function, we duplicate the block and we keep the same logical tests as in the original block but the outputs contain the derivative flow.
- In the case of "black box" block, the mathematical equations are not accessible and nothing can be done. So one need to retreat them using finite difference.

These differents rules are detailed in (John, Clara and François 2017)

3. FINITE DIFFERENCE METHOD

This is an approximation of $\frac{dY}{dp_i}$ using first order Taylor formula. (Griffith, D. V. and Smith M. 1991)

$$\frac{dY}{dp_i} = \frac{Y(p_1,\ldots,p_i+\Delta p_i,\ldots,p_N) - Y(p_1,\ldots,p_N)}{\Delta p_i} \quad (1)$$

Where N is the number of parameters and Δp_i the step of derivation of parameter p_i.

It is therefore sufficient to disturb each parameter in order to obtain an approximation of the Jacobian and Hessian.

This method is very simple to implement because it does not require modification of the direct model, and can intervene completely independently. On the other hand, it is very slow for a large number of parameters, since it requires N + 1 simulations. It will therefore be of little interest for a large number of parameters. In addition, it can be very inaccurate in case of a highly dynamic system

4. PARAMETERS IDENTIFICATION MODEL

The identification process is to find numerically a set of model parameters, which correlates the best possible predictions and experimental results. It is based on minimizing the difference between the recorded model response and the given experimental result. Such a difference can never be zero. However, the rule states that when the difference is smaller, the set of parameters is better. In this work, the identification of the model parameters uses the block diagram of Figure3 (Davide A. 2004).

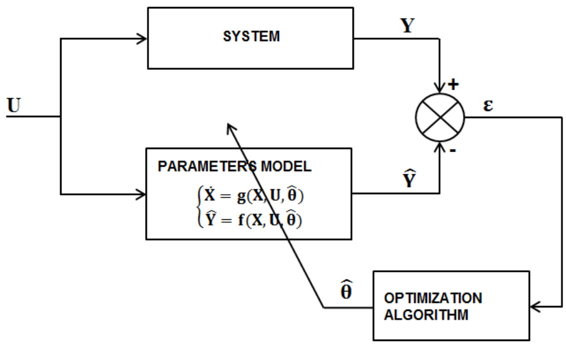


Figure 3: Parametric identification model

Where $\hat{\theta}$: model parameters. The output $\hat{Y}$ is obtained by simulation of the model parameters. The identification is done by minimization of the quadratic criterion function of the output error ε. The optimization algorithms that we use are those based on gradients, for their speed of convergence. The model parameters of PMSM is described in the forthcoming paragraph 4.1.

4.1. PMSM modeling

The modeling of the PMSM is already the subject of many works. A model in the Park reference is used in this work.
The stator dq-winding Voltages can be expressed as follows (Sturtzer and E. Smichel 2000; Abdessemed and M. Kadjoudj 1997):

$$V_{sd} = R_s.i_{sd} + L_s \frac{di_{sd}}{dt} - L_s.\omega.i_{sq} \qquad (1)$$

$$V_{sq} = R_s.i_{sq} + L_s \frac{di_{sq}}{dt} + L_s.\omega.i_{sd} + K_A.\omega \qquad (2)$$

Where, i_{sd}, i_{sq} are the stator currents, V_{sd} and V_{sq}, the stator voltages, R_s and L_s denote respectively stator and inductance cyclic stator, p and ω are the number of pair of poles and the speed of the equivalent dq-windings (in electrical rad/s) in order to keep the d-axis always aligned with the stator magnetic axis (Kimbark 1995), K_A is a coefficient characterizing the machine(maximum flux megnet).
The speed ω is related to the actual rotor speed Ω as:

$$\omega = p.\Omega \qquad (3)$$

In the normal speed range below the rated speed, the reference for the d-winding current is kept zero (ids = 0). The electromagnetic torque can be expressed as follows:

$$C_{em} = p.K_A.i_{sq} \qquad (4)$$

The differential equation which characterizes the mechanical behaviour of the unit harnesses – PMSM is given by (Cardenas and Dobson 1996; Gerqud 2002):

$$C_{em} - C_r = J\frac{d\Omega}{dt} + f\Omega \qquad (5)$$

Where J *and* f are respectively the inertia and friction coefficient of the machine, *Cr* is the load torque.
The simulation block diagram for the PMSM is shown in Figure 4.

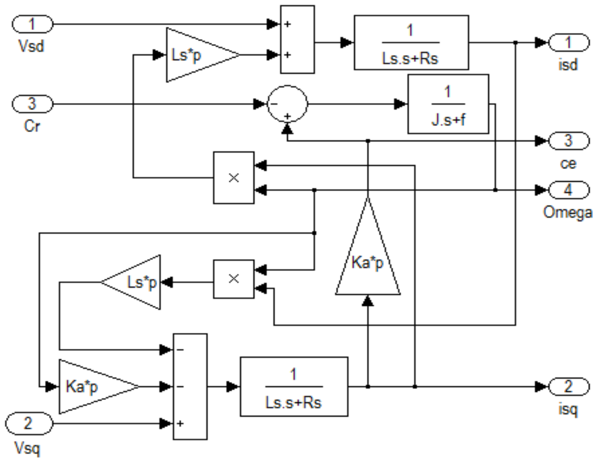


Figure 4: PMSM block diagram

In this diagram, the parameters to be identified are the electrical parameters (Ls, Rs, Ka) and the mechanical parameters (J, f).

4.2. Diagram block for the identification of electrical parameters

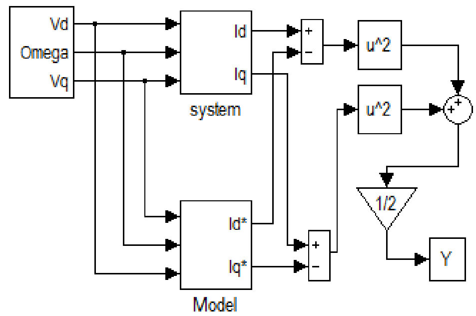


Figure 5: Diagram for identification of electrical parameters

After computing the first derivatives and the Hessian, using the Diffedge Methodology in section 2, we have the following block diagram

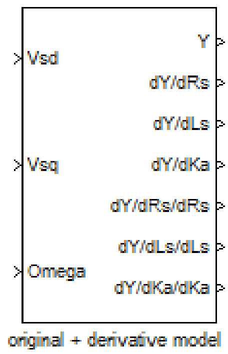


Figure 6: Original model + derivative model with respect to parameters Ls, Rs and Ka

4.3. Diagram block for the identification of mechanical parameters

For identifying the mechanical parameters, we used closed loop speed control.
The mechanical behavior is described using the mechanical equation (5). Thus, a PI type corrector is sufficient to establish the speed loop with the desired dynamics. Considering the perfectly regulated current, the current loop will be equivalent to a unit gain and the speed loop diagram is given in Figure 7

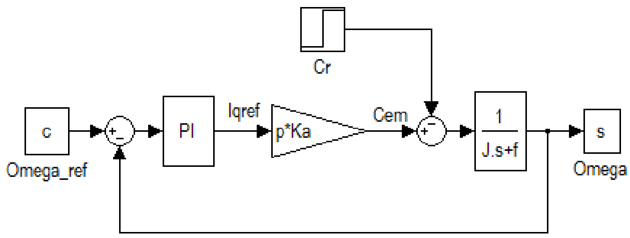


Figure 7: Speed loop controller

Still using the Diffedge Methodology described in section 2, we have the block diagram of Figure 8 for the identification of mechanical parameters:

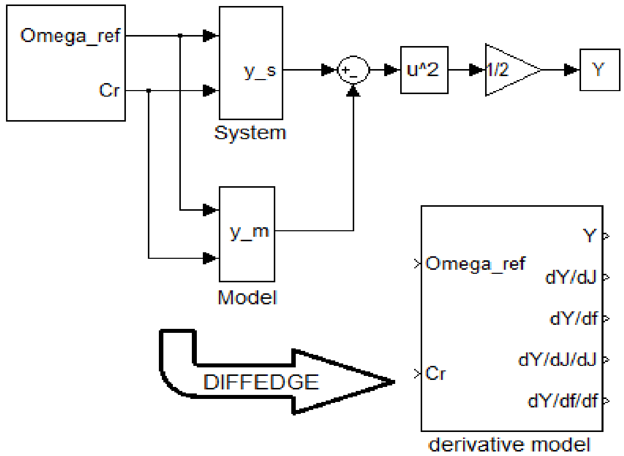


Figure 8: Original model + derivative model with respect to parameters J and f

5. NUMERICAL APPLICATIONS

The proposed new identification strategy is applied to a three-phase motor of 1kW, 200 V, 6000 rpm.

Initially, a database for the excitation signals and output signals of the system is digitally composed using the parameters summarized in (Table 1)

Table 1: Parameters used to create the database

Model parameters		
Parameters	Values	Unit
Ls	$1.365\text{x}10^{-3}$	H
Rs	0.424	[illegible]
Ka	0.117	V/(elect.rad/s)
J	$3.4\text{x}10^{-4}$	kg.m²
f	10^{-4}	N.m.s

An artificial experimental space is generated by simulating the model with this set of parameters which will be named afterwards nominal parameters.

For each model (Figure 6 and Figure 8), we launched two optimization processes. The first uses the Jacobian calculated by finite differences. The second uses the formal Jacobian calculated by Diffedge. Once Jacobian calculations are complete, the algorithm setting is the same for both processes

5.1. Identification of electrical parameters

In addition to the nominal parameters, the necessary data is that of Vd, Vq and the mechanical speed, to carry out the identification.

The nominal value of the mechanical speed is used. The references of the voltages Vd and Vq are identical and which are trapezoids of variable amplitude.

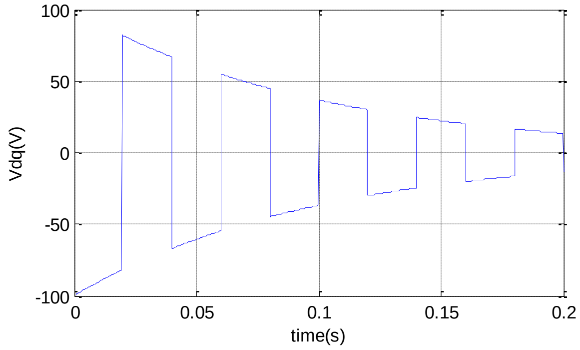


Figure 9: Reference voltages Vdq

Multiple numerical tests were performed to select the optimal derivation step for each parameter. Figure 10 shows the comparison of the derivatives with respect to the parameters Ls (here dp= Ls/Ls-nom).

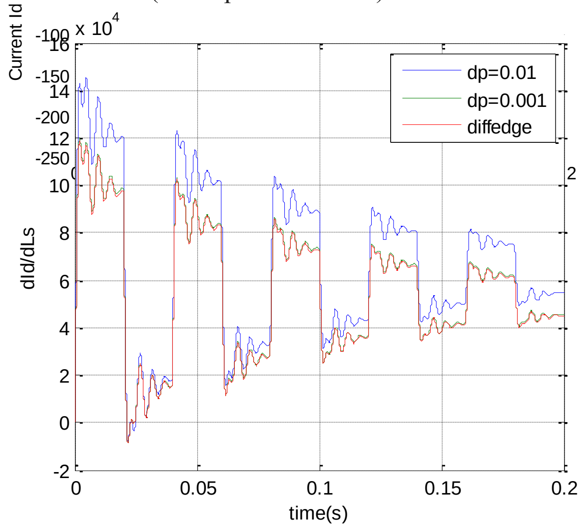


Figure 10: Influence of the step derivation

In this figure, it can be seen that the accuracy of the finite difference method depends on the choice of the step derivation. We opt for the choice of dp = 0.001 to perform the identification. With this choice, we have the same precision as Diffedge. The optimization algorithm used is GAUSSS-NEWTON (Dennis, J.E., More, J.J. Holland, M. 1977).

Starting from the same initial values of the parameters, the two optimization processes are started. The optimized parameters are summarized in (Table 2)

Table 2: Initial and optimized electrical parameters

Model Parameters	Ls (H)	Rs (Ω)	Ka V/(elect.rad/s)
Initial value	$1.00\text{ x}10^{-3}$	0.650	0.085
Optimized with FDM	$1.17\text{ x}10^{-3}$	0.387	0.131
Optimized with DFG	$1.29\text{ x}10^{-3}$	0.442	0.097
Nominal value	$1.36\text{x}10^{-3}$	0.424	0.117

The values of the optimized parameters are almost identical whether finite difference derivation (FDM) or Diffedge derivation (DFG) are used. We have taken as stopping criterion: norm of the objective function at 10^{-4}.

Proc. of the Int. Conference on Modeling and Applied Simulation 2018,
ISBN 978-88-85741-07-2; Bruzzone, De Felice, Frydman, Longo, Massei and Solis, Eds.

In this identification, the two methods are of equivalent precision. On the other hand, the method using Diffedge is much faster (with a calculation time saving of a factor 2.5). Note that in both methods the estimated inductance Ls has the highest relative error. This is because the estimation of this parameter depends on the derivatives of the currents which are calculated numerically and can introduce additional noise.

.

5.2. Identification of the mechanical parameters

The identification of the mechanical parameters is done in a closed loop of speed. The speed reference is also a variable amplitude trapezoidal signal that starts with $\Omega_{ref} = 628 rad/s$

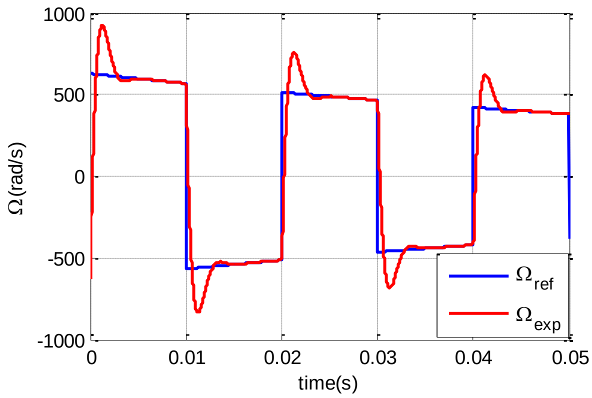


Figure 11: Reference and output speed

These results were obtained using the set of nominal parameters. They are considered as experimental data to identify the mechanical parameters.
After launching the identification process the two methods gave the following results:

Table 3: Initial and optimized mechanical parameters

Model Parameters	J (kg.m²)	f (N.m.s)
Initial value	2.50×10^{-4}	1.50×10^{-4}
Optimized with FDM	3.37×10^{-4}	0.99×10^{-4}
Optimized with DFG	3.38×10^{-4}	0.98×10^{-4}
Nominal value	3.40×10^{-4}	10^{-4}

Note that the mechanical parameters are better estimated than the electrical parameters. Indeed, the signals obtained in speed loop are less noisy than those of currents.
The speed has a very slow dynamic compared to the current and it does not present noise like that of the current. It is for this reason that we did not need to present the influence of choice of derivation steps.
At equal accuracy, the capture of computation time has further shown that the Diffedge-based method is twice as fast as that using the finite difference.

5.3. Simulation validation of identification results

To validate the obtained parameters (electrical and mechanical parameters), a simulation of the complete diagram of a PMSM vector control was carried out.

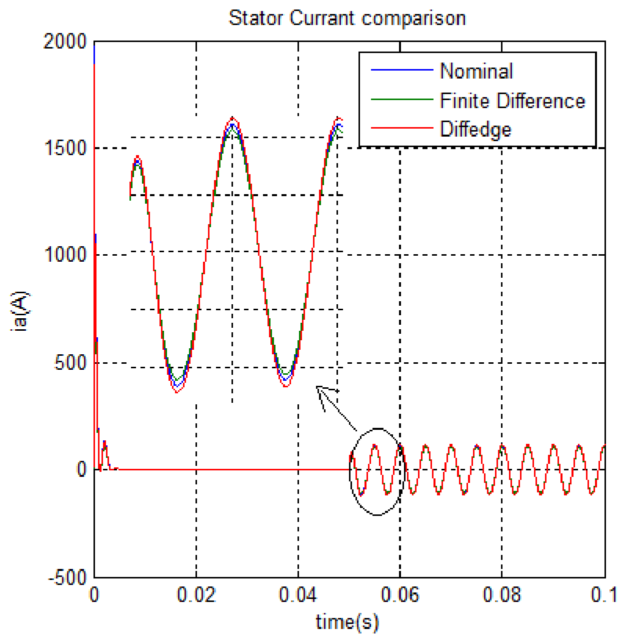


Figure 12: Comparison of stator currents

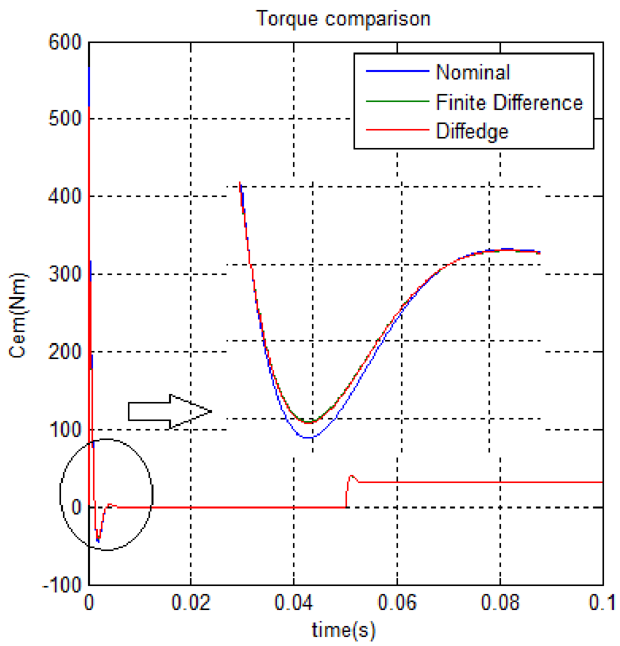


Figure 13: Comparison of the electromagnetic torque

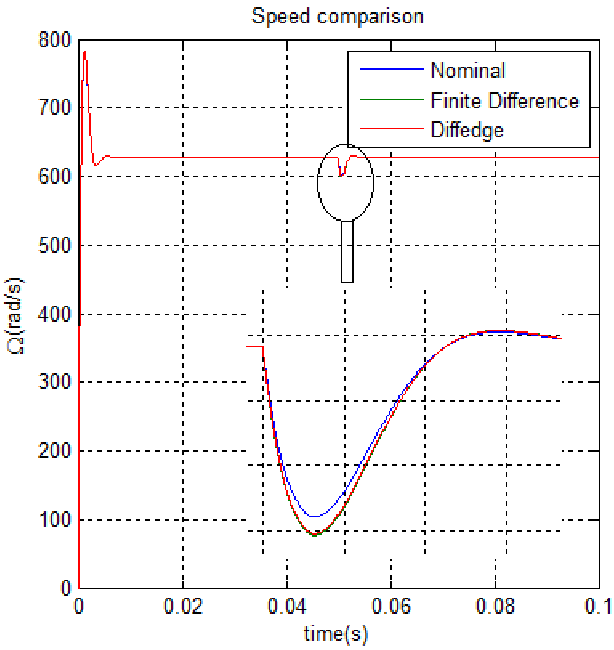


Figure 14: Comparison of the mechanical speed

The results presented in Figure 12 to Figure 14 were obtained by simulating the behavior of the during starting, subject to a load torque at t = 0.05s Through the results Optimization given the same tips as nominal parameters.

6. CONCLUSION

The objective of this work is to propose a new strategy to optimize the identification parameters of a PMSM model. Hence, diffedge methodology is used, method based on the formal calculation through block diagrams. The gradient and hessian calculation for the optimization algorithm is done in the same simulation environment as the model. This approch is tested to identify electrical and mechanical parameters of the PMSM.

The results obtained are rather encouraging, since, with equal precision, the calculation time for the Diffedge method is much lower than that obtained by finite differences.

It is recognized that this methodology shows its ability to optimize the model parameters both the electrical parameters and the mechanical parameters.

Consequently, the predicted responses describe faithfully the experimental results.

REFERENCES

Gilbert R.J.C, Le Vey G. and John M., 1991. La différentiation automatique de fonctions représentées par des programmes Tech rep 1557 INRIA

John M. and Cambois T., 2004. "Differentiation, sensitivity analysis and identification of hybrid Models under Simulink", Symposium Techniques Avancéeés et Stratégies Innovantes en Modélisation et Commandes Robustes des Processus Industriels, Martigues, 21 & 22 septembre 2004.

Bastogne T., Thomassin M. and John M., 2007. "Selection and identification of physical parameters from passive observation. Application to a winding process" Control Engineering, Practice, Elsevier, 2007, 15 (9), pp.1051-1061.

John M. and Clara M., 2004. Automatic differentiation of hybrid models Illustrated by Diffedge Graphic Methodology. APPEDGE 18–22, rue d'Arras 92000 Nanterre, France

Griffith, D. V., Smith, M., 1991 « Numerical Methods for Engineers », Blackwell Scientific Publication

Davide A, 2004. « Identification des paramètres du moteur à induction triphasé en vue de sa commande vectorielle », Mémoire de maitrise en génie électrique, Université de Laval, Quebec, Décembre 2004.

Kimbark E.W, 1995. Power system stability : Synchronous Machines, ISBN 0-7803-1135-3, IEEE Press.

Sturtzer G. and Smichel E., 2000. Modélisation et Commande des Moteurs Triphasés., Edition Ellipses, Paris.

Cardenas-Dobson R., 1996. Control of Wind Turbine Using a Switched Reluctance Generator. PhD Thesis, University of Nottingham.

Dennis J.E, More and Holland J., 1977. Quasi-Newton methods, motivation and theory. SIAM Review, Vol. 19, No. 1, pp. 46-89.

AUTHORS BIOGRAPHY

Rakotoarisoa A. Jean Claude. Born in 1976. Doctor Engineer, He works in the laboratory of Automatic. His theme consists of optimization and identification, applied to electrical, mechanical systems and renewable energies

Razafinjaka N. Jean. Born in 1956. Doctor Engineer HDR. He works now in the Laboratory of Automatic, Polytechnic Higher School of Diego Suarez, Madagascar. His field of research relates to the advanced commands applied to electric machines, electromechanical systems and renewable energies.

Hadjee G. Eric Born in 1975. Doctor Engineer, He works in the Electricity and Electrical Network Laboratory. His theme consists of optimization and FACTS, applied to electrical network renewable energies.

A STUDY OF FOOT STRUCTURE FOR HUMANOID ROBOT ON ROUGH TERRAIN

Van-Tinh Nguyen[(a)], Hiroshi Hasegawa[(b)]

[(a), (b)]Graduate School of Engineering and Science
Shibaura Institute of Technology, Japan
[(a)]School of Mechanical Engineering
Hanoi University of Science and Technology, Vietnam

[(a)]nb16508@shibaura-it.ac.jp, [(b)]h-hase@shibaura-it.ac.jp

ABSTRACT
This paper presents a foot structure for humanoid robot locomotion on rough ground called an adaptive foot which is a novel structure proposed in our previous study when considering the robot locomotion on flat ground. Realizing stable walking on rough ground is essential, because humanoid robots are expected to assist humans in not only a flat known domestic environment but also on rough ground surfaces with terrain variances while maintaining their tasks. The mentioned foot structure equips the robot with a good adaption. It enables the foot to increase the contact points and improve the stability on complex surface. Moreover, this structure also reduces the effect of impact force on the robot foot at toe off period in the walking. In this study, a gait pattern is generated by an approximated optimization method based on Response Surface Model (RSM) and Improved Self-Adaptive Differential Evolution Algorithm (ISADE). The result is experimentally validated through dynamic simulation on Adams (MSC software, USA) with the Kondo robot.

Keywords: humanoid robot, rough ground, foot structure, gait pattern, optimization

1. INTRODUCTION

In recent years, the application of a humanoid robot is spread out into many fields in the human life. Some of the humanoid robots are even well-known in all over the world (Sakagami et al. 2002; Ogura et al. 2006; Ishida et al. 2004). The primary purpose is for the humanoid robot to walk on flat ground, but not on rough surfaces which are often encountered in our living environment. The reason is simple: walking controller implementation for humanoid robots is extremely challenging even on flat terrains due to the many degrees of freedom to be controlled simultaneously. Nowadays, development of application in human daily life is forcing humanoid robots to walk in natural or artificial environments, which do not have perfectly flat surfaces prepared for the robots to walk on. In this paper, our ultimate target is to solve humanoid robot walking problem on corrugated surface where bipedal locomotion is not very stable.

Surveying the existing literature, the study witness very few publications dealing with humanoids walking on rough terrains. Walking on uneven surfaces such as slopes and inclines have received much more attention (Wang et al. 2014; Ali et al. 2012; Seven et al. 2011). A number of other works study biped motion concerned with walking on level surface such as up and down stair (Sheng et al. 2009; Qin et al. 2013).

A few researches focus on a gait generation method for biped locomotion on rough ground. In detail, Sang-Ho Hyon and Gordon Cheng presented an adaptive control method for humanoid robot on unknown rough terrain. The adaption is achieved by an optimally-distributed anti-gravitational forces (Hyon and Cheng 2007). In the same way, Mitsuharu Morisawa et al. proposed a biped locomotion control for uneven terrain with narrow support region. In their work, a walking capability of the robot is enhanced by using the information of a support region (Morisawa et al. 2014). In (Zheng et al. 2013), Y. F. Zheng et al. proposed two new types of gaits named “step-over” and “ski-type” to overcome challenge when humanoid robot moves on uneven surfaces.

On the other hand, some researchers have realized that the conventional foot with a rigid and flat sole has sufficient support polygon area, which is a convex polygon of minimum area including all contact points, only when the robot walks on flat ground. On rough surface, contact states often become one-point contacts or line-contacts, thus the support polygon becomes too small, centre of gravity (CoG) point moves out of the support polygon, and causing the robot unstable. As a result, some previous papers have developed a new foot structure for a humanoid robot to adapt to a rough environment. For instance, to realize stability on complex ground surface, Moyuru Yamada et al. developed a biped robot with a point-contact type foot with springs. It suppressed the impact force at foot landing and provided a stable contact states on rough terrain by adapting geometrically to complex surfaces (Yamada et al. 2011). Likewise, Yokomichi and Ushimi (2012) proposed a new foot structure with three toes with frictional locking and unlocking mechanism to avoid the influence of reaction force on biped robot when walking on unknown rough terrain such as a craggy place. In the same way, this paper proposed a novel foot structure with toes inspired by human foot. The mentioned structure helps the humanoid robot easily adapt to rough environment, for example, a corrugated surface. It enables the foot to increase the contact points and improved the stability. The rest of this paper is organized into four sections. Section 2 describes the mechanical structure of the humanoid robot. The principle of gait pattern generation is in Section 3. Section 4 shows the results of simulation experiments by dynamic simulation on Adams. Finally, Section 5 includes some brief conclusions and future works.

2. EXPERIMENTAL ROBOT MODEL

2.1. Overview of Structural Design

In this study, the proposed model is built based on the KHR-3HV robot of Kondo Kagaku Company which is the third generation of a humanoid robot developed by this company. The KHR-3HV robot has the weight of 1.5kg, the height of 401.05mm and up to 22 DOFs with 17 actual servos and 5 dummy servos. However, in this work, robot legs are concentrated. Thus, upper body joints are fixed and lower body have 10 controlled joints for the legs as shown in Figure 1.

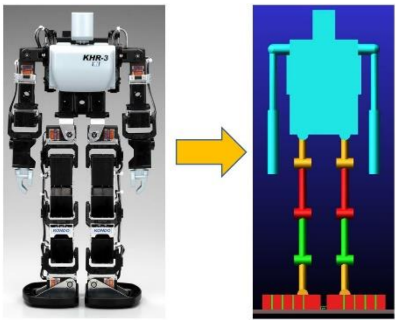


Figure 1: Real robot and proposed robot

2.2. Foot Mechanism

In the locomotion, the human foot support area continuously varies on the sole of foot as depicted in Figure 2. The black area is position where supports forces areas. Wherewith, LR is heel only in loading response, MSt is foot flat in mid stance, TSt is forefoot and toes is terminal stance, and PSw is medial forefoot in pre-swing. Perry and Burnfield (2010) found that toe contact with ground is quite variable. The onset of toe involvement followed insolated forefoot support by 10% of the stance period. In this period, toe pressures differ markedly with the greatest pressure of the big toe. It ranged between 30% and 55% of that at the heel. Thus, the big toe has an important role in the human walking, especially the toe-off period.

By this idea, the seven foot mechanisms are designed for simulation experiments inspired by human foot as depicted in Figure 3. The paper plans to consider the effect of the novel foot structures on the robot walking behavior with the corrugated ground which has two waves with the height of 4mm. The length of uneven segment is 90mm.

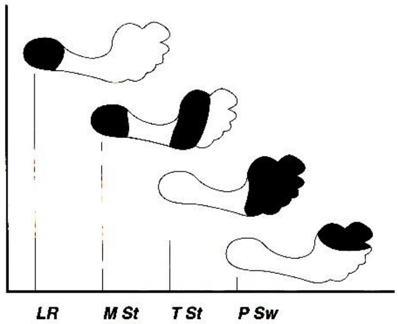


Figure 2: Sequence of foot support areas during stance

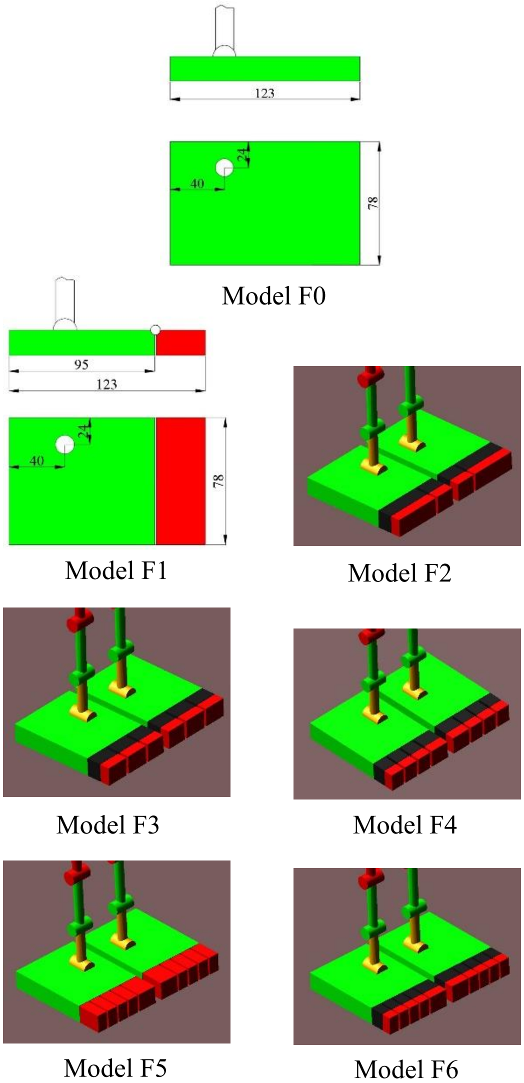


Figure 3: Foot mechanism

In Figure 3, Model F1 shows the basic parameters of the robot's feet. Chockalingam and Ashford (2007) have proved that average ratio between the foot length and heel varies from 1.196 to 1.426. In this model, the length of heel and foot are 95mm and 123mm, respectively. Hence, the length of toes will be 28mm. The ankle installation position is determined based on the real robot. These parameters are used for all the remain models.

As proven in (Nerakae and Hasegawa 2014), the biped robot whose big toe width ratio per foot equals 0.28, has the longest walking distance when big toe length is fixed and this ratio is similar to the ratio of the human's feet. Thus, the width of the big toe and the feet were designed of 22mm and 78mm. This ratio is applied to model F2, F3, F4, F5 and F6. Spring stiffness coefficients is respectively set of 0.39N.mm/deg, 0.20N.mm/deg to big toe joint and remain toe joints.

The walking behavior of the robot with the adaptive foot structure is illustrated in a cycle described in Figure 4. Toe mechanism is expected to have a bending motion and enables the robot to overcome the obstacles.

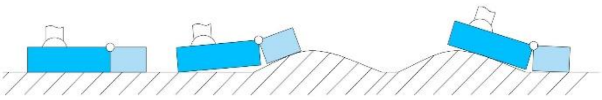


Figure 4: Adaptive walking behavior

In the second experiment, the paper investigated the effect of the torsion stiffness of the passive joint on the biped walking distance. It is considered in a predefined range as described in Table 1.

Table 1: Torsion spring stiffness coefficient

No	d(mm)	T(Kg.mm/210°)
B1	9.14	3.23
B2	9.25	4.38
B3	9.30	5.88
B4	9.37	7.60
B5	9.53	10.60
B6	11.71	13.71
B7	11.81	16.82
B8	11.81	17.40
B9	11.96	22.81
B10	14.12	27.19
B11	14.33	38.14
B12	14.45	45.28
B13	14.68	58.41
B14	19.05	85.26
B15	19.41	120.63
B16	23.55	133.30
B17	23.77	162.57
B18	23.98	191.83
B19	28.40	256.24
B20	28.60	295.87

3. GAIT PATTERN GENERATION

3.1. Definition of Joint Angle

The joint angles are defined as depicted in Figure 5 and these specifications are described in Table 2.

3.2. Gait Function

Basing on the human walking pattern as depicted in (Whittle 2007), this study assumes that the robot control data is generated by the gait function as trigonometric function shown in Equation (1). By changing a, b, c, d coefficients, the gait functions will be created to allocate to each joint of the biped robot.

$$\varphi_i\,(t) = a_i + b_i.\cos(\omega t) + c_i.\sin(\omega t) + d_i.\cos(2\omega t) \quad (1)$$

Where a, b, c, d are coefficients, t is time, ω is angular velocity and i is index of joint.

In toe mechanism, due to considering a reduction in energy consumption of the robot, the passive joint is selected as a toe joint. Consequently, φ_{8r}, φ_{8l}, φ_{9r} and φ_{9l} have a value in the range from 0° to 30°. Their values depend on the robot geometric posture as well as impact forces when the robot performs its motion.

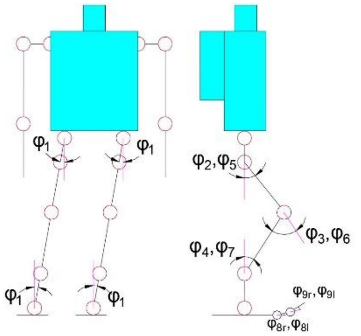


Figure 5: Definition of joint angles

Table 2: Angle specification

Angle	View plane	Leg	Joint	Value
φ_1	Frontal	Both	Hip & ankle	-15° to 15°
φ_2	Sagittal	Right	Hip	-50° to 50°
φ_3	Sagittal	Right	Knee	0° to 60°
φ_4	Sagittal	Right	Ankle	-50° to 50°
φ_5	Sagittal	Left	Hip	-50° to 50°
φ_6	Sagittal	Left	Knee	0° to 60°
φ_7	Sagittal	Left	Ankle	-50° to 50°
φ_{8r}	Sagittal	Right	Proximal phalanx	0° to 30°
φ_{8l}	Sagittal	Left	Proximal phalanx	0° to 30°
φ_{9r}	Sagittal	Right	Distal phalanx	0° to 30°
φ_{9l}	Sagittal	Left	Distal phalanx	0° to 30°

3.3. Optimization Procedure

The concept of the optimization process is described in Figure 6.

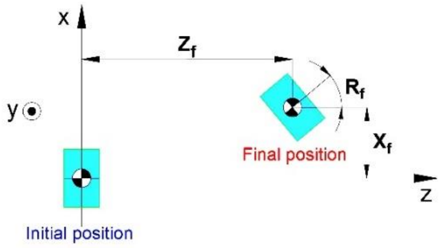


Figure 6: Optimization parameters

Figure 6 describes the parameters of the optimization process. Where X_f, Z_f, and R_f are respectively side distance, straight distance, and angle of rotation at the final position of the robot. Design variable vector, objective function, constraint function and penalty function are defined as described in Equation (2-7).

Design variables (DVs):

$$x = [a_i, b_i, c_i, d_i],\ i = 1 \div 4 \tag{2}$$

Constraint functions:

$$h_1(x) = 243.53 - Y_f = 0 \tag{3}$$

$$h_2(x) = N - 360 = 0 \tag{4}$$

Where Y_f is distance from CoG to the ground. N is a total simulation step.
Objective function:

$$f(x) = -Z_f \rightarrow \min \tag{5}$$

Penalty function:

$$P(x) = \sum_{i=1}^{2}[h_i(x)]^2 \tag{6}$$

Modified objective function:

$$F(x) = -Z_f + \gamma.P \rightarrow \min \tag{7}$$

There are two constraint functions. Equation (3) ensures the robot not to slip at the final framework in the simulation. In Equation (4), N is equal to 360 to check the success of the simulation. In Equation 7, γ is a penalty coefficient set to 1000. Before applying ISADE for the optimization problem, Z_f and Y_f are approximated by RSM through random sampling. The approximated optimization process is depicted as Figure 7.

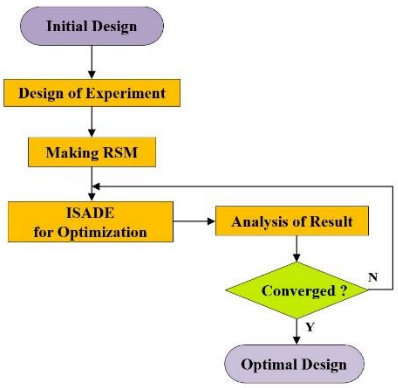


Figure 7: Overview of optimization

4. SIMULATION RESULT

4.1. First Experiment

In all experiments, the robot motion is simulated in five cycles. One cycle is set up to 1.2 seconds. Thus, five cycles spend on 6.0 seconds. Next, 1.2 seconds is used for checking robot stability. In this simulation, one step takes 0.02 second, so the total number of steps is 360. The simulation result is shown in Table 3.

Table 3: Simulation result

Model	Distance		Rotation
	X_f(mm)	Z_f(mm)	R_f(deg)
F0	-34.50	206.17	7.28
F1	-90.61	300.78	41.31
F2	-43.96	299.86	25.50
F3	-100.55	301.23	-0.633
F4	-117.11	298.65	36.27
F5	-6.45	316.13	10.89
F6	82.85	262.48	15.07

The trajectory of the CoG point is shown as in Figure 8.

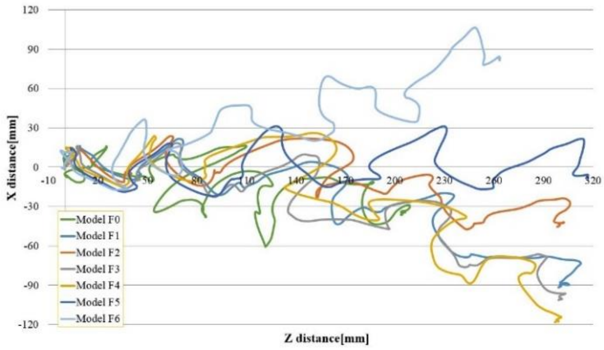


Figure 8: Result of the first experiment

In Table 3 and Figure 8, as can be seen that all the trajectories of the robot's CoG are approximately the waveform of the circular function which is similar to that of the human beings. With the model F0 having no toe mechanism, it is encountered to climb up the obstacles due to the rigid foot structure, thus, the walking distance is unexpected. By contrast, the robot with toe mechanism can overcome the obstacles easily. When the number of toe increases, the foot is more flexible to enhance the number of contacting point. Therefore, the robot performs the steady walking. However, in this case, if the toe mechanism has so many passive joints such as model F6, the number of contacting point will be so big, the foot got stuck in concave part of terrain. As a result, the walking distance declines. In consideration of straight walking and distance, model F5 with X_f, Y_f, R_f of -6.45mm, 316.13mm, and 10.89° has the best performance. Accordingly, this model is selected to do the research further. The walking behavior of model F5 is depicted in Figure 9. As can be seen, to overcome the obstacles on the terrain, the robot performs the bending motion of the toe which cultivates the contacting points and confine the sudden change of the ground reaction force, thus, enabling the robot to walk steadily.

The waveform of the gait functions allocated to all joints of the humanoid robot is show in Figure 10.

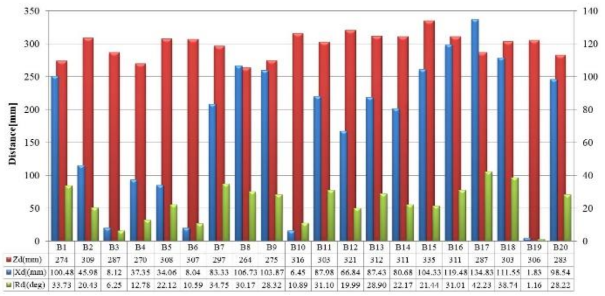
4.2. Second Experiment

The result of the second experiment is shown in Figure 11. As can be seen that all model can walk on the rough ground. Side distance and angle of rotation in the simulation have a little difference when comparing with

Proc. of the Int. Conference on Modeling and Applied Simulation 2018,
ISBN 978-88-85741-07-2; Bruzzone, De Felice, Frydman, Longo, Massei and Solis, Eds.

the calculation results on account of the approximating method. When increasing the torsion stiffness, the toe mechanism becomes harder and the robot has a trend to slip on the ground, thus, the side distance goes up as well. On the contrary, the foot structure becomes more flexible to adapt to the waving terrain, as a result, the side distance witnesses a decline.

In consideration of walking straight and distance, model B10 with the stiffness coefficient of 27.19 Kg.mm/210º has the best performance.



	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15	B16	B17	B18	B19	B20
Zd(mm)	274	309	287	270	308	307	297	264	275	316	303	321	312	311	335	311	287	303	306	283
|Xd|(mm)	100.48	45.98	8.12	37.35	34.06	8.04	83.33	106.73	103.87	6.45	87.98	66.84	87.43	80.68	104.33	119.48	134.83	111.55	1.83	98.54
|Rd|(deg)	33.73	20.43	6.25	12.78	22.12	10.59	34.75	30.17	28.32	10.89	31.10	19.99	28.90	22.17	21.44	31.01	42.23	38.74	1.16	28.22

Figure 11: Result of the second experiment

5. CONCLUSIONS

A foot structure with toes for a small humanoid robot is proposed in this study when considering the robot walking on the corrugated surface. The locomotion of the robot is simulated by multi-body dynamics simulation software, ADAMS. As a result, all the models can walk stably with the bending of toe which is similar to that of human beings. Model F5 obtaining the best performance, which has one big toe, four small toes and no proximal interphalangeal joint. In the second stage, the paper investigated the effect of spring stiffness on the walking distance to find out the consistent stiffness for the passive joint of the robot's toe.

As for the future work, the paper tries to consider the effect of the other structure parameter such as length of toe on the walking behavior of the robot and increasing the height of the obstacles.

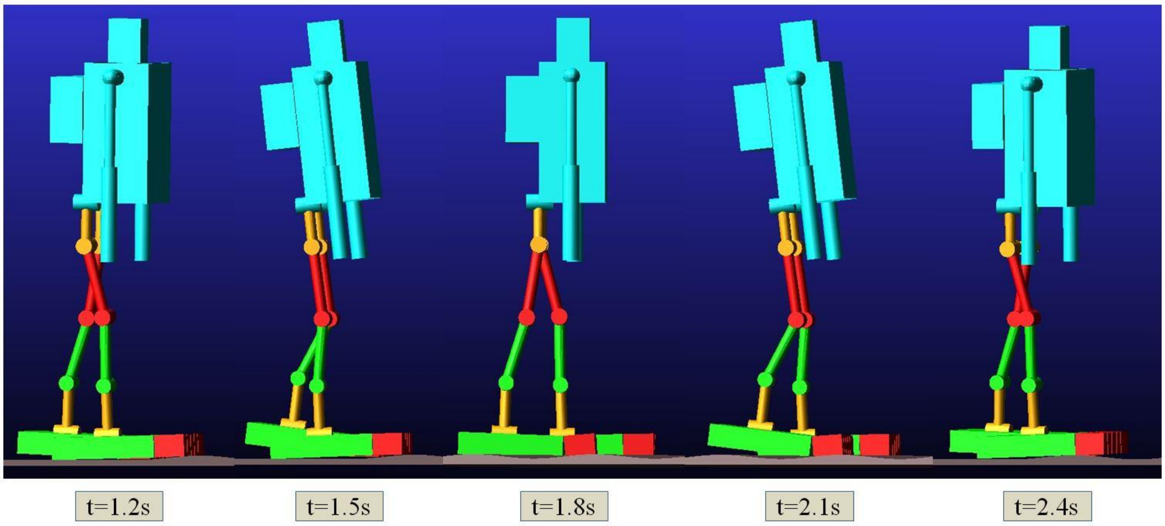
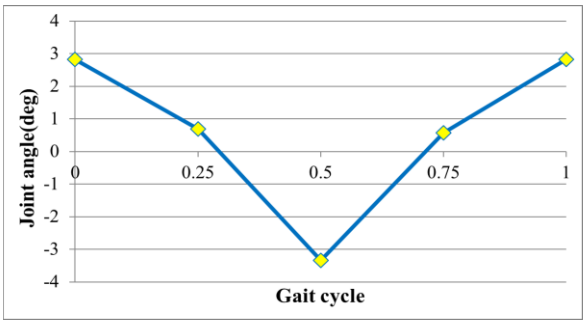
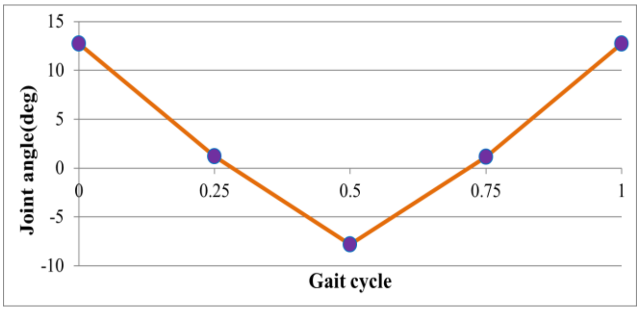


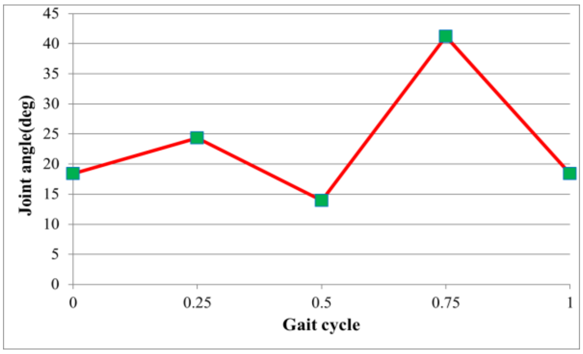
Figure 9: Robot walking behavior on rough ground



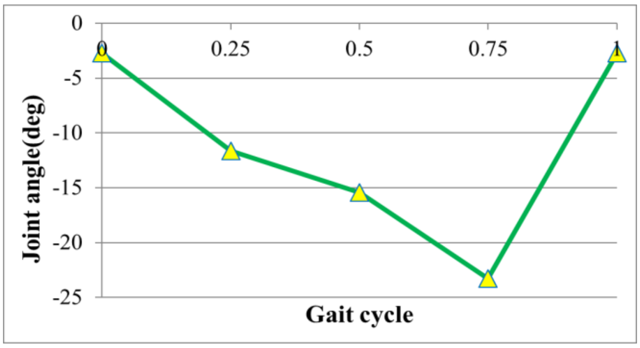
(a) A cycle of gait function (hip and ankle roll joint angle)



(b) A cycle of gait function (hip pitch joint angle)



(c) A cycle of gait function (knee pitch joint angle)



(d) A cycle of gait function (ankle pitch joint angle)

Figure 10. Waveform of the gait function

REFERENCES

Sakagami Y., Watanabe R., Aoyama C., Matsunaga S., Higaki N., and Fujimura K., 2002. The intelligent ASIMO, system overview and integration. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2478–2483, Switzerland.

Ogura Y., Aikawa H., Shimomura K., Morishima A., Lim H. O., and Takanishi A., 2006. Development of a new humanoid robot WABIAN-2. Proceedings of IEEE International Conference on Robotics and Automation, pp. 830-835, Orlando, Florida.

Ishida T., 2004. Development of a Small Biped Entertainment Robot QRIO. Proceedings of the Fourth Symposium Micro-Nano mechatronics for Information-Based Society, pp. 23-28, Nagoya, Japan.

Wang H., Zheng Y. F., Jun Y., and Oh P., 2014. DRC-Hubo Walking on Rough Terrains, Proceeding of the IEEE International Conference on Technologies for Practical Robot Applications, Woburn, USA.

Ali F., Motoi N., and Kawamura A., 2012. An Improved Trajectory of Biped Robot for Walking along Slope, Proceeding of the 12th IEEE-RAS International Conference on Humanoid Robots, pp. 140-145, Osaka, Japan.

Seven U., Akbas T., Fidan K. C., Yilmaz M., and Erbatur K., 2011. Humanoid robot walking control on inclined planes, Proceedings of the IEEE International Conference on Mechatronics, pp. 875-880, Istanbul, Turkey.

Sheng B., Huaqing M., Qifeng L., and Xijing Z., 2009. Multi-Objective Optimization for a Humanoid Robot Climbing Stairs based on Genetic Algorithms, Proceeding of the IEEE International Conference on Information and Automation, pp. 66-71, Macau, China.

Qin Z., Chang-xiang F., Tao Y., and Kamiya Y., 2013. Action Generation of a Biped Robot Climbing Stairs, Proceeding of the IEEE International Conference on Mechatronics and Automation, pp. 1069-1073, Takamatsu, Japan.

Hyon S-H., and Cheng G., 2007. Simultaneous adaptation to rough terrain and unknown external forces for biped humanoids. Proceeding of the 7th IEEE-RAS International Conference on Humanoid Robots, pp. 19-26, Pittsburgh, USA.

Morisawa M., Kita N., Nakaoka S., Kaneko K., Kajita S., and Kanehiro F., 2014. Biped Locomotion Control for Uneven Terrain with Narrow Support Region, Proceeding of the IEEE/SICE International Symposium on System Integration, pp. 34-39, Tokyo, Japan.

Zheng Y. F., Wang H., Li S., Liu Y., Orin D., Sohn K., Jun Y., and Oh P., 2013. Humanoid Robots Walking on Grass, Sands and Rocks, Proceeding of IEEE International Conference on Technologies for Practical Robot Applications, pp. 1-6, Woburn, USA.

Yamada M., Sano S., and Uchiyama N., 2011. Point-Contact Type Foot with Springs and Landing Control for Biped Walking on Rough Terrain, Proceeding of the IEEE International Conference on Robotics and Biomimetics, pp. 2355-2360, Phuket, Thailand.

Yokomichi T., and Ushimi N., 2012. A Study of the Sole Mechanism of Biped Robots to Rough Terrain Locomotion, Journal of Robotic and Mechatronics, Vol. 24, No. 5, pp. 902-907

Perry J., and Burnfield J. M., 2010. Gait Analysis: Normal and Pathological Function, 2nd ed., Slack Incorporated, p. 82

Chockalingam N., and Ashford R., 2007. Selected foot length ratios in a non-pathological sample, Revista Internacional de Ciencisa Podológicas, 1 No. 2 pp. 25-30

Nerakae K., and Hasegawa H., 2014. Big toe sizing design of small biped robot by using gait generation method. Applied Mechanics and Materials, 541-542 pp. 1079-1086

Whittle M. W., 2007. An Introduction to Gait Analysis. 4th ed., Oxford, p. 59

AUTHORS BIOGRAPHY

Van-Tinh Nguyen received the B.E. (2012) from Hanoi University of Science and Technology (HUST), Vietnam, and M.E. (2016) from Shibaura Institute of Technology (SIT), Japan. Besides, he works as a lecturer at School of Mechanical Engineering of HUST. Currently, he is pursuing Dr. Eng. degree in Functional Control Systems at SIT. His research interests include optimization system design, multi-body systems, humanoid robot, and evolutionary algorithm.

Hiroshi Hasegawa received his B.E. in 1992 and M.E. in 1994 from Shibaura Institute of Technology, Japan. He received Dr. Eng. In 1998 in mechanical engineering from Tokyo Institute of Technology, Japan. He has been working at Shibaura Institute of Technology, and currently is a Professor at the Department of Machinery and Control System, College of Systems Engineering and Science. He is a member of JSEE, JSME, ASME, JSCES, JSST and KES. His research interests include computer-aided exploration, creativity of design and systems engineering.

CATEGORIC SIMULATION OF PRODUCTION FLOWS

David Kruml [(a)], Jan Paseka[(b)]

[(a),(b)]Department of Mathematics and Statistics, Masaryk University
Kotlářská 2, 611 37 Brno, Czech Republic

[(a)]kruml@math.muni.cz,[(b)]paseka@math.muni.cz

ABSTRACT

We develop a notion of categoric simulation of production flows. The flow is geometrically depicted as a certain subset in three modes (dimensions) — mass, space, and time. Each of the modes could be organized to a tree structure. We want to see the modes by means of an abstract mathematical notion — a category — which naturally captures the hierarchical/net structure of a mode even that it could be very complex. As a highly abstract notion, categoric simulation provides a broad perspective on simulation as something that can be developed and changed and thus has the ability to be a fundamental method of representing flows. We argue that this methodology has the potential of providing elegant, simple, and efficient solutions to problems arising in the manufacturing, as well as other applications concerning multivalued or varying flows in nets.

Keywords: categoric simulation of production flow, modes of production flow, signal, stack validation algorithm

1. INTRODUCTION

In this article we develop a notion of *categoric simulation* of production flows. This means that we want to see simulation as an abstract mathematical structure - a *category* - consisting of *objects* - places that each processed unit passes at an accurate time, and *morphisms* - devices such as production processes, transfers or stacks. Simulation has been successfully adopted in numerous studies related to manufacturing system design and operation. With simulation, it is also possible to analyze several scenarios and to consider a wide range of performance measures. Categoric simulation is a highly flexible tool which enables us to evaluate effectively different alternatives of system configurations and operating strategies to support decision making in the manufacturing context.

Our categoric model provides the structure of graphs as e.g. in Choi and Kang (2013) or Tanchoco (1994) but adds a *hierarchy* of composed devices. The rationale for developing such a framework is similar to the category-theoretic approach to quantum theory based on Abramsky and Coecke (2004). As the starting point of quantum theory, usually the seminal work Birkhoff and von Neumann (1936) is cited. The idea of Birkhoff and von Neumann has been to gain an understanding of quantum physics from the perspective of a particular non-classical "logic". The primary motivation of Abramsky and Coecke was in this case rather closely related to actual physical phenomena, the aim is to construct a simplified, technically lightweight description of quantum algorithms and protocols. The category-theoretic approach focuses on the processes and the composition of systems as central notions. In our paper we demonstrate that the formalism designed to model quantum circuits can well be used to model classical processes — objects and morphisms describe the flow of events, and the tensor product corresponds to parallel processes.

Our approach is, generally speaking, within the tradition of using category theory as a modern language for the description of networks. Examples come from programming as in Căzănescu and Stefănescu (1990), quantum protocols as in Abramsky and Coecke (2004), electrical circuits as in Baez and Fong (2016) or Petri nets as in Meseguer and Montanari (1990).

The type of category suitable in this context is exactly what has been proposed in the context of quantum protocols: dagger compact symmetric monoidal. It is consequently this framework that we may use for the application that we have in mind: we can simulate the production flow in manufacturing. The reason of doing so is actually of a practical nature: even though the data with which we deal is finite, it is too large to allow for straightforward computations. Hence we need ways how to simplify the complexity of composed devices and to this end the categorical approach has proved to be quite useful.

We argue that it is possible to identify at least three types of simulation models for different kinds of manufacturing:

- a *motion model* — the manufacturing is regular in all aspects,
- a *signal model* — the production is still mass and regular on processes but the processes need not be synchronized, and
- a *snapshot model* — the production is irregular or the number of processed units is small.

The three models analysed in the article in Section 4 seem to be widely shared in the literature (e.g. in Al-Turki, Ayar, Yilbas and Sahin (2014) or Tanchoco (1994)) and in practice. Typically, a firm production may refer to all three types. We want to show how to integrate them, possibly by transforming the structure and functionality of the total production of the firm. Our approach allows machine flexibility and routing flexibility. Machine flexibility means to manufacture different types of products on a machine whereas routing flexibility means the manufacturing of the operations on different machines. Hence our methodology can be also used for flexible manufacturing systems as in Shanker and Srinivasulu (1989).

This paper is a part of the research project *Research and Development of Information System for Value Flow Optimization and Production Planning*. The research project develops both a *complete methodology* and a *software program* for a direct simulation of production flow in manufacturing. The purpose of the methodology is to support simulation engineers to structure their projects and provide helpful experiences for robust modeling, verification, validation, and analysis of the production flow, using structured modeling and transparent code. The target application for the intended methodology is a detailed modeling of manufacturing systems. The production engineers that use the results will obtain a detailed and real time information on how they can improve their area of the production flow and processes.

The rest of the paper is organized as follows: Section 2 describes in detail how the production flow propagates as a certain 2-dimensional manifold through a 3-dimensional mass-space-time state space. Section 3 explains the background to the solution methodology, followed by the details regarding the proposed solution methodology. Section 4 describes the proposed solution methodology with respect to three basic modes of simulation models.

Desctiption of proposed algorithm on the stack validation problem is provided in Section 5, followed by an illustrative example in Section 6. Finally, Section 7 concludes the paper.

In the whole paper, we assume for simplicity that the production is *perfect* with no errors or variance and that all times can be derived from the following parameters:

- *Process time PT* is time of working one unit in a given process.
- *Cycle time CT* is time between transits of two consecutive units through a node.
- *Changeover time CO* is time between two *regular regimes* of a machine.
- *Idle time IT* is time spend by a unit in a stack.

For the simulation purposes, we do not assume that other times come to a play. A comprehensive review of simulation for manufacturing system design and operation which focuses on the application of discrete event simulation can be, e.g., Negahban and Smith (2014).

2. MODES OF A PRODUCTION FLOW

The production flow propagates as a certain 2-dimensional manifold through a 3-dimensional mass-space-time. The dimensions are:

- *Mass N* is a number of processed units.
- *Space l* expresses how the flow progress through a production network.
- *Time t* has its usual physical meaning.

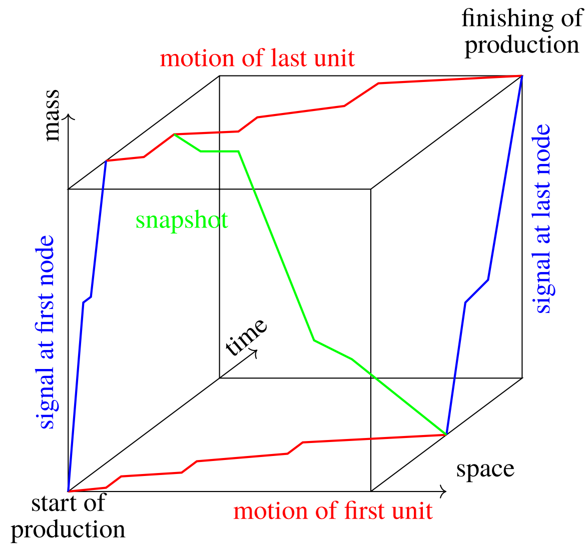


Figure 1: Mass-space-time

Each point (N, l, t) of the flow expresses the statement: "Nth unit passed node l at time t." The production can be put into an enveloping cuboid starting at vertex "nothing done, input node, time zero" and finishing at vertex "all done, output node, last item just finished". Proportions of the cuboid could be associated with economic value of the production — more mass means better, more space means more added value, thus also better, more time means worse. Hence, *value v* could be naïvely expressed as

$$v = \frac{Nl}{t}.$$

The flow can be studied by perpendicular cuts:

- A space-time cut represents a *motion* of one unit through the path.
- A time-mass cut represents a *signal* at one node, i. e. a function which express a number of units passed in time.
- A space-mass cut represents a *snapshot* of overall progress at one moment.

It is interesting that each of the concepts provide some of basic types of simulation.

Thanks to the geometric interpretation we can also mimic steps of production planning:

- a *technological design* of the production network, i. e. the spatial mode of the flow,
- a *parameter specification* of used devices, i. e. an extension of the technological design in a space-time plane,
- *customer orders specification*, i. e. the flow structure along the mass mode, and
- a *schedule of jobs*, i. e. finishing the plan by placing the job boxes to the mass-space-time.

3. CATEGORIC LANGUAGE AND TREE ORGANISATION

Production planning starts with a design of the network, i. e. the *spatial* mode of the flow. It expresses connections among consequent processes or stacks and forms the structural component of a *value stream mapping* chart.

The network could be naturally described in terms of graph theory but we try to do that in a more modern way in terms of category theory. A *category* is an abstract mathematical structure consisting of *objects* and *morphisms*. An object is an analog of a vertex and express a node of the network, i. e. a place or a state that each processed unit passes at an accurate time. A morphism is an analog of an oriented edge (arrow) between two objects and expresses any change of the state. This changes include production processes, transfers, or even stacks (the unit state changes from "deployed" to "picked-up") and we refer to those as *devices*. An electrical circuit is a good analogy: The morphisms/devices are the components, the objects are their connections.

The main benefit of categoric modeling is provided by *composition* of morphisms — once we have two consequent morphisms, then there is a composed morphism. The composition satisfies associativity, i. e. more than two morphisms can be composed "independently on parenthesizing". Thanks to the composition, we can consider a production line as one device instead of a serie of devices.

Since the production network need not consist of a single path, we have to consider extra *tensor* operation $\otimes$ for parallel composition. Categories with such structure are called *monoidal*. The operation is defined both on objects and morphisms and in our setting should be understood just as a formal grouping. The interference of the serial and parallel composition is a subject of axioms and rules of monoidal categories and is thoroughly discussed in Coecke and Paquette (2011) and Baez and Fong (2016). Naïvely speaking, these rules specify which network designs provide equivalent circuits.

The categoric model provides the structure of a graph but adds a *hierarchy* of composed devices. This is useful for organization and control of larger networks — one can see a factory as a network of a small number of large "subfactories" and then "scroll down" to elementary devices. Since the devices can be composed via more equivalent ways, it is upto the designer to select one preferred way of organization. But then we obtain a *tree* of devices and this has a perfect correspondence with the *work breakdown structure* (WBS, see the example in Section 6).

As usual, the category language does not bring some concrete solution but it makes mathematical aspects of manufacturing much clearer. Moreover, the hierarchical approach is suitable also for other modes of the flow — mass and time. Mass is naturally organized by product types, resources, jobs, etc. Each of the divisions can be refined to a tree structure, e. g. product types have variants and colors, customer demands etc. That is, the mass tree evolves through space as the units are rearranged, reordered, regrouped, or renamed. On the other hand, time seems to be universal and there is no need to consider the tensor for parallel grouping. Thus we see it as a tree of time units, e. g. months, weeks, days, shifts, hours, and minutes. The hierarchization of all modes enables to divide the mass-space-time to smaller segments. Hence our simulation algorithm can be performed by a divide and conquer principle and with a high efficiency.

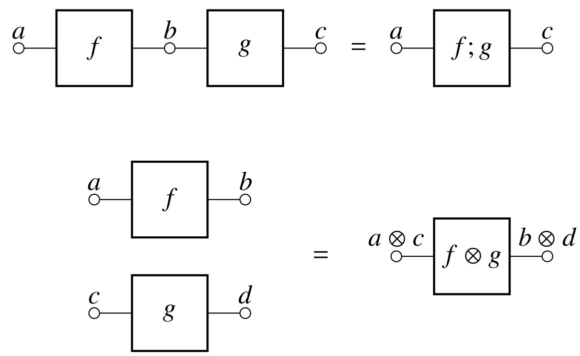


Figure 2: Serial and parallel grouping

4. SIMULATION MODELS

Given the above considerations, now we want to distinguish between three simulation models of production flow characterized by different degrees of regularity. These we call *motion model* (the manufacturing is regular in all aspects), *signal model* (the production is still mass and regular on processes but the processes need not be synchronized) and *snapshot model* (the production is irregular or the number of items is small).

4.1. MOTION MODEL

The motion model is suitable for manufacturing which is regular in all aspects, e. g. that of production lines. There are no substantial differences in the way how all units are produced and they are assumed to spend equal times in processes and stacks. In particular, all the processes share the same cycle time CT. One can simulate a motion on just one unit and copy it to all the space-time cuts with delays in multiples of CT. The motion is determined merely by process times PT_i and idle times IT_i. Since the production is homogeneous in mass and time, simulation advances along the spatial mode and has complexity $O(l)$.

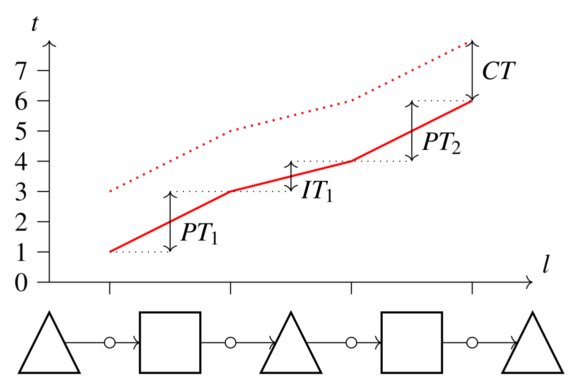


Figure 3: Motion model

Proc. of the Int. Conference on Modeling and Applied Simulation 2018,
ISBN 978-88-85741-07-2; Bruzzone, De Felice, Frydman, Longo, Massei and Solis, Eds.

4.2. SIGNAL MODEL

The signal model is suitable for production which is still mass and regular on processes but the processes need not be synchronized. That is, the processes could operate under different speeds and we need high-capacity stacks to eliminate the differences. Faster processes are finished sooner, their work is compressed to time-bounded *jobs*, and the unused capacity can be offered to another production. The produced units do not share the idle times and thus they have different motions. However, we can still use the regularity on processes and recover the flow from the time-mass cuts. In contrast to the motion model, there is no obvious way how to predict times for the jobs. We assume that these times are a subject of *schedulling* and this provides an extra input of the production plan.

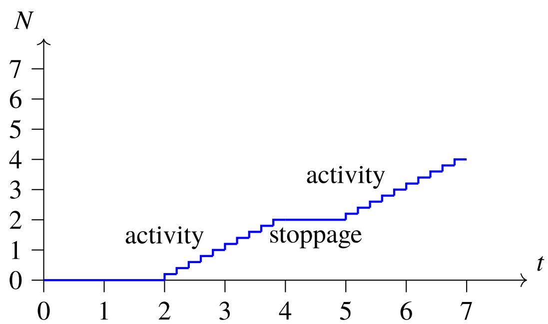


Figure 4: Signal model

Once the schedule is created, we emulate the signals for all nodes of the network. Since the processes work under regular regimes and each node is adjacent to precisely one process, the signal is uniquely determined. This idea uses a different philosophy than traditional push and pull systems. For the sake of simulation, we assume that a process always gets and produces "what it wants" and that a failure is detected on a stack as an overflow or underflow error. Instead of a process message "I want an item and there is no one in the input stack" or "I worked an item and there is no space for it in the output stack", the message is created by a stack in the sense "I am overflowed" or "the process took me an item which I did not have".

In the signal model we construct the flow both from the network design and the schedule of jobs. We need to *validate* the flow, i. e. decide whether it is designed correctly. The problem reduces to checking functionality of stacks: If all the stacks support correctly the surrounding processes during the whole simulated period, then the plan is aproved. If some stack reports a failure, then the plan is rejected, the planner is notified of space and time of the error, and can promptly reconsider.

The signal model is perhaps a new way of simulation and for repetitive processes keeps some advantages. It is still necessary to construct a signal at every node, hence the complexity along the spatial mode is also $O(l)$. But the flow has a regular behaviour at each node, thus relation between mass and time can be described by information compressed to complexity $O(\log N)$. This principle is used also by an iterative validation algorithm which will be described in Section 5. At all, the complexity of signal simulation is $O(l \log N)$.

4.3. SNAPSHOT MODEL

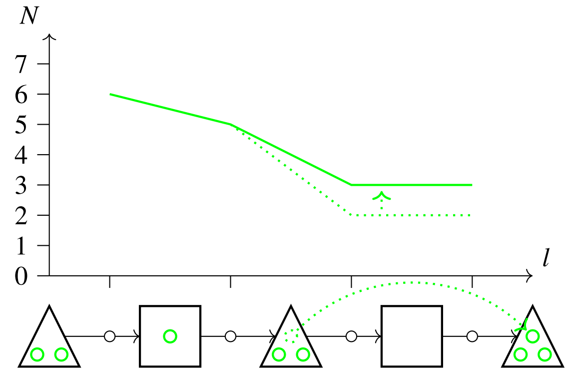


Figure 5: Snapshot model

The snapshot model is a reasonable choice for other types of production, i. e. when we lack regularity or a number of items is small. The tree organisation of production queues is not effective, the items can be simulated individually, and this is done by traditional *discrete-time simulation*. Let us recall that time is divided by *events*, i. e. moments when an item transits a node. In our setting, this happens whenever the item is picked up from a stack or delivered there. A node counter is increased.

5. STACK VALIDATION

Using the "compositional philosophy", more processes in a line could be considered as a single process where the process times or idle times are summed up, while more stacks in a line could be considered as a single stack with a summed capacity. In this way we end up with a simplified network where processes and stacks alternate. As mentioned, we assume that process activity is scheduled, thus we *anticipate* process behavior, and finally, we *create* signals on all input and output nodes of each process. This provides the collection of signals at all nodes.

Since the signals are created individually, one can easily implement changes on devices along the mass mode, e. g. the units in the queue can be reordered, regrouped, or renamed. (For example, a LIFO stack inverts the order the stored units.) The material deploy and shipping could be also modeled as signals. These considerations make the model versatile for multivalued flows, high customization, or supply chain control.

A signal could be written as a formal word in two variables x, o for produced units or gaps (time units), respectively. For example, word *ooooxooxooxoooo* would stand for a signal where four gaps *oooo* (period of doing nothing) are followed by three-time repeated sequence *xoo* (of one unit and two gaps) and finished by two additional gaps *oo*. Such a word could be written in more compact form $o^4(xo^2)^3o^2$ and this provides tree of subwords:

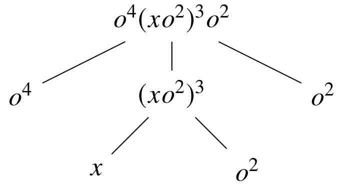


Figure 6: Tree of subwords

Each of the subwords could be approximated just by *volumes* of x and o, i. e., by the number of units, or gaps, respectively, in the subword. E. g., subword $(xo^2)^3$ has x-volume $1 \cdot 3 = 3$ and o-volume $2 \cdot 3 = 6$. The volumes are calculated inductively from leaves to a root of the tree. The geometrical meaning is that the volumes demarcate the smallest rectangle to which the signal segment surely embeds "without looking inside". It represents a *rough knowledge* about the subword and enables to work with it less precisely but much quickly than in a complete detail.

The tree organization of signals is effective whenever the production is regular in the sense that processes work in *repetitive regimes* with known parameters (process times, cycle times). Notice that the regularity may appear also on higher levels of a tree — we can add planned shutdowns, or use product leveling. That is, an already designed pattern of production can be repeated in larger time scales and this increases only a power of a subword, not a complexity of the word.

Once the signals are prepared, we can run a validation algorithm for stacks. In contrast to active and deterministic processes, stacks are passive and behave nondeterministically. By assumption, we know signals on inputs as well as outputs, and we have to decide whether they are balanced that the stock is kept in given bounds. The algorithm performs from roots to leaves, that is, in an opposite way than the preparation of signals. The first iteration is very rough and provides a simple guess about the stock development in time. Further iterations bring more and more precision as we advance to leaves of the signal trees. But the point is that they are performed not on the whole time interval but only on segments where previous iterations did not give decisive answers. That is, whenever the approximation is sufficient to say that the stack surely works or surely fails on a certain interval, the interval is not more tested. In the remaining cases we go to a lower level of the trees and make a more precise approximation.

The classical time simulation works "in a complete detail" along the whole time interval. The iterative validation algorithm performs vertically along the tree structure and improves the guess only when necessary. It can be compared to precision strikes versus carpet bombing. When both the regularity of production and the number of produced items reach an interesting level, the *method should substantially accelerate the simulation*. The complexity depends on a way how the signal words are compressed to the power form and this is what we informally guessed by $\log N$.

6. EXAMPLE

Let us demonstrate the model on a simple example. We consider a network consisting of stacks a, b, c, d, e, f and processes g, h, k, l:

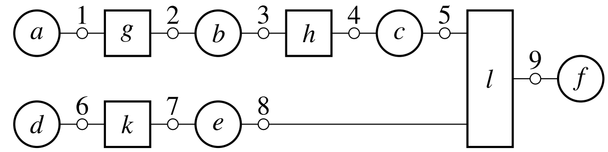


Figure 7: Simple network

The network can be written by many algebraic terms, e. g. $(a \otimes d); (g \otimes k); (b \otimes e); (h \otimes 1); (c \otimes 1); l; f$ or $((a; g; b; h; c) \otimes (d; k; e)); l; f$ where 1 stands for an identity morphism (which "does nothing").

WBS usually specifies only processes and a preferred way of grouping. In the example, we can consider the following one:

Table 1: Work breakdown structure

process id	term
0	$((g; h) \otimes k); l$
1	$(g; h)$
1.1	g
1.2	h
2	k
3	l

Assume that the assembly process l completes 1 unit from the lower branch with a batch of 2 units from the upper branch (creating 1 unit of another kind) and that its process time is 10 and cycle time 4. The other processes are specified as follows:

Table 2: Process parameters

process	process time	cycle time	batch size
g	3	3	1
h	8	2	1
k	80	100	10

Our wish is to produce 50 units, i. e. to increase the stock at f by 50. But we can specify the plan by setting initial and final stocks on all stacks, for example:

Table 3: Stack parameters

stack	range	initial stock	final stock
a	0–200	150	0
b	0–150	10	20
c	0–300	0	40
d	0–100	75	25
e	0–100	0	0
f	0–100	0	50

From this, we calculate numbers of processed units and times:

Table 4: Jobs

process	units	batches	time
g	150	150	450
h	140	140	286
k	50	5	480
l	100/50/50	50	206

Finally, a planner organizes jobs to a schedule:

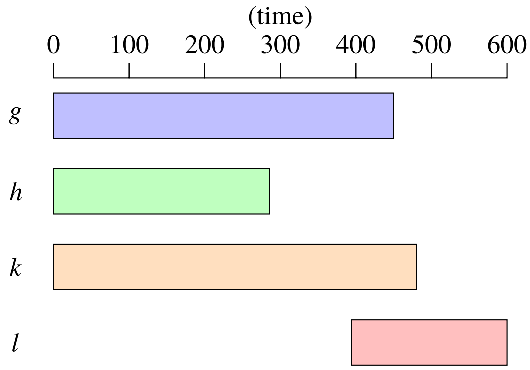


Figure 8: Schedule of jobs

We get signals:

Table 5: Signals

node	word
1	$(xo^3)^{150}o^{150}$
2	$(o^3x)^{150}o^{150}$
3	$(xo^2)^{140}o^{320}$
4	$o^8(xo^2)^{140}o^{312}$
5	$o^{400}(x^2o^4)^{50}$
6	$(x^{10}o^{100})^5o^{100}$
7	$o^{80}(x^{10}o^{100})^5o^{20}$
8	$o^{394}(xo^4)^{50}o^6$
9	$o^{400}(o^4x)^{50}$

Let us demonstrate a validation of stack b: Its input is the signal at node 2, the output is the signal at node 3. The stack signal is calculated as a subtraction of these signals.

The words of the signals at nodes 2 and 3 can be rewritten as $(o^3x)^{150}o^{150} = (o^3x)^{93}o \cdot o^2x(o^3x)^{56} \cdot o^{150}$ and $(xo^2)^{140}o^{320} = (xo^2)^{140} \cdot o^{170} \cdot o^{150}$, respectively to distinguish three periods. In the first period both input and output are active. The output works faster (due to a shorter CT), hence the stock decreases. In the second period only the input is active and the stack is filled. In the third period there is no activity and the stock remains constant.

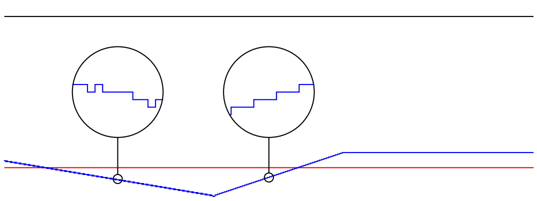


Figure 9: Validation of stack b

The red and black curve are bounds of the stack, the blue curve is the stock signal. It is obvious that the stack fails shortly after the start of production because process h is faster than g and the stock is low. (In practice, h would starve.)

In the first iteration, both words are guessed by tangent lines and we get very rough estimations:

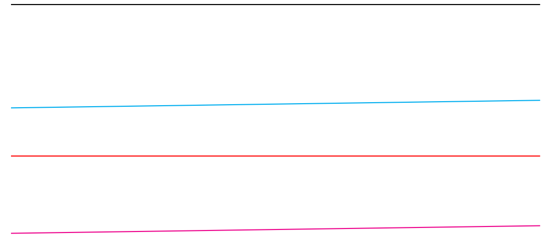


Figure 10: The first iteration of validation

The upper or lower guess is depicted by a cyan or magenta line, respectively. Thus we know that the stack is not overfilled but do not know whether it is overdrawn. This holds for all points of the time interval.

We have to continue with the second iteration where the words are split to subwords $(o^3x)^{150}$ and o^{150}, or $(xo^2)^{140}$ and o^{320}, respectively. This divides the time interval to three periods on which we subtract pairs $((o^3x)^{93}o, (xo^2)^{140}), (o^2x(o^3x)^{56}, o^{170})$, and (o^{150}, o^{150}) as discussed before. We guess them separately by tangent lines and get an approximation much closer to reality:

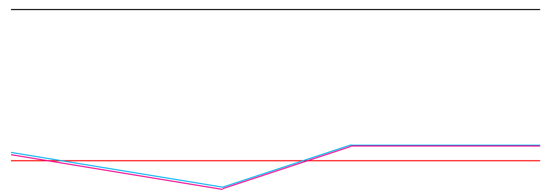


Figure 11: The second iteration of validation

The iteration could continue by a third step on the small areas where the lines cross bounds but we can stop it because we already have an answer for the validation question — the stack surely fails by being overdrawn.

The error is immediately reported and a planner can reconsider about time plans for processes g, h. (Intuitively, he would do better with postponing the job for h.)

In the example, we considered a simple production of one kind of products. However, the model would operate also with heterogeneous flows. In such a case, changeover times would appear in the schedule as certain "obstacle boxes".

7. CONCLUSION AND FUTURE RESEARCH

Simulation methods of production flow in manufacturing are well known for possessing a large variety of objectives and constraints.

The main purpose of the present paper was to preview a categoric methodology for a direct simulation of this task. A categoric model for real time simulation of production flow, which clarifies the way how larger parts can be built from smaller pieces, is proposed.

We present the flow as a certain 2-dimensional manifold through a 3-dimensional mass-space-time state space. We then explain in detail the connection between our categoric language and the description of production flow.

For real time simulation of production flow, the authors suggest using composition of three types of simulation models rather than using common state unified approach. These models represent various degrees of regularity of a production flow. Their complexity can be summarized as in Table 6.

Table 6: Complexity of production modes

model	complexity
motion	$O(l)$
signal	$O(l \log N)$
snapshot	$O(lN)$

In the signal model, which we think is an important model today, it is also crucial to be able to manage and handle the validation of the flow.

This paper presents a *Stack validation algorithm* (SVA) that allows the production engineer to obtain a detailed and real time information about the proposed production flow. The proposed SVA enhances the applicability of traditional approaches to simulation.

We applied our model to an example to demonstrate the model efficiency.

This study makes a contribution in successfully materializing the afore mentioned task by obtaining a test version of a simulation program "VALUE FLOW" of production flow.

Future research potential of the proposed methodology could be significant. One direction for example, is to investigate how the spaces of signals and behaviour of processes can be described in an algebraic manner.

Another important direction of potential research is a development a probabilistic generalisation of the perfect production model. The real production is biased by many random events and a superposition of their effects could result to a very complex random variable. However, the flow of the probabilistic model can still be depicted as a "fuzzy cloud" in the mass-space-time and we believe that a categoric modeling is possible.

Many random events can be encoded to words of signals just by random variables in powers. For example, let $(xo^3)^{100}$ be a perfect signal representing a production cycle of time length 3 repeated 100 times. When the cycle time is not perfect, one can replace it by a random variable Y (e. g. that of normal distribution with mean 3) and this results to word $(xo^Y)^{100}$. When the process makes a waste, one can write $(x^Y o^3)^{100}$ to express that the unit x is either produced or not and here Y is a random variable with a Bernoulli distribution (i. e. a discrete distribution taking non-zero probabilities only at 0 and 1). Finally, $(xo^3)^Y$ would be an example of production where the cycle is perfect but the number of repetitions is random.

Thus the random events could appear on any level of the signal hierarchy and could affect any mode. We currently work on a simplified calculus for their aggregation. Instead of a definite yes/no answer, the resulting model of flow should provide a probabilistic guess of success.

ACKNOWLEDGEMENTS

Research of the first (corresponding) author was supported by the project New approaches to aggregation operators in analysis and processing of data, Nr. 18-06915S by Czech Grant Agency (GAČR). Research of the second author was supported by the project Research and Development of Information System for Value Flow Optimization and Production Planning financed by Ministry of Industry and Trade: project CZ.01.1.02/0.0/0.0/15_019/0001370.

REFERENCES

Abramsky S., Coecke B., 2004. A Categorical Semantics of Quantum Protocols, Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science, pp. 415–425. 17-17 July 2004, Turku (Finland).

Al-Turki O., Ayar T., Yilbas B. S., Sahin A. Z., 2014. Integrated Maintenance Planning in Manufacturing Systems. Cham: Springer International Publishing.

Birkhoff G., von Neumann J., 1936. The logic of quantum mechanics. The Annals of Mathematics, 37 (2), 823–843.

Baez J.C., Fong B., 2016. A compositional framework for passive linear networks. arXiv preprint arXiv:1504.05625. Available from: https://arxiv.org/pdf/1504.05625.

Bonchi F., Sobociński P., Zanasi F., 2017. The Calculus of Signal Flow Diagrams I: Linear relations on streams. Information and Computation 252, 2–29.

Căzănescu V.-E., Stefănescu G., 1990. Towards a new algebraic foundation of flowchart scheme theory. Fundamenta Informaticae 13, 171–210.

Choi B. K., Kang D., 2013. Modeling and Simulation of Discrete-Event Systems. Hoboken, New Jersey: John Wiley & Sons, Inc.

Coecke B., Paquette É. O., 2011. Categories for the practising physicist. In: Coecke B., ed. New Structures for Physics. Berlin: Springer, Lecture Notes in Physics, vol. 813, 173–286.

Meseguer J., Montanari U., 1990. Petri nets are monoids. Information and Computation 88, 105–155.

Negahban A., Smith J., 2014. Simulation for manufacturing system design and operation: Literature review and analysis. Journal of Manufacturing Systems 33 (2), 241–261.

Shanker K., Srinivasulu A., 1989. Some solution methodologies for loading problems in a flexible manufacturing system. International Journal of Production Research 27 (6), 1019–1034.

Tanchoco J., 1994. Material flow systems in manufacturing. Dordrecht: Springer.

AUTHORS BIOGRAPHY

DAVID KRUML

is an assistent professor at Department of Mathematics and Statistics of Masaryk university in Brno, Czech Republic. He recieved his Ph.D. in 2003. He is primarily interested in algebraic quantum structures, category theory, and game theory. He currently works as a key person on project "Research and Development of Information System for Value Flow Optimization and Production Planning" financed by Ministry of Industry and Trade: project CZ.01.1.02/0.0/0.0/15_019/0001370 (2017–2019).

JAN PASEKA

is an associate professor at Department of Mathematics and Statistics of Masaryk university in Brno, Czech Republic. He recieved his Ph.D. in 1990. His professional field of interests includes non-commutative generalization of topology, quantum and many-valued structures. He was a principal investigator of several research projects. He currently supervises at Masaryk university the project "Research and Development of Information System for Value Flow Optimization and Production Planning" financed by Ministry of Industry and Trade: project CZ.01.1.02/0.0/0.0/15_019/0001370 (2017–2019).

DISCRETE EVENT SIMULATION APPLIED TO THE ANALYSIS OF THE CASH-DESKS UTILIZATION IN A SELECTED SHOP OF THE RETAIL CHAIN

Martina Kuncova[(a)], Marketa Skalova[(b)]

[(a),(b)]University of Economics Prague, Czech Republic, Faculty of Informatics and Statistics, Dpt. Of Econometrics

[(a)]martina.kuncova@vse.cz, [(b)]SkalovaMarketa@seznam.cz

ABSTRACT
Simulation model is usually used for the process modelling and analysis of the systems where other mathematical techniques are hardly to use. Discrete event simulation could be applied at those type of processes where the change in the system does not occur continuously but only when a significant event happens. It is typical for various queuing systems analysis. This contribution deals with the application of simulation program SIMUL8 to the analysis of the cash-desk utilization in a selected shop of the retail chain. The main aim is to model the system and to use the simulation model to set the optimal number of cash-desks and test several "what-if" situations. The advantage of SIMUL8 consists of its simplicity and interpretability of achieved results. Our results showed that in case of bigger purchase more than 16 cash desks are necessary to open to achieve less than 10 minutes customers' waiting times.

Keywords: discrete even simulation, cash-desks, retail shop, SIMUL8

1. BACKGROUND AND RELATED WORK

Simulation modeling belongs to the suitable instrument that can be used in the real-world situations to better understand the reality or to make a responsible decision. Simulation can be characterized as a technique for imitation of the real situation, process or activity in order to study the reality or to find a solution to a problem (Banks 1998). Models can be created in various types of software depending on the problem modeled and on the time changes when the simulation is run. When no time sequences are needed, only Monte Carlo simulation application could be efficient especially for iterative evaluation of a deterministic model. But real simulation is usually made via discrete event simulation model or continuous simulation.
Discrete event simulation (DES) could be used in various situations (Goldsman and Goldsman 2015): to model queueing systems (services), to model manufacturing systems or to model inventory systems. In queuing systems such as shops or other services customers arrive randomly and may wait in lines to be served. In a typical queuing system' simulation model the length of the queues or waiting times are analyzed and limited so as not to lose customers (Dlouhy et. al 2011). Not only shops but also helpdesk centers are simulated to allow the improvement of helpdesk services (Manoel, Bouzada and Alencar, 2017). The same situation happens in polling stations where long queues might harm the voting process (Hang Au et al. 2017). In manufacturing systems, where parts or products are processed in various sequences at different stations also the queues of the products are necessary to analyze because of the limited space (Fousek, Kuncova and Fabry 2017). Similar is the case with inventory systems, where random quantities of a certain product are purchased by customers and taken from a store, which influences he policy of ordering products. O'Kane et al. (2000) show the importance of discrete-event simulation for the decisions to increase in total production output. Masood (2006) investigates how to reduce the cycle times and increase in the machine utilization in an automotive plant. Montevechi et al. (2007) show the meaning of simulation experiments representing different scenarios and company strategies.

The model which is well prepared can answer a lot of different question such as "What can happen if we change..?", "Are we able to fulfill our customers' needs?" etc.
In the simple queuing system a mathematical model could be used, such as M/M/1 or M/M/C model, where 1 or c is the number of serving centers and the customers' inter-arrival times and the operating times could be modeled via exponential distribution (Gross et al. 2008). But when other conditions are necessary to model (different probability distributions for the operating times, different inter-arrival times during a day, several rules for the queuing) the usage of the simulation model is more suitable to describe the situation and to find the average queue lengths and the average waiting time.

2. INTRODUCTION INTO THE PROBLEM

As one of the authors partly worked in a retail shop we were asked to analyze the situation of the cash desks to be opened as occasionally large queues arise, resulting in growing dissatisfaction with customers who leave the shop without purchasing. The manager of the store wanted to find out how many cash desks should be opened so that most customers would not wait for more than 10 minutes. The store provided us with a customers' data that are available in their database. One of the requirements was also the processing of this data, as this is not done regularly (for example to create ABC analysis). Based on an agreement with the company's management and according to our experience SIMUL8 software could be used to solve this problem.

The main aim of this contribution is to create a simulation model for a shop in a retail chain to find out how many cash-desks are needed in real life satiation and how the increase in the customers' arrivals influence the length of the queues in front of the cash-desks. Data for the model were taken from the real shop and from the authors' experience based on the work in a retail shop. No data about employees' wages was not provided neither do the other operating costs, so the costs are not part of the analysis.

Our main contribution lies not only in the model creation but also with the description of the processed, data analysis and estimations of the length of the purchase activities. Afterwards the retail store could use the model to test other situations that might appear.

3. MATERIAL AND METHODS

Since we chose SIMUL8 program to create the model with, we first describe the model creation environment and the basic parameters and afterwards the data we worked with.

3.1. SIMUL8 Software

SIMUL8 is a software package designed for Discrete Event Simulation or Process simulation. It has been developed by the American firm SIMUL8 Corporation (www.simul8.com). The software has started to be used in 1994 and every year a new release has come into being with new functions and improved functionality. It allows user to create a visual model of the analyzed system by drawing objects directly on the screen of a computer. SIMUL8 belongs to the simulation software systems that are widely used in industry and available to students (Greasley 2003).

SIMUL8 operates with 6 main parts out of which the model can be developed: Work Item, Work Entry Point, Storage Bin, Work Center, Work Exit Pont, Resource (Concannon et al. 2007; Fousek, Kuncova, and Fabry 2017).

Work Item: dynamic object(s) (customers, products, documents or other entities) that move through the processes and use various resources. Their main properties that can be defined are labels (attributes), image of the item (showed during the animation of the simulation on the screen) and advanced properties (multiple Work Item Types).

Work Entry Point: object that generates Work Items into the simulation model according to the settings (distribution of the inter-arrival times). Other properties that can be used in this object are batching of the Work Items, changing of the Work Items' Label or setting of the following discipline (Routing Out).

Storage Bin: queues or buffers where the Work Items wait before next processes. It is possible to define the capacity of the queue or the shelf life as a time units for the expiration.

Work Center: main object serving for the activity description with definition of the time length (various probabilistic distributions), resources used during the activity, changing the attributes of entities (Label actions) or setting the rules for the previous or following movement of entities (Routing In / Out).

Work Exit Point: object that describes the end of the modeled system in which all the Work Items finish its movement through the model.

Resource: objects that serve for modelling of limited capacities of the workers, material or means of production that are used during the activities.

All objects (except of resources) are link together by connectors that define the sequence of the activities and also the direction of movement of Work Items. The movement is animated in a simulation run which might be helpful to see the flow of items through the system. The suitability of the model can be easily assessed. When the structure of the model is verified, a number of trials can be run under different conditions. Then, the performance of the system can be analyzed statistically. Values of interest may be the average waiting times or utilization of Work Centers and Resources (Shalliker and Rickets 2002).

SIMUL8 can be used for various kinds of simulation models (Concannon et al. 2007). The case studies can be seen also on the website www.simul8.com. Our experience shows that SIMUL8 is easy to learn and easy to use. It can serve not only for the modeling of different services (Dlouhy et al. 2011) including healthcare services and patients treatment (Anderson et al. 2017; Lebcir et al. 2017), but also for the simulation of various production processes (Ficova and Kuncová 2013; Fousek, Kuncova and Fabry 2017), to the improvement of the lean manufacturing system (Omogbai, Salonitis 2016).

3.2. Description of the processes in a shop

The selected shop belongs to the big shops of an international retail chain. Not all data necessary for the model were available and so it was necessary to estimate some operating times. Data are from the first three month of the year 2016. As the customers have the customer card, it is possible to create the ABC analysis. This analysis divides the customers into 3 groups according to their monthly purchase. Type C customers are small customers with up to 10,000 CZK purchases. Shopping from 10,000 to 40,000 CZK are B-type buyers, and buyers over 40,000 CZK are type A customers.

When analyzing customer behavior, our first question was: "How many customers are arriving and when?" To find a answer to this question, the number of visits (minimum, maximum and average) for each ABC group and day had to be set according to the customers data (Table 1). It was also important to determine the percentage of groups, which shows the last line in the first table.

Table 1: ABC analysis - visit per day by day

Day/Avg.	A	B	C	SUM
Monday	200	440	1499	2139
Tuesday	200	449	1692	2341
Wednesday	208	466	1913	2587
Thursday	209	500	2091	2800
Friday	205	492	2456	3192
Saturday	128	373	2242	2743
Sunday	109	329	1678	2116
SUM	1259	3048	13611	17917
%	7%	17%	76%	100%

Subsequently, the distribution of arrivals by time and day within one week was analyzed. This fact is very important because type A and B customers (or big entrepreneurs) go on a working day in the morning, while the members of the group C goes rather in the late afternoon or at the end of the week. The distribution of customers into the selected day periods is described in Table 2.

Table 2: Distribution of customers within one week based on weekday and time

Day /Time	5:00-9:59	10:00-13:59	14:00-18:59	19:00-22:00	SUM
Mon	2.01%	3.78%	4.95%	1.37%	12.10%
Tue	2.25%	4.18%	5.37%	1.45%	13.25%
Wed	2.36%	4.72%	5.84%	1.72%	14.64%
Thu	2.52%	5.00%	6.38%	1.94%	15.84%
Fri	2.49%	5.29%	6.88%	2.01%	16.68%
Sat	2.14%	4.39%	6.26%	1.63%	15.52%
Sun	1.39%	4.31%	4.99%	1.28%	11.98%
SUM	15.15%	32.77%	40.68%	11.40%	100%

There are 16 classic cash desks and 4 self-service cash desks at the shop. Cash desks 56 to 59 are self-service, only one employee supervises all four cash desks. Although there are 16 cash desks available, they are not all open all the time. Cash desks 1, 2, 15 and 16 are opened only for Christmas, when a lot of customers come. Customers with a gold or silver card might take precedence over other customers in cash desks 1 to 8 that is why they prefer these cash desks.

Table 3 shows the preferences of the customers from the group A, B and C from the sets of cash desks point of view. The table shows that most large business owners (A and B group) have gold or silver cards, and therefore they prefer cash desks 3 to 8. On the other hand, C-group customers use end-of-shelves or self-service cash desks.

Table 3: Cash desks selection order for customers from a group A, B, C

Priority of customers	A	B	C
Cash desks 3-8	1	1	3
Cash desks 9-14	2	2	1
Cash desks 56-59	3	3	2

3.3. Building a simulation model

The process of building of a simulation model (Fig.1) can be divided into several steps (Dlouhy et al. 2011; Sharma 2015):

- problem formulation
- objectives settings
- decision about the type of the model
- conceptualization of the problem
- data collection
- software selection
- building a simulation model
- verification and validation of the model
- model testing and change of inputs
- results description
- documents or reports creation

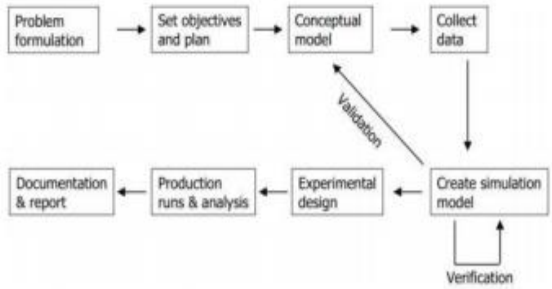


Figure 1: Steps in a process of the simulation model building (Sharma 2015)

The main problem and objective of our analysis is to create a model of the selected shop and to find out the usage of the cash desks in the given shop based on the estimated probability distribution of the time of purchase payment. Data was taken from the shop database but part of the data is not know and must be estimated according to the authors' knowledge and experience. Discrete even simulation was selected as a type of a model that can answer the main questions as the M/M/c and similar queuing models do not take into account special conditions and links between the activities. The conceptual model seems to be easy as there are only few activities in the process: the customers come, choose what to buy, stand in the selected queue, pay for the purchase and leave the shop. For the simulation model SIMUL8 software was chosen. It was necessary to change the data from Table 1 into inter-arrival times to simulate the "customers come" situation (because SIMUL8 does not use the probability distribution of the number of arrivals as the standard setting). As higher usage of the store is visible during the week, the simulation model includes 1 average working day (the weekend days were not modeled) and the inter-arrival times are estimated via triangular distribution (Table 4). The setting of one time period in SIMUL8 is visible in Figure 2.

Table 4: Inter-arrival times of customers (in minutes)

	Day time period	5-10	10-14	14-19	19-22
Triangular distribution	Lower	0.7	0.3	0.28	0.61
	Mode	0.83	0.34	0.34	0.74
	Upper	1.05	0.45	0.34	0.92

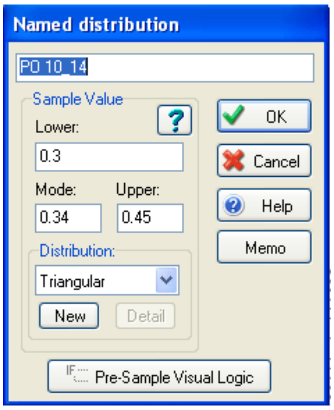


Figure 2: Setting of the inter-arrival times in SIMUL8

The activities "choose what to buy" were separated according to the group of the customers (A, B or C). Table 5 shows the estimated time for this activity (again triangular distribution was applied). Figure 3 shows the setting in SIMUL8.

Table 5: Settings for the length of the purchase activity

Triangular distribution	Lower (minutes)	Mode (minutes)	Upper (minutes)
Group A	15	30	60
Group B	15	25	50
Group C	10	20	40

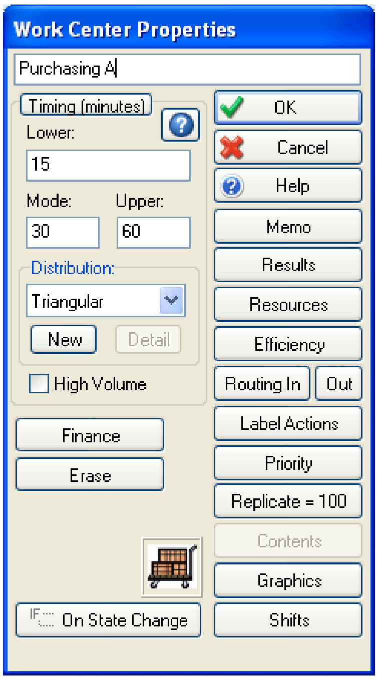


Figure 3: Lenght of the purchase setting (Group A)

For the next part where customers choose the queue the priorities from Table 3 were used and afterwards the discipline of the shortest queue selection was applied. According to this setting in SIMUL8 the entities (customers) prefer only the priority 1 direction if not inaccessible. One of the ways how to prevent access is to set the maximum length of a queue. In the first model the maximum was set to be equal to 6 customers (so the seventh customer tries to find another shortest queue which is shorter), in the second model the maximum queue length in the cash desks 3 to 8 (for A and B groups especially) was set to 3 customers only.

The last important activity in the model is "pay for the purchase". As there were no real data available, it was necessary to estimate the length of this activity. According to our knowledge and experience we have tested two possibilities: the first situation where fewer items are purchased, the second when the purchase is bigger and the payment period is longer (Table 6).

Table 6: Time for the payment at a cash desk

Cash desks	Time for the payment (triangular distribution)			
	Purchase size	Lower (minutes)	Mode (minutes)	Upper (minutes)
3-8	Small	1	3	7
	Big	2	4	10
9-14	Small	1	3	7
	Big	2	4	10
56-59	Small	2	4	8
	Big	3	5	10

4. RESULTS

As it was mentioned above two situations were tested: first when the maximal length of any queue is 6 customers. Model 1 is shown in Figure 4.

The model was verified and validated. First a visual check has been made that the model is as realistic as possible. After entering all the data, the model was tested several times to find out if the selected probability distributions provide the results known from the reality (experience of the author and of the retail shop manager). Small changes were necessary to fit the expected results if we keep the same number of cash desks that are usually opened. Afterwards the model could be used for simulation experiments.

The results showed that the usage of the cash desks 3-8 is much higher than the usage of the cash desks 9-14 (when small purchase size was tested) and also the queuing times in front of the cash desks 3-8 were long (Table 7). Most of the customers were served at the cash desks 3-8 till 10 minutes waiting (Figure 5) but the average time is too big for them to wait (as cash desk 3-8 are preferred by A and B group). The queuing time at self-serviced cash desks 56-59 is realistic (Figure 6 - here no maximum for the queue length was used)

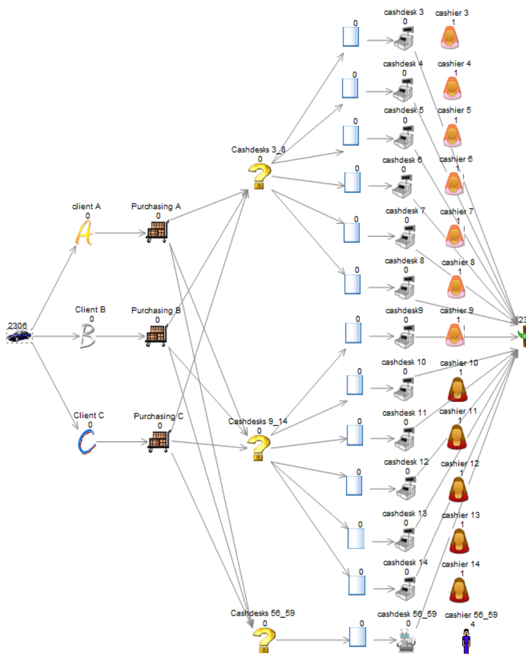


Figure 4: Model 1 in SIMUL8

Table 7: Results (max.6 people in a row, small purchase size)

	Avg. usage	Avg.queuing time (minutes)	Max.queing time (minutes)
Cash desks 3-8	62%	3.80	15.81
Cash desks 9-14	28%	0.22	2.97
Cash desks 56-59	57%	1.57	10.35

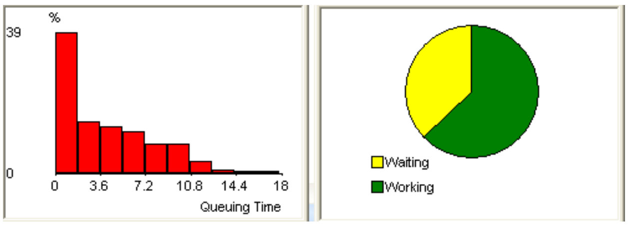


Figure 5: Cash desk 5 – queuing time and utilisation

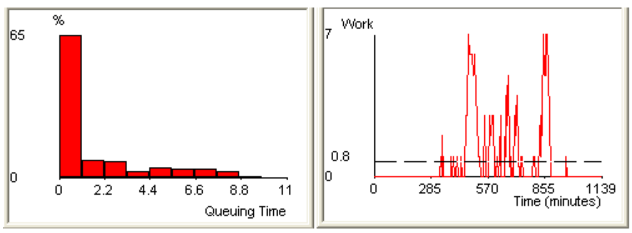


Figure 6: Cash desks 56-59 – queuing time and queue length

According to these results the second model was designed with maximal length of the queues in front of the 3-8 cash desk equal to 3 customers (as customers from group A and B prefer these cash desks but if another cash desk 9-14 has shorter queue they choose this one). If all the queues to cash desks 3-8 are full (3 customers in each) the customer chooses with the probability 80% one of the cash desks 9-14 (with the shortest queue) of goes to the self-serviced cash desks 56-59 (20% probability).

The second model is on Figure 7. The results show that the queuing times better fits the reality (Table 8) and although the queuing time for the self-serviced cash desks is longer, still 99% of customers is served till 10 minutes waiting.

For the small size purchase the experiments with model 2 were taken just to find out how many cash desks could be closed in order to have max. 10 minutes of waiting time for more than 95% of all customers in all queues. The results of these experiments showed that only 1 of the cash desks 3-8 and one from the 9-14 could be closed. In this situation 99% of all customers had the maximum waiting time lower than 10 minutes.

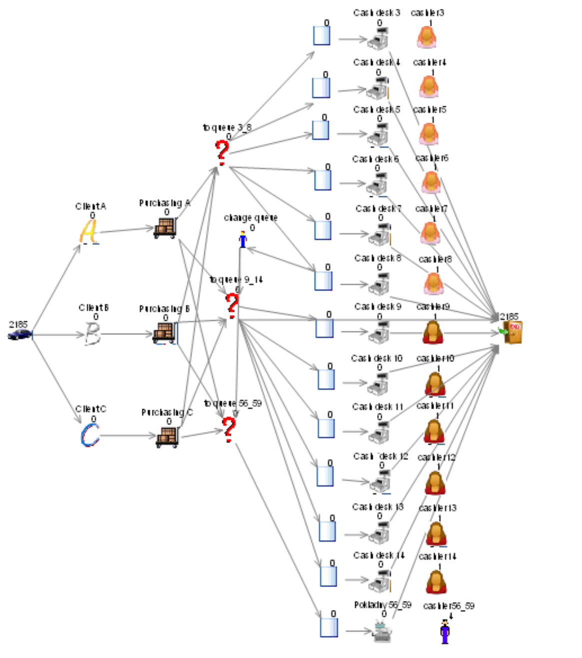


Figure 7: Model 2 with max. 3 customers in a row to cash desks 3-8

Table 8: Results of the model 2 (small purchase size)

	Avg. usage	Avg.queuing time (minutes)	Max.queing time (minutes)
Cash desks 3-8	54%	0.89	3.00
Cash desks 9-14	33%	0.43	5.28
Cash desks 56-59	59%	2.13	12.13

The last part of the analysis was aimed at the increase of the payment times. If the times are higher (triangular distribution 2-4-10 resp. 3-5-10 minutes used), then the waiting time at the cash desks 9-14 starts to be very high (when maximum queue length for the cash desks 3-8 was 3 customers) – see Figure 8, and also the usage of these cash desks is higher (Figure 9). It is necessary to open more cash desks. If the cash desk 1 and 2 are

opened, 65% of customers were waiting till 10 minutes. If also the cash desk 15 is opened the percentage of the customers served till 10 minutes waiting was 85%.
If we release the maximum length at all the queues, then 80% of customers at the cash desks 3-8 and 99 % of customers at other cash desks were waiting max. 10 minutes (Figure 10 and 11).

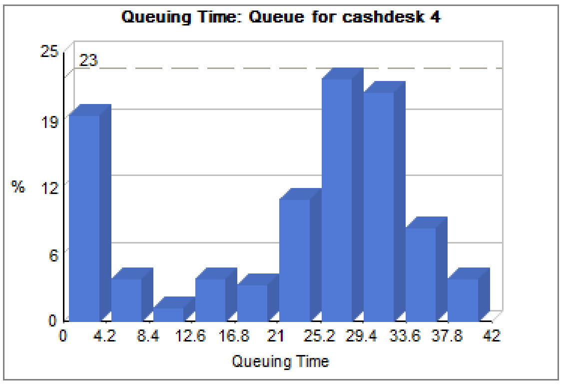


Figure 8: Results – queuing time for customers at cash desk 4 (big purchase, 16 cash desks)

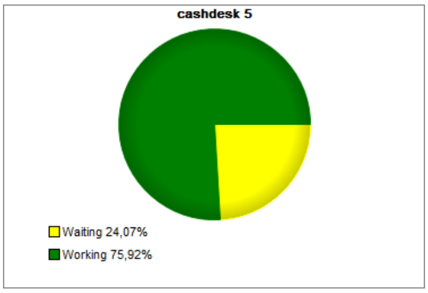


Figure 9: Example of the usage of the cash desk 5 (big purchase)

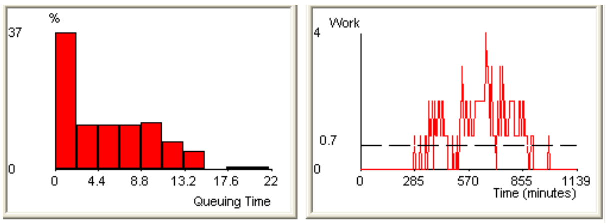


Figure 10: Queuing time and the queue length at cash desks 3 (no max.queue size, big purchase, 19 cash desks available)

Figure 11: Queuing time and the cash desk 9 usage 58% (no max.queue size, big purchase, 19 cash desks available)

5. CONCLUSION

The aim of the paper was to create a simulation model in SIMUL8 software to analyze the situation in one of the shops of the retail chain and to test the number of cash desks necessary to be opened for the customers. As the results of the first model did not fully correspond with the reality we created the second model which seemed to be much more suitable. On this model two situations were tested for the small and big size purchases. The simulation results showed that in case of the small size purchase (and the smaller times for the payment) it is possible to close 2 cash desk and to operate with the 14 cash desks (4 of them self-serviced). But when the purchase size is big and the payment times are longer, the number of cash desks currently open (16 cash desks) is insufficient, and it is necessary to open another 3 (out of 4 possible). In this case the waiting times of more than 85% of all customers are lower than 10 minutes. We can conclude that discrete event simulation is a good tool to analyze queuing systems and SIMUL8 might be suitable software for that kind of modeling and analysis.
The limitations of the model and the results are given by the lack of data about the behavior of customers inside the store (the length of the purchase, the length of payment). To use the model in future we recommend to the managers to try to track these times and then modify data (or the probability distributions) in the model. The model can also be used to analyze operating costs or to investigate expected losses on early exit of customers without making a purchase.

ACKNOWLEDGMENTS

This work was supported by the grant No. F4/57/2017 of the Faculty of Informatics and Statistics, University of Economics Prague.

REFERENCES

Anderson, G., H., Jenkins, P., J., McDonald, D., A., et al., 2017. Cost comparison of orthopaedic fracture pathways using discrete event simulation in a Glasgow hospital. BMJ Open 2017(7). Available from: https://bmjopen.bmj.com/content/7/9/e014509, [accessed 10 July 2018]
Banks, J. 1998. Handbook of Simulation. USA: John Willey & Sons.
Concannon, K., et al., 2007. Simulation Modeling with SIMUL8. Canada:Visual Thinking International.
Dlouhy, M., Fabry, J., Kuncova, M. and Hladik, T., 2011. Simulace ekonomických procesů (in Czech). 2. ed. Brno: Computer Press.
Ficova, P. and Kuncova, M., 2013. Looking for the equilibrium of the shields production system via simulation model. Proceedings of the Conference Modeling and Applied Simulation, pp. 50-56. September 25-27, Athens (Greece).
Fousek, J., Kuncova, M. and Fabry, J., 2017. Discrete event simulation – production model in SIMUL8. Proceedings of

Greasley, A., 2003. Simulation modelling for business. Innovative Business Textbooks. London: Ashgate.
Goldsman D. and Goldsman P., 2015. Discrete-Event Simulation. In: Loper M. (eds) Modeling and Simulation in the Systems Engineering Life Cycle. Simulation Foundations, Methods and Applications. London: Springer, 103-109.
Hang Au, Ch., Xu, Z., Wang, L. and Fung, W.S.L., 2017. Establishing a three-step model of designing the polling stations for shorter queue and smaller waiting time: a case study using computer simulation, Journal of Information Technology Case and Application Research, 19(4), 225-245.
Lebcir, R., Demir, E., Ahmad, R., Vasilakis, Ch., and Southern, D., 2017. A discrete event simulation model to evaluate the use of community services in the treatment of patients with Parkinson's disease in the United Kingdom. BMC Health Services Research - BMC series – open, inclusive and trusted. 17(50). Available from: https://doi.org/10.1186/s12913-017-1994-9, [accessed 10 July 2018]
Manoel, L.G., Bouzada, M.A.C. and Alencar, A.J., 2017. Computer Simulation Improving the IT Helpdesk Problem Management: A Systematic Literature Review. International Business Management 11(1), 68-77.
Masood, S., 2006. Line balancing and simulation of an automated production transfer line, Assembly Automation, 26 (1), 69 – 74.
Montevechi, J.A.B., et al., 2007. Application of design of experiments on the simulation of a process in an automotive industry. WSC'07 Proceedings of the 39th Conference on Winter Simulation, NJ, USA: IEEE Press Piscataway, 1601-1609.
O'Kane, J.F. et al., 2000. Simulation as an essential tool for advanced manufacturing technology problems. Journal of Materials Processing Technology 107/2000, 412 – 424.
Omogbai, O. and Salonitis, K. 2016. Manufacturing system lean improvement design using discrete event simulation. Procedia CIRP – Conference on Manufacturing Systems 57, 195 – 200.
Shalliker, J. and Ricketts, C., 2002. An Introduction to SIMUL8, Release nine. School of Mathematics and Statistics, University of Plymouth.

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 real life problems.

Marketa Skalova: She is a final year student at the University of Economics in Prague. She obtained bachelor's degree in field Mathematical Methods in Economics. As a master's degree she has chosen branch of Econometrics and Operations Research. She also works at the retail shop.

EVALUATION OF RFID TECHNOLOGY FOR REAL TIME DRILL PIPES EFFICIENT MANAGEMENT IN HARSH ENVIRONMENTS: A DISCRETE EVENT SIMULATION MODEL

Yasmina Maizi[(a)], Ygal Bendavid[(b)], Tarik Belarbi[(c)]

[(a)&(b)]École des Sciences de la Gestion – Université du Québec à Montréal, IoT Lab, CRI²GS
[(c)] Mystone Energy

[(a)]maizi.yasmina@uqam.ca, [(b)]bendavid.ygal@uqam.ca, [(c)]tarik.belarbi@mystonenergy.ca

ABSTRACT
Drilling sites operations in the Oil and Gas industry require multiple assets including mud pumps and drill pipes. The later represents one of the most critical type of assets, since pipes are dig down the surface of the earth up to few km. If one of these drill-pipes breaks, the whole site is compromised and such an incident could cost millions of dollars. In the past few years, RFID technology has been used to improve maintenance operations and prevent such incidence to occur. While the potential of the technology seems very interesting in regards to process efficiency, realistic impacts still need to be assessed and validated in the Oil and Gas field. In this paper we present a simulation model as a way to assess the impact of using such a technology for improving drill-pipes management.

Keywords: Simulation, RFID Technology, Oil and Gas, Drill-Pipes.

1. INTRODUCTION

Looking back to the last decades, one can note that Oil and Gas industry went through several up and down cycles due to market volatility. Hence, putting Oil and Gas companies all over the world under high competitive pressure to reduce their operational costs, grow their business and increase their profitability. Even though the pressure is high, these companies are still not taking full advantages of advances in technologies that offer real time monitoring of assets conditions and fields services operations in order to reduce companies' expenses.
Efficient asset management and life cycle management are crucial to any company and especially to Oil and Gas companies where, any asset downtime or failure have heavy operational and hence financial consequences. For example, drill string failure due to asset fatigue is the most common and costly type of drilling failure, which can cost up to millions of dollars to be fixed, and sometimes can lead to the shutdown of the well. Hence, the objective of this paper is to explore the impact of using RFID technology for improving Oil and Gas drilling operations, in particular for drill-pipes management.

The reminder of this paper is organized as follows: in section 2, we review the literature on RFID technology and application in the Oil and Gas Industry. Section 3 describes the business case for real-time drill-pipes efficient management. First, we describe the problem from the field. Second, we define the objectives of the solution with respect to the focal company requests. Third, we evaluate RFID technological options for drill-pipes tracking in harsh environments. Fourth, we propose a simulation model that demonstrates how the proposed RFID technology could satisfy focal company requirements. We conclude this section by explaining the expected outcomes of our simulation model. Finally, section 4 presents our concluding remarks and future work.

2. RFID IN OIL & GAS

2.1. RFID Technology

Although RFID systems are well established in sectors such as retail, the harsh conditions of the Oil and Gas environment makes it more complex to implement and use. Therefore, it is important to describe in further details the technical aspects of RFID systems to better understand their potential and limitations.
An RFID system is a multi-layer system of a hardware and software combination that integrates into an enterprise back end system (Bendavid 2012). The first layer allows the automatic identification of tagged assets (e.g., drill pipes, mud pumps, etc.) by automatically capturing RFID tags IDs. In the last few years, passive Ultra High Frequency (UHF) RFID tags (powered by the reader RF) have increasingly been used in logistics application due to their relative low cost and great performance with a read range up to few meters in metal assets (see **Figure 1**) Companies such as *Confidex, Omni ID, HID or Xerafy* have recently proposed such tags. Semi passive tags equipped with embedded sensors (powered by a battery) can be used in applications where condition of the products need to be captured and traced.

Figure 1: On Metal RFID Tag for Drill-pipes (taken at RFID Journal Live 2018)

On the other hand, active tags are used for Real Time Location Systems (RTLS) applications since they have a read range up to few hundred meters. They are used for yard management application to track mobile asset's movements and location.

Various types of RFID readers can be used to capture the data from RFID tags, including fixed readers, portable handheld readers and more recently reliable vehicle mounted readers. In the Oil and Gas context, fixed RFID readers can be used in specific locations (e.g., pipe ramps to the well, drill-pipe rig) to read tags as the assets enter or leave a specific zone. This data is then transferred to a data management layer using wired/wireless communication network. For asset identification and maintenance purpose, RFID portable handheld readers are mostly used in Oil and Gas.

2.2. RFID Technology in Oil and Gaz

Recent advances in technologies have provided innovative solutions for real time assets tracking and efficient management allowing organizations to insure sustainable operational excellence through real-time analysis of operational conditions and performance indicators. In the last few years, Radio Frequency Identification (RFID) technology has been made available in the utility sector for maintenance and inspection applications (Swedberg 2016), in particular for managing pipes and other equipment for Oil and Gas operators (Swedberg 2013). The technology allows workers to locate and access historical data regarding each asset (e.g., a pipe) before beginning maintenance tasks.

While in the Oil and Gas industry, tracking metal pipes and other downhole equipment is more complex, because of the very harsh operations conditions, some companies have started exploring the technology back in the early 2010 (Redlinger 2011). Hence, in the last few years, some vendors have proposed very robust solutions for asset tracking and management. For example, NOV Tuboscope asset identification and management solution uses a Low Frequency (LF) 125 kHz passive RFID tag (TracTag tag -TracID™) to identify drill-pipes and other equipment used in extreme conditions (NovCom 2008). Its small physical dimensions enable mounting the tag in downhole equipment by including it in a slot without jeopardizing equipment integrity. In term of process, the tag is read by a stationary RFID reader in close proximity or with a handheld reader at a maximum distance of 10 cm. Leaving the operators with a small reading distance that can considerably affect the performance of operations, especially in the context of drilling. More recently, other types of technology have been made available, including reading tag from further distances and allowing to revise the actual maintenance process. In 2015, Xerafy, a Hong Kong based company introduced a metal passive UHF RFID Tag (Xplorer), ATEX-certified for downhole pipe identification in the Oil and Gas industry (Xerafy 2017). Like others, the tag can be embedded in downhole tubulars such as drill-pipes (see **Figure 2**). According to the company, the tag has a read range up to five feet with a fixed reader compared to previous LF and HF RFID downhole tags which can only be read in close proximity. For instance, pipes can be identified during drilling operations/tripping, the tag can be read by a fixed reader (antenna) or by an operator using a handheld reader as the pipe is rotating.

Figure2: On Metal RFID Tag in a Drill-Pipe (Xerafy, 2017)

Although real-time asset tracking and management in field services for Oil and Gas industry are a well-known issue, addressing it remains a considerable challenge. Moreover, the effectiveness and efficiency of the previous cited solutions are all claimed by the vendors and are not estimated under a neutral basis. Apart from the figures provided by the vendors, it is very difficult to estimate properly the impact of using such technologies for real-time asset management.

More recently, researchers at Cal Poly (California Polytechnic State University) have found that a drone equipped with passive UHF RFID reader can read tags attached to steel drills/utility pipes at an accuracy rate of 95-100 % at a distance of 8-12 feet (Swedberg, 2018).

3. BUSINESS CASE FOR REAL-TIME DRILL-PIPES MANAGEMENT

3.1. Context

The starting point of this research work was a request from a private Canadian company (Mystone Energy) that specializes in the supply of equipment and parts used in all stages of drilling and well services. They requested from us a tracking solution for localization, inventory and operational purposes, that will be offered to their clients. This solution should be appropriate for equipment and accessories subject to extreme working conditions, in construction, mining and Oil and Gas field operations. For the purpose of this paper, we focus on a tracking solution for typical tubular equipment that are used in drilling operations in the Oil & Gas field.

Drilling operations management begins when the geographical coordinates of the wells to be drilled are identified. Once these coordinates are known, a list of wells to be drilled is built and the drilling rigs are moved from one drilling site to another. Whenever a drilling rig moves from one location to another, it brings all equipment with it, including all the drill-pipes. This equipment is stored in metal boxes and transported by truck from one site to another. See **Figure 3** for a typical drill site and **Figure 4** for typical drilling equipment, stored in the rig area and ready to be transported.

Figure 3: Typical Drilling Site

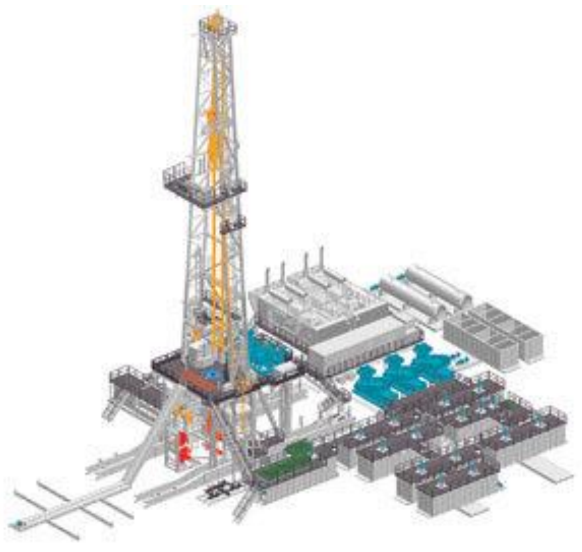


Figure 4: Typical Drilling Rig Site Layout

Each drilling rig has a quantity of drill-pipes for its operations. For this project, we are talking about a total of 140 drilling devices requiring 8 000 to 9 000 meters of drill-pipes each. On average, about a hundred of drilling rigs work in parallel. **Figure 5** displays drill-pipes that are required and stored in a typical drilling site.

Figure 5: Drill-Pipes on a Rig

Currently, tubular equipment maintenance is done with respect to API 7G and DS1 standards. According to these standards, the tubulars are classified and the damaged threads are repaired. For drill-pipes used under normal conditions, the control is done every 1 500 hours of use whereas for those used under severe drilling conditions, the control is done every 1 000 hours.

The inventory of drill-pipes is done manually by identifying and counting all drill pipes present on a rig. The identification of drill-pipes is a big issue. Drill-pipes are sent to the drilling site with an inspection report containing the manufacturers' serial numbers or an internal number given by the organisation in charge of the inspection. The drilling site operations manager keeps a manual track of all drill-pipes identification numbers. However, due to several frictions with the abrasive wells, drill-pipes tend to damage and identification numbers are erased during the ascent phase. Quite often, this leads to several problems, especially when the operations manager keeps several reports from different inspection groups. An attempt to a solution was provided to resolve this problem – using a special device for punching numbers on drill-pipes. But unfortunately, without any success. The identification numbers always tend to fade.

3.2. RFID Process for Drill-Pipes

We propose to investigate the use of RFID technology to identify and track tubular equipment and to link all relevant operational information to a particular piece of equipment. With the help of this technology, each piece of equipment will be scanned (i) at the picking location, ii) before entering the well, iii) when it comes out of the well, iv) when it is stored and v) when it undergoes maintenance tests. Any information produced as a result of the maintenance tests will be attached to each tracked

Proc. of the Int. Conference on Modeling and Applied Simulation 2018,
ISBN 978-88-85741-07-2; Bruzzone, De Felice, Frydman, Longo, Massei and Solis, Eds.

asset and available through the integrated information system. The equipment condition will be traced continuously so as to avoid any possible accident upstream. The reading of the tubular at the entrance of the well will allow a validation of its condition and thus avoid accidents that can lead to important costs and significant downtimes.

3.3. Simulation Model

The focal company requested tracking solutions for one of its clients located in North Africa. This client is ranked as the 4th Gas producer in the OPEP and 10th Gas producer worldwide. We are presently in the data collection phase in order to build a simulation model that will help us evaluate the impact of using RFID technology for real-time drill-pipes efficient management. The data concerns 148 wells and 140 drilling equipment. At this stage of the research, we are conducting a preliminary pilot study that evaluates the advantages of using tracking solutions on a specific site (Hassi-Messaoud). Hereby, we describe the steps used to conduct this preliminary study. These steps are illustrated in **Figure 6**: i) Choose the appropriate RFID technology. ii) Tag placement to incorporate the RFID tag into existing drill-pipes. iii) Map the main drilling operations with the corresponding available data for this specific region. iv) Build a base-case simulation model that mimics the present mode of operations using ARENA software. The base case model includes the actual manual inventory management process. v) Verify and validate the base-case model with statistics provided by the focal company. vi) Design the simulation model for the future mode of operations. This model includes the appropriate RFID technology for real-time drill-pipes tracking. vii) Based on the simulation models, compare the future mode of operation to the actual mode of operations.

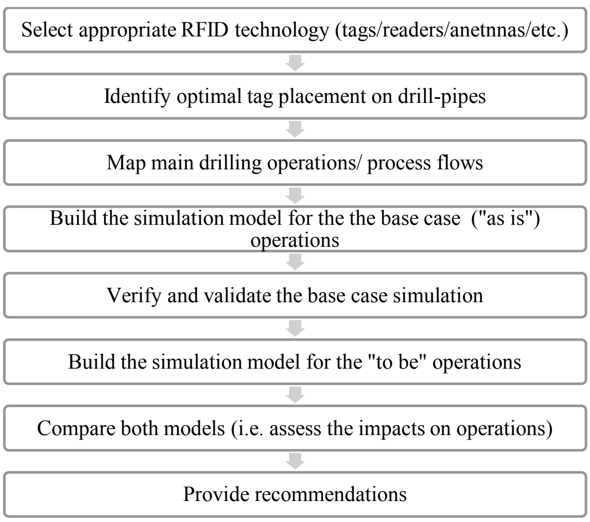


Figure 6: Research Methodology

3.3.1. Drilling Operations Process

The simulation models developed represent the drilling operations process considering two states. The first one is the present mode of operations where there is no real-time tracking of drill-pipes. The second one represents the future mode of operation, where RFID tags are integrated to drill-pipes. The drilling operations process can be summarized as follows: The main company (wells owner) approves the drilling contract and sends a drilling request to the contractor. Once a rig is available, the latter is moved either from the yard or from a previous drilling site with its lot of 8 000 to 9 000 meters of drill-pipes. Once the rig reaches the drilling sites, it is fixed and prepared for drilling operations. The site owner sends an inspecting company to inspect drilling conditions as well as general state of drilling equipment, including drill-pipes. The inspection can either succeed or fail. If the latter fails due to inappropriate drill-pipes state (this is due to inappropriate maintenance reporting as well as inappropriate drill-pipes identification and tracking information), the contractor has to either replace identified drill-pipes, or in the worst case, replace the complete lot. In both cases, these inspection failures will imply drilling operations to be delayed from 3 days to 45 days, incurring costs ranging from USD$ 30,000 to USD$ 1,680,000. Once the inspection is approved and the drilling operations start, it is very frequent to see an interruption of these operations due to drill-pipes failures. Again, these failures are due to inappropriate drill-pipes identification and tracking information as well as inappropriate maintenance information. In this case, the failures can imply delays ranging from 8 days to 18 months, incurring costs ranging from USD$ 528,000 to USD$ 45,000,000. It is obvious that an appropriate tracking and real-time management is critical to reduce these downtimes costs. **Figure 7** displays the drilling operations process.

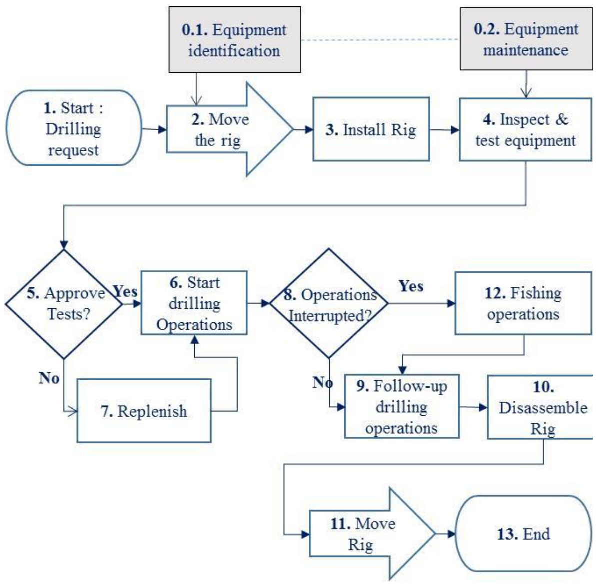


Figure 7: Drilling Operations Process

3.4. Models Assumptions

Based on discussion conducted with Mystone Energy, for this preliminary study, the following assumptions were made:

i) we consider only one contractor that can drill around 15 wells per year (in the best conditions);

ii) we consider only drill-pipes and heavy weights equipment downtimes;

iii) contractors can use two drilling rigs at the same time in the considered region.

3.5. Models Outcomes

We first built the baseline model to analyze the present mode of operations. After validation with the focal company, we identified the two following output parameters: (i) total drilling time and (ii) total number of drilled wells. The model was run for a period of three years and replicated 100 of times. **Table 1.** shows that the output parameters of the present mode of operations model are not different from collected data.

Table 1: Output Parameters Comparison

	Collected data	Simulation model outputs
Number of drilled wells	15	15
Total drilling time (days)	76	76.23

After validating the present mode of operations (PMO) simulation model, we built the future mode of operations (FMO) model by integrating RFID tags into drill-pipes in order to automatically identify and track them. The analysis of the simulation model results showed that the integration of such a technology reduces considerably the drill-pipes downtimes delays. Indeed, with the help of the RFID technology, we are able to detect out of service or inappropriate drill-pipes before their failure in the well. Actually, we are able to detect them, in the worst case at the entry of the well. The simulation results show a reduction of 74,68% of total delays due to inspection refusal and a reduction of 84.42% of total delays due to drill-pipes failures within the well. This reduction of failures is possible due to the real-time tracking and management of these equipment. **Table 2.** displays these results.

Table 2: Simulation Results

Total delays (days)	PMO	FMO	Improvement
Inspection refusals	295.462	74.807	74.681 %
Failures within the well	3857.591	600.94598	84.422%

3.6. Future Developments

In this preliminary study, we wanted to show that using RFID technology to identify and track drill-pipes could considerably reduce operations costs by detecting out of service drill-pipes before their failure in the well. Actually, our purpose was to detect them, in the worst case at the entry of the well. Moreover, we aim to show that an appropriate real-time monitoring of operations and assets via appropriate technologies will ensure a continuous improvement of operational processes and achieve sustainable operational excellence in the field of Oil and Gas. The next step for this research work is to i) first consider the complete region with 148 wells and ii) evaluate the impact of introducing such a technology for identifying and tracking different rig drilling equipment as well as equipment present on the rig structure but not used in the wells.

4. CONCLUSION

As discussed in this paper, RFID technologies have been widely used across industries and have enabled a variety of application. However, in the Oil and Gas industry, the implementation of RFID technology still represents a considerable challenge due to the harsh environments conditions where operations are managed. The objective of this research was to explore the realistic potential of RFID technology for improving drill-pipes management. In this paper, we proposed the frame of a simulation model for assessing and validating the real impact of such a technology. We tested this frame on a pilot study that considers a small drilling region. The results of this pilot study showed that integrating RFID technology for real-time drill-pipes management can considerably reduce downtimes delays and consequently reduce the operational costs. The next step of this research is to complete the data gathering phase for the remaining regions in order to include all the 148 wells, build the simulations models and run different scenarios to evaluate the efficiency of RFID enabled maintenance policies using real-time operational data. We believe that the implementation of such a solution will certainly improve drill-pipes and heavy weights equipment daily management. However, the validation of the success of such an implementation will be achieved once the results of the simulation runs will be completed providing relative KPIs and ROI of such an implementation on the 148 existing wells to drill.

REFERENCES

Bendavid Y., 2012. Positioning RFID Technologies in the Enterprise Information Systems Portfolio: a Case in Supply Chain Management. International Journal of Automated Identification Technology, Vol. 4, No 1, pp. 11-24.

NovCom, 2008. Rugged and Reliable at 4000F available from: https://www.nov.com/stories/tracid.aspx

Redlinger T., 2011. RFID drill pipe heads offshore. Offshore Magazine. Available from : https://www.offshore-mag.com/articles/print/volume-71/issue-11/drilling-completion/rfid-drill-pipe-heads-offshore.html

Swedberg C., 2018. Drone RFID Read Rate Hovers Near 100 Percent in Oil Fields. RFID Journal. Available from: http://www.rfidjournal.com/articles/view?17582

Swedberg C., 2016. Utility Company Uses RFID to Identify Underground Pipes, Wires. RFID Journal. Available from: http://www.rfidjournal.com/articles/view?15188

Swedberg C., 2013. EchoRFID Offers Views Into Buried Oil and Gas or Utility Pipes. RFID Journal. Available from: http://www.rfidjournal.com/articles/view?10751

Xerafy., 2017. Xerafy Introduces First Embedded Metal RFID Tag for Oil & Gas Downhole Pipe Identification. Available from: http://www.xrefay.com

AUTHORS BIOGRAPHY

Dr. Ygal Bendavid is the director of the IoT Lab. and a professor in the Department of Management and Technology/School of Management (ESG) at the Université du Québec à Montréal (UQAM). Dr. Bendavid is a founder of RFID Academia. He holds M.Sc. and Ph.D. degrees in industrial engineering from the École Polytechnique de Montréal. As a specialist in RFID for supply chain management, he is a frequent presenter at RFID Journal LIVE! Conferences.

Dr. Yasmina Maïzi is an associate professor at the Management and Technology Department / School of Management (ESG) at the Université du Québec à Montréal (UQAM). She holds a M.Sc. and Ph.D. degrees in industrial engineering from Conservatoire National des Arts et Métiers (France). She has an extensive professional experience in simulation, data analysis and optimization. She develops complex simulation models to analyze improve or redesign systems and processes for various industry sectors.

Tarik Belarbi has 25 years of working experience in different fields of Oil and Gas sector. He is the chairman of Mystone energy, a company that supplies equipment and parts necessary for the oil and gas industry, in all stages of drilling and well services.

SIMULATION IMPROVES SERVICE AND RESOURCE ALLOCATION AT A SALON

Ganapathi Baliada Ramesh[(a)], Brittany Harju[(a)], Daniel Scipione[(a)], Kristina Vujic[(a)] ,Edward J Williams[(a)]

[(a)]Business Analytics, College of Business, University of Michigan – Dearborn, Dearborn, MI, USA

[(a)]gramesh@umich.edu, [(a)]bharju@umich.edu, [(a)]dscip@umich.edu [(a)]kvujic@umich.edu, [(a)]williame@umich.edu

ABSTRACT

Simulation, historically first used in manufacturing industries, has steadily and deservedly expanded its reach into helping service businesses evaluate and implement solutions to problems such as slow service, overutilized or underutilized resources, suboptimal scheduling, and inefficient workflow. Such service industries have included banks, retail stores, hospitals and clinics, hotels, call centers, and credit unions. In the present work, the authors used discrete-event process simulation to analyze and resolve such problems at a salon. This salon provides a complex mix of services such as haircuts and/or hair coloring, massages, manicures and pedicures, and other spa services.

This simulation modeling and analysis project successfully addressed issues such as size of the waiting area, asymmetry of utilizations across various resources, ways of handling schedules and arrival rates differing by days of the week, and best allocation of priorities among customers and tasks.

Keywords: discrete-event process simulation, work measurement, resource allocation, decision support system

1. INTRODUCTION

Historically, discrete-event process simulation was first and very extensively used to improve manufacturing operations, and extensions to warehousing and supply chain operations (e.g., stochastic optimization of a large-scale inventory-routing problem, as documented in (Lefever, Aghezzaf, and Hadj-Hamou 2018)) naturally followed. In recent years, simulation has repeatedly proved its ability to improve service operations, such as those of banks, credit unions, public transport (e.g., the work of (Otamendi and Pastor 2006) to improve bus stations), hospitals and clinics, hotels, call centers, and retail stores (e.g., the work of (Williams, Karaki, and Lammers 2002)). These improvements often arise through discovery and implementation of better scheduling policies, improved allocation of resources, redesign of process flow, or more effective allocation of typically scarce capital to fund investments such as new equipment. For example, (Baskaran, Bargiela, and Qu 2013) applied simulation to the effective scheduling of nurses. Granting of consumer credit was made faster and more efficient through the work of (Chen et al. 2004). Likewise, (Sivaramakrishnan et al. 2016) describes the successful use of simulation to improve operations at a highly specialized takeout restaurant.

The simulation study described in this paper sought improvements to a hair salon and spa, located in a suburban city in the Great Lakes region of the United States. This salon is already a highly successful "small business enterprise," and its clientele is expanding. The wide-ranging concerns of its management included size of its waiting area, uneven resource utilization (over-utilization of some resources coupled with under-utilization of others), the understandable disappointment and annoyance of would-be customers who had to be turned away, plus typical queueing performance metrics such as average and maximum time in queues. Furthermore, the "color bar" (a small work area of the salon, equipped with a sink, where the stylists go to mix the color that will be applied to the client's hair) was already notorious as a point of congestion, since nearly half of all customers' requests require that a worker spend preparatory time there. As is now rarely the case in large companies, but quite often the case in small independently owned enterprises, the salon owner and manager was unfamiliar with process simulation and its capabilities. When approached by a member of the project team, she was thus delighted to hear that "simulation will let us see the effects of making a change by looking at a computer, rather than by spending money and time making a highly committal and disruptive change that might not work." The University of Michigan – Dearborn's College of Business actively collaborates with local enterprises, as was done in this case study. Such collaboration helps business students understand ongoing enterprise concerns, a major advantage of teaching discrete-event simulation to business students (Ståhl 1996).

In the following sections, we (1) present an overview of the salon's operations, (2) describe the collection and analysis of input data, (3) discuss the building, verification, and validation of the simulation model, (4) highlight key results from this model obtained by output analysis, and (5) present conclusions, recommendations to the salon's management, and candidates for possible future work.

2. OVERVIEW OF SALON OPERATIONS

The salon under study provides a wide variety of services, and therefore is divided into two major operational areas. The hair salon provides haircuts, hair coloring, and hair styling. The spa provides massages, facials, manicures, and pedicures. A small minority of customers, on one visit, may receive multiple services from either or both parts of these areas (a possibility not addressed in this study, except for the scenario in which a customer who has received hair styling may or may not elect to visit hair coloring). Upon appointment request several months in advance, the business will reserve its facilities for use by an entire bridal party. Management of this type of business involves numerous knotty complexities – financial, operational, and interpersonal, as extensively documented by (Tezak and Folawn 2011). For example, the workers (more than twenty) have a wide range of skill sets: hair stylists (the majority), nail technicians, spa therapists, skin care specialists, masseuses, eyelash extension specialists, and airbrush makeup specialists. Management was particularly interested in salon performance metrics on weekends: The salon is open 8am-4pm Saturdays and 11am-5pm Sundays.

Customers arrive originally at a reception area, and then are directed to the waiting room. Depending on the service the customer wishes to receive, either a stylist or a spa worker will come to the waiting room, greet the customer, and escort him or her (a large majority of the customers are women) to the appropriate point in the styling area or the spa. After completion of the styling or spa service, the customer returns, without escort, to the reception area. There, before exiting, the customer will pay the bill and perhaps book a future appointment. As this study began, the manager felt herself under duress to decide quickly "Should the waiting area be expanded further or scaled back (it had recently been expanded), and how will such a decision impact both operational areas?"

3. DATA COLLECTION AND ANALYSIS

On becoming acquainted with the capabilities of discrete-event process simulation, the salon manager immediately and eagerly anticipated opportunities to improve important system performance metrics, such as:

1. Reducing the percent of clients who, upon seeing a nearly full waiting room, balk;
2. Reducing the frequency with which salon workers must stay late to finish services to their customers, and the length of time they must stay for this reason;
3. Reducing the mean and the standard deviation of the times clients receiving a specified treatment stay in the salon (customers appreciate predictability relative to their other plans for the day);
4. Keeping the utilization of various salon workers in the 70%-80% range.

Members of the analysis team, taking care *not* to instigate the Hawthorne effect (Kroemer and Grandjean 1997), visited the spa to unobtrusively observe its operations. Very remarkably and commendably, the manager had numerous data readily available. These data included schedules, varying by day of week, for the styling workers, the spa workers, and the receptionist; arrival rate of customers by hour of the business day, and a detailed breakdown of percentage of customers desiring each of eight categories of service, as shown in Table 1. Overall, three-quarters of customers visit the salon and one-quarter visit the spa.

Table 1: Percent of Customers Desiring Services

Service Requested	% of Customers
Salon	(aggregate ¾)
Hair coloring only	15
Haircut only	18¾
Haircut and hair coloring	33¾
Miscellaneous salon service	7½
Spa	(aggregate ¼)
Manicure	5
Pedicure	8¾
Massage	5
Miscellaneous spa service	6¼

Additional data included service times for all the pertinent service areas (some often shorter than five minutes, and others as long as an hour and half, and highly variable), and a blueprint of the salon floor plan. Using this blueprint to determine distances, plus a typical walking speed of 1.4 meters per second (not treated as variable across customers), permitted calculation of transit times (not negligible) between various locations within the salon. Data was also provided concerning the possibility of balking (Bhat 2015). Specifically, prospective customers typically balked if five or more people were already in the waiting area. Furthermore, the reception desk personnel would, as diplomatically as possible, "balk" customers whose requested service would clearly last well past closing time. This latter situation is very rare, partly because arrivals are by appointment, not "walk-in," and partly because if operations in the salon become backlogged, the reception desk personnel, being apprised of this, will telephone the customer ahead of appointment time to give the disappointing news.

Each of the service-time data sets was analyzed using the specialized distribution-fitting software tools Stat::Fit® (Leemis 2002) and @RISK® (Clemen and Reilly 2014). Upon viewing the histograms of these distributions and receiving an explanation of the three parameters of a triangular distribution (minimum, mode, maximum), the salon manager felt comfortable with their use – a valuable step toward model credibility.

4. MODEL DEVELOPMENT, VERIFICATION, AND VALIDATION

Members of the project team concurred in the choice of the Simio® software [SIMulation with Intelligent Objects] (Prochaska and Thiesing 2017, Smith, Sturrock, and Kelton 2017) to construct a model of the salon's operations. Simio® provides constructs such as the Server (to model, for example, the pedicure station), the Worker (who can both escort a customer to the pedicure station and then perform the pedicure, and the Resource (e.g., the shampooing apparatus required to shampoo a customer's hair). Sequence Tables allow each Entity (customer) to visit the proper locations in the correct order, based on the type of customer (see Table 1). Virtually no incremental work is required to create a helpful animation. A screen shot of the model appears as Figure 1 in Appendix A, and one of the model's internal representation of customer sequences appears as Figure 2 in Appendix B.

At the client's explicit request, the project team built two separate models (the second largely a clone of the first), one for Saturday operations and one for Sunday operations.

An extensive array of techniques are available for model verification and validation (Sargent 2004). Among them, the analysis team undertook all of the following:

1. Temporary elimination of all randomness from the model (replacing each probability distribution by its mean) and checking results arithmetically;
2. Allowing only one entity (customer) to enter the model and checking the logical route it followed (this was done for each of the eight customer "types" in Table 1);
3. Directional analysis (Radosiński 1996), in this context, increasing customer arrival frequency and checking that queue lengths and times in queues increase as expected;
4. Executing the model time-step by time-step while closely observing the animation;
5. Using actual historical data on customer types and their arrival times (obtained from the previous week) and checking the model's predicted performance metrics against recollections of that day.

As is usual, these techniques exposed several modeling errors; after correction of those errors, model results closely (within 6%) matched historical observations. At this point the salon manager expressed confidence in the model – thus the model achieved credibility.

5. EXPERIMENTATION AND RESULTS

The salon is clearly best represented as a terminating (as opposed to a steady-state) system, opening at a specified hour each day and running until at least closing time (typically a little beyond, to allow the workers time to complete services to the last few customers). Therefore, Saturday runs lasted just over eight hours and Sunday runs just over six hours. In addition to the "base case" representing current salon operations as used for verification and validation, five additional scenarios were studied. These scenarios explored the effects of potential changes to the work schedules of the stylists, the seating capacity of the waiting area, and the capacity of the color bar. Each of these scenarios was run for ten replications, the team having verified that this number of replications would generate 95% confidence intervals, sufficiently narrow to distinguish among alternative proposals clearly (Currie and Cheng 2016).

The salon manager asked three specific questions relative to the simulation results and the scenarios chosen:

1. Should the salon increase the capacity of the waiting area beyond the current value of six customers?
2. Should the salon rearrange the scheduling of the stylists to have more than two (the current policy) stylists on duty near the end of each business day – even slightly beyond closing time?
3. Should the salon increase the capacity (i.e., the number of concurrent mixes of hair coloring possible) of the color bar?

The experimentation with the model and these scenarios provided the following unequivocal answers, obtained by comparing the results of the various scenarios via hypothesis tests and/or one- or two-way analysis of variance [ANOVA]:

1. No, the current waiting area capacity of six was completely adequate.
2. No, asking a third stylist to be available up to and just beyond closing time would be of no incremental benefit.
3. Yes, increasing the capacity of the color bar from three to six concurrent hair-coloring preparations would provide significant performance improvements listed below.

These answers were of high value as follows:

1. *Not* increasing the capacity of the waiting area avoided disruption and expense, and would leave space available so a "coffee and tea" area could be moved there – such an area being an amenity competitors had recently begun offering, and hence clients were coming to expect.
2. Not requiring additional stylists to be available late in the business day would avoid lowering morale among the stylists.
3. Moving the "coffee and tea" area to be adjacent to the waiting area will create space to expand the color bar. The doubling (from three to six) of color bar simultaneous capacity will decrease time a client waits in the styling chair by 19% and will decrease time a stylist stands idle waiting to work at the color bar by 27%. From the psychological viewpoint of a client, time spent waiting in the styling chair is "duller" than time spent in the waiting area: During the latter waiting time, a client can conveniently read a magazine, sip coffee or tea, or chat casually with other clients.

6. CONCLUSIONS AND FUTURE WORK

The salon manager, having become acquainted with simulation as a consequence of this study, is highly pleased with the usefulness of the results, especially since (1) the duration of the study was less than three months and (2) the study never disrupted salon operations in the slightest. She plans a prompt increase in the color bar capacity.

Plans for further work include the following:

1. "Cloning" either of these models to create a "Monday through Friday" model.
2. Adding downtimes to the simulation model to drive improvements in contingency planning.
3. Allowing modeling of more general possibilities of clients desiring more than one service.
4. Investigating whether time required at the reception area upon arrival is indeed independent (as currently assumed) of the type of service to be provided. (Interestingly, in view of the current model structure, which makes extensive use of internal data tables, revoking this assumption will require only an updated table entry, but no change to underlying model logic.) Likewise, reception area service time at departure (e.g., paying the bill and scheduling a future appointment) may differ from service time upon arrival.
5. Adding the ability to model different walking speeds – currently, all movement is 1.4 meters per second. It is plausible that stylists walk faster when not escorting a client, and that clients themselves have different walking speeds.
6. Investigating the degree to which length of waiting queues affect service times – the salon manager considers it highly likely that knowing of a long waiting line will provoke either a stylist or a spa worker to accelerate service slightly, if only by virtue of less "chit-chat" with the client.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the cooperation and assistance of the salon manager and workers in providing data, explaining details of the salon's operation, and specifying the performance metrics of primary concern. Additionally, cogent criticisms from two anonymous referees have helped the authors significantly improve the paper.

REFERENCES

Baskaran, G., Bargiela, A.,Qu, R. 2013. A study of cost effective scheduling of nurses based on the domain transfer method. Proceedings of the 27th European Conference on Modelling and Simulation, 309-314. May 27-30, Ålesund (Norway).

Bhat, U. 2015. An introduction to queueing theory: Modeling and analysis in applications. 2nd ed. Basel, Switzerland: Birkhäuser

Chen, H-N., Jin, J., Rajavelu, G., and Reichenbach, C. 2004. Modeling and simulation of consumer credit originations processes. Proceedings of the 2004 Winter Simulation Conference, Volume 2, 1821-1826. December 5-8, Washington D.C. (United States of America).

Clemen, R. and Reilly, T. 2014. Making hard decisions. 3rd ed. Mason, Ohio (United States of America): South-Western.

Currie, C. S. M. and Cheng, R. C. H. 2016. A practical introduction to analysis of simulation output data. Proceedings of the 2016 Winter Simulation Conference, 118-132. December 11-14, Arlington Virginia (United States of America).

Kroemer, K. H. E. and Grandjean, E. 1997. Fitting the task to the human. 5th ed. London (United Kingdom): Taylor & Francis Limited.

Leemis, L. 2002. Stat::Fit Fitting continuous and discrete distributions to data. ORMS Today 29 (3), 52-55.

Lefever, W., Aghezzaf, E-H., and Hadj-Hamou, K. 2018. Stochastic optimization of a large-scale inventory-routing problem with transshipment through introduction of effective simulation steps. Proceedings of the 2018 Industrial Simulation Conference, 93-97. June 6-8, Ponta Delgada (Portugal).

Otamendi, J. and Pastor, J. M. 2006. Dynamic analysis of the waiting area in a public station. Proceedings of the 20th European Conference on Modelling and Simulation, 338-343. May 28-31, Bonn (Germany).

Prochaska, K. and Thiesing, R. 2017. Introduction to Simio. Proceedings of the 2017 Winter Simulation Conference, 4410-4419. December 3-6, Las Vegas, Nevada (United States of America).

Radosiński, E. 1998. The validation of an industrial firm's simulation model. Proceedings of the 8th European Simulation Symposium, Volume 1, 206-210. October 24-26, Genoa (Italy).

Sargent, R. G. 2004. Validation and verification of simulation models. Proceedings of the 2004 Winter Simulation Conference, Volume 2, 1821-1826. December 5-8, Washington D.C. (United States of America).

Sivaramakrishnan, S., Chandrasekaran S., Dhanapal J., Sekar, P., and Williams, E. 2016. Simulation improves operations at a specialized takeout restaurant. Proceedings of the 30th European Conference on Modelling and Simulation, 59-65. May 31 – June 3, Regensburg (Germany).

Smith, J., Sturrock, D., and Kelton, W. 2017. Simio and simulation: Modeling, analysis, applications. 4th ed. Sewickley, Pennsylvania (United States of America): Simio LLC.

Ståhl, I. 1996. Teaching discrete-event simulation to business students. Proceedings of the European Simulation Multiconference, 399-402. June 2-6, Budapest (Hungary).

Tezak, E. and Folawn T. 2011. Successful salon & spa management, 6th ed. Clifton Park, New York: Milady.

Williams, E. J., Karaki, M., and Lammers, C. 2002. Use of simulation to determine cashier staffing policy at a retail checkout. Proceedings of the 14th European Simulation Symposium, 172-176. October 23-26, Dresden (Germany).

AUTHOR BIOGRAPHIES

Baliada Ramesh Ganapathi is a graduate student in Business Analytics enrolled with the College of Business at the University of Michigan-Dearborn. Prior to this, he studied mechanical engineering as his first graduate degree program at MVIT University, India. He has extensive experience in data analytics and data visualization with a keen eye for improving processes for companies large and small. He was a Research Assistant with the iLabs, a center for innovative research at the University of Michigan- Dearborn, where he used data analysis to advise one of Michigan's largest non-profit organizations, ACCESS. He has also had the opportunity of working with some of Michigan's top companies, including J.D. Power and DTE Energy. At DTE Energy, he was instrumental in the automation process of weekly reporting and helped in improving the efficiency of the distribution operations system. He is a student member of the Indian Graduate Student Association (IGSA) and has worked with non-profit organizations like Habitat for Humanity and Gleaners Food Bank. He was also the recipient of the Non-Resident Graduate Tuition Differential Scholarship (NRTD) for Fall 2017 and Winter 2018. His email address is gramesh@umich.edu.

Brittany Harju is a graduate student in business analytics enrolled with the College of Business at the University of Michigan-Dearborn. Her bachelor's degree is in operations management from the Kelley School of Business at Indiana University. She has experience in data governance and quality as a Master Data Specialist at Carhartt, Incorporated. She was also responsible for data accuracy and data integrity at Trinity Health as a Business Applications Specialist. Her email address is bharju@umich.edu.

Daniel Scipione is a graduate student in the M.S. Business Analytics program enrolled with the College of Business at the University of Michigan-Dearborn. Prior to this, Daniel earned a B.S. in biology at the University of Michigan-Ann Arbor. Daniel has always had an interest in statistics, science, and mathematics. He decided that the Business Analytics program matched his analytical background. During his time at the University of Michigan - Dearborn business school, he completed an internship during the summer of 2017 as a data analyst with Fiat Chrysler Automobiles. Daniel will be graduated in May of 2018. He just recently secured a full-time job offer with Urban Science in Detroit, MI, and will begin his job May 1, 2018. His e-mail address is dscip@umich.edu.

Kristina Vujic is a graduate student in the Business Analytics program enrolled with the College of Business at the University of Michigan-Dearborn. Her first graduate degree is in Economics from Wayne State University where she also completed her undergraduate degrees in Mathematics and Economics. She currently works as an engineering business lead in the Budget & Forecasting group for Global Propulsion Systems at General Motors Corporation. Prior to this, she was a trade preference analyst with STTAS qualifying goods for free trade under various free trade agreements. Her email address is kvujic@umich.edu.

Edward J. Williams holds bachelor's and master's degrees in mathematics (Michigan State University, 1967; University of Wisconsin, 1968). From 1969 to 1971, he did statistical programming and analysis of biomedical data at Walter Reed Army Hospital, Washington, D.C. He joined Ford Motor Company in 1972, where he worked until retirement in December 2001 as a computer software analyst supporting statistical and simulation software. After retirement from Ford, he joined PMC, Dearborn, Michigan, as a senior simulation analyst. Also, since 1980, he has taught classes at the University of Michigan, including both undergraduate and graduate simulation classes using GPSS/H™, SLAM II™, SIMAN™, ProModel®, SIMUL8®, or Arena®. He is a member of the Institute of Industrial Engineers [IIE], the Society for Computer Simulation International [SCS], and the Michigan Simulation Users Group [MSUG]. He serves on the editorial board of the International Journal of Industrial Engineering – Applications and Practice. During the last several years, he has given invited plenary addresses on simulation and statistics at conferences in Monterrey, México; İstanbul, Turkey; Genova, Italy; Rīga, Latvia; and Jyväskylä, Finland. He served as a co-editor of Proceedings of the International Workshop on Harbour, Maritime and Multimodal Logistics Modelling & Simulation 2003, a conference held in Rīga, Latvia. Likewise, he served the Summer Computer Simulation Conferences of 2004, 2005, and 2006 as Proceedings co-editor. He was the Simulation Applications track coordinator for the 2011 Winter Simulation Conference. A paper he co-authored with three of his simulation students won "best paper in track" award at the Fifth International Conference on Industrial Engineering and Operations Management, held in Dubai, United Arab Emirates, in March 2015. His email addresses are ewilliams@pmcorp.com and williame@umich.edu.

APPENDIX A

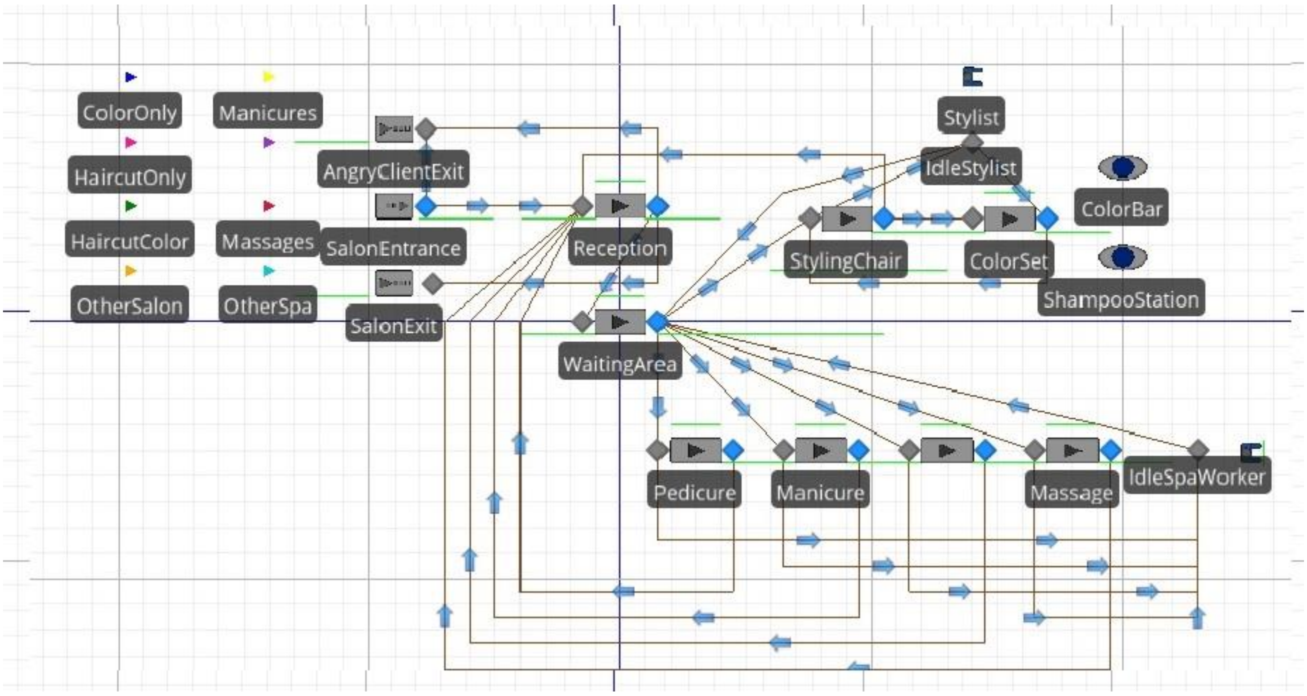
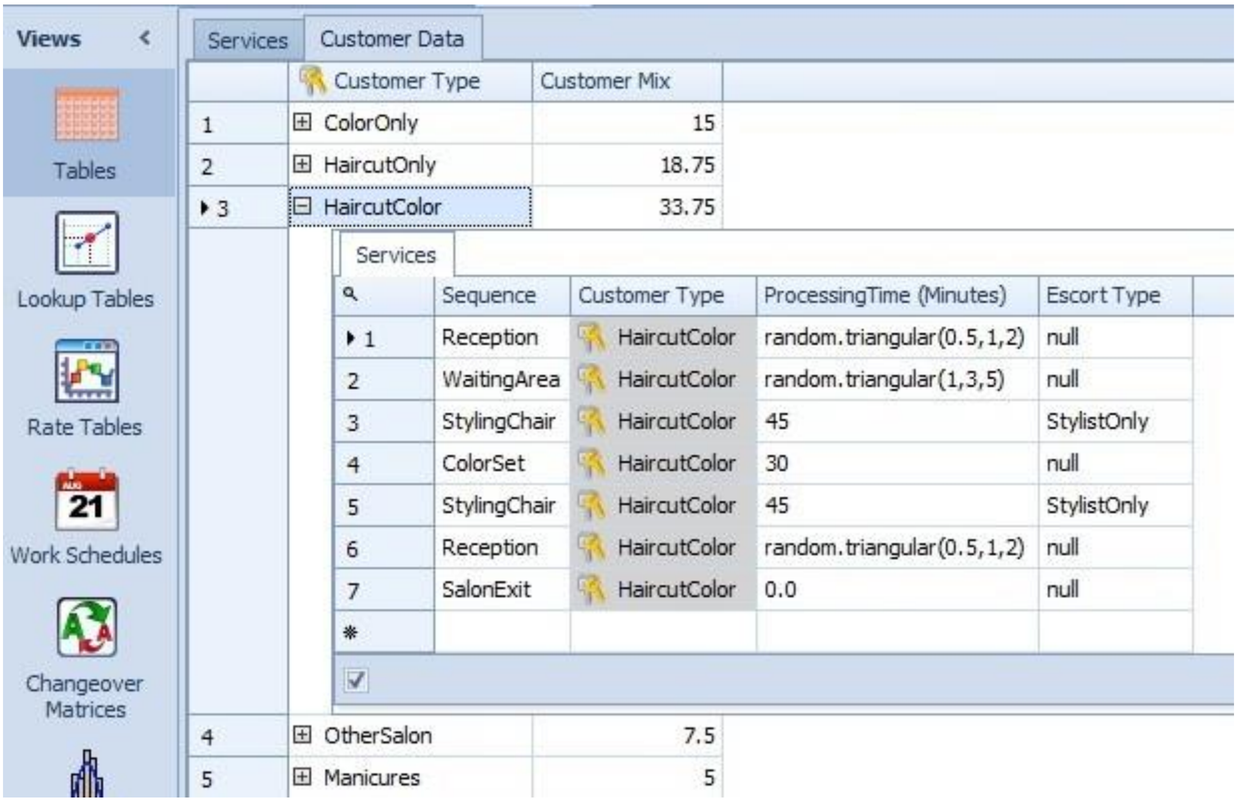


Figure 1: Screen of the Simio® Model Layout

APPENDIX B



	Customer Type	Customer Mix
1	⊞ ColorOnly	15
2	⊞ HaircutOnly	18.75
3	⊟ HaircutColor	33.75

Services

	Sequence	Customer Type	ProcessingTime (Minutes)	Escort Type
1	Reception	HaircutColor	random.triangular(0.5,1,2)	null
2	WaitingArea	HaircutColor	random.triangular(1,3,5)	null
3	StylingChair	HaircutColor	45	StylistOnly
4	ColorSet	HaircutColor	30	null
5	StylingChair	HaircutColor	45	StylistOnly
6	Reception	HaircutColor	random.triangular(0.5,1,2)	null
7	SalonExit	HaircutColor	0.0	null

	Customer Type	Customer Mix
4	⊞ OtherSalon	7.5
5	⊞ Manicures	5

Figure 2: Simio® “Relational Database” Internal Data Representation

Proc. of the Int. Conference on Modeling and Applied Simulation 2018,
ISBN 978-88-85741-07-2; Bruzzone, De Felice, Frydman, Longo, Massei and Solis, Eds.

A STUDY ON PASSENGER-TO-DECK ASSIGNMENT RULES FOR MULTI-DECK ELEVATOR SYSTEMS

Satoshi Goto[a], Shunji Tanaka[b]

[a],[b]Kyoto University

[a]goto@rotary.kuee.kyoto-u.ac.jp, [b]tanaka@kuee.kyoto-u.ac.jp

ABSTRACT
This study addresses multi-deck elevator systems that utilize elevator cars with multiple decks. In the group control of such systems, it is necessary to determine the assignment of passengers to decks as well as the assignment to cars. We propose a heuristic rule for the passenger-to-deck assignment to achieve better group control. Using a model of multi-deck elevator systems, we conduct computer simulation to examine its effectiveness.

Keywords: multi-deck elevator system, group control, passenger-to-deck assignment, heuristic rule

1. INTRODUCTION
An elevator system is one of the most important vertical transportation facilities. Along with the increase of high-rise buildings, the necessity for efficient elevator systems is also increasing nowadays. Although the most straightforward way for improving the transportation capability is to install additional elevator shafts (and elevator cars), it is not always possible due to limitation of floor space. Other solutions are a double-deck elevator system (DDES) that uses elevator cars composed of two decks and a multi-car elevator system (MCES) that uses several independent cars in each shaft. To take full advantages of such systems, however, a more intelligent group control method is crucial than that for a single-deck elevator system (SDES). In this paper, we consider a group control problem for multi-deck elevator systems (MDESs), as a generalization of DDESs, whose cars possibly have more than two decks.
The group control for SDESs has been widely studied for a long time. Research papers typically aim to assign them to appropriate cars so as to minimize waiting time of passengers. For example, Tanaka et al. (2005a, 2005b) proposed a branch-and-bound algorithm for an SDES with a single car. They showed by computer simulation that a destination hall call system (Strakosch 1998) that enables passengers to register their destination floors directly before boarding a car can considerably improve the transportation capability under heavy passenger traffic. Hiller et al. (2014) formulated the group control problem of an SDES with a destination hall call registration system under immediate assignment (IA) and delayed assignment (DA). The IA assigns a passenger to a car immediately after his/her call is registered, while in the DA the assignment can be delayed until a car arrives at the floor where he/she is waiting. Ruokokoski et al. (2015, 2016) also provided formulations (routing formulation and assignment formulation) of the group control problem for an SDES. There are also some studies on DDESs, granted that they are not so many as those for SDESs (Sorsa, Siikonen, and Ehtamo 2003; Hirasawa, Eguchi et al. 2008; Mabu, Yu et al. 2010). In these studies, the assignment of passengers to decks is determined by soft computing approaches such as a genetic algorithm and a genetic network programming. For a real-time application of group control, however, it is desirable to construct a simple and good enough heuristic rule. Tanaka (2014) decomposed the group control problem for MDESs into a passenger-to-car assignment problem and a passenger-to-deck assignment problem, and investigated the latter problem. The objective of this problem is to assign passengers to individual decks under the assumption that the assignment of passengers to cars is already fixed. In Tanaka (2014), a simple heuristic rule was proposed for this problem as well as a mixed-integer programming (MIP) formulation. However, the rule did not yield a good performance compared with the MIP approach.
The purpose of this study is to propose a better passenger-to-deck assignment rule. Then, using a model of MDESs, we conducted computer simulation to examine its effectiveness.

2. MULTI-DECK ELEVATOR SYSTEM
Suppose that an MDES with a single car composed of N_D decks is installed in an N_F-story building as illustrated in Figure 1. The decks are numbered as deck 0, deck 1, …, deck $(N_\mathrm{D}-1)$ from the bottom to the top. There are $(N_\mathrm{D}-1)$ garage floors under the ground (1st) floor so that passengers waiting at the 1st floor can board the topmost deck (deck $(N_\mathrm{D}-1)$). For simplicity, the capacity of the decks is ignored.
In this MDES, destination hall call registration is assumed: all passengers register their destination floors prior to boarding at each floor.

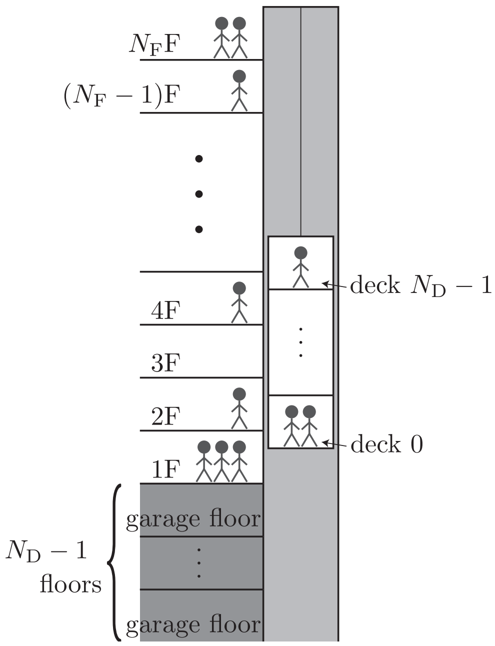


Figure 1: Multi-deck Elevator System

Our objective is to assign passengers who register their calls to the decks of the car so as to transport them as efficiently as possible. To this end, we consider a problem to assign passengers waiting at floors to decks every time when a new passenger registers a hall call. Here, we assume that after the assignment is determined, the car serves the assigned passengers using the selective-collective rule (Strakosch 1998), which is adopted in most elevator systems. In this rule, the car serves passengers in the following order:

1. Every passenger whose direction coincides with the current direction of the car and whose origin floor is ahead of the car.
2. Every passenger whose direction is opposite to that of the car.
3. Remaining passengers.

An example is shown in Figure 2.
Assume that passenger p who travels from floor O_p to D_p registers a call at time t_p. The information about passengers that is available for the system at t_p is

$$\mathcal{P}^{\text{on}} := \{(i, t_i, O_i, D_i, d_i) \mid \text{passenger } i \text{ is on board the car}\} \quad (1)$$

$$\mathcal{P}^{\text{w}} := \{(i, t_i, O_i, D_i, d_i) \mid \text{passenger } i \text{ is waiting at a floor}\} \quad (2)$$

where d_i is the assigned deck of passenger i. The state of the car at t_p is also known to the system, and it is represented by the current floor I, direction $H \in \{\text{UP}, \text{DOWN}\}$, and door state $U \in \{\text{OPEN}, \text{CLOSE}\}$.

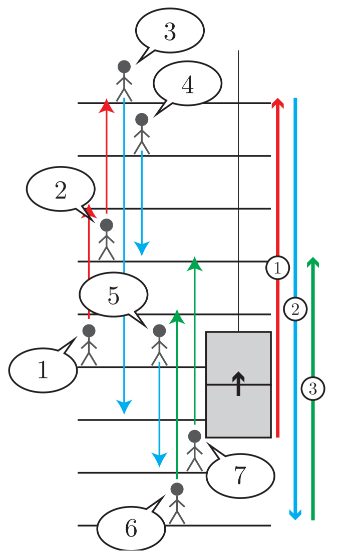


Figure 2: Selective-collective Rule

3. PASSENGER-TO-DECK ASSIGNMENT RULES

3.1. Simple Assignment Rule by Tanaka (2014)

In Tanaka (2014), a simple passenger-to-deck assignment rule was proposed. This rule is based on the observation that the transportation efficiency depends on total car travel time, which is determined by the number of car stops and the total car travel distance. When a new passenger registers a call, it assigns him/her to the deck that minimizes the number of car stops. Ties are broken by the total car travel distance (shorter is better). If it is also the same, the deck which reaches the destination floor of the passenger first is chosen. Specifically, we assign passenger p to the deck as follows:

1. Identify the decks to which passenger p can be assigned and let $\mathcal{C}_p$ be the set of such decks.
2. For each deck in $\mathcal{C}_p$, calculate the number of car stops. If the deck is uniquely determined, let d_p be that deck and terminate the procedure.
3. For each deck in $\mathcal{C}'_p$ $(:= \{d \in \mathcal{C}_p \mid d$ minimizes the number of car stops$\})$, calculate the total car travel distance. If the deck is uniquely determined, let d_p be that deck and terminate the procedure.
4. The deck in $\mathcal{C}''_p$ $(:= \{d \in \mathcal{C}'_p \mid d$ minimizes the total car travel distance$\})$ that reaches the destination floor of passenger p first is identified, and let d_p be that deck. More specifically, the highest deck is chosen if the destination of the passenger is upward, and the lowest deck otherwise.

Note that the number of stops and the total travel distance can be calculated only with the information on passengers and the car state.

3.2. Proposed Rule

The rule provided in 3.1 determines only the assignment of the passenger who newly registered a call, meaning that once a passenger is assigned to a deck, it is never changed. On the other hand, the proposed rule re-assigns the passengers who have not yet boarded the car as well. That is, the rule re-assigns the passengers in $\mathcal{P}^{\mathrm{w}}$ at the same time. In other words, the previous rule is for the IA, whereas we response a rule for the DA. Our rule first classifies passengers $\mathcal{P}^{\mathrm{w}} \cup \{p\}$ into three groups according to the selective-collective rule: those ahead of the car whose directions coincide with the current direction of the car (group 1), those in the opposite direction (group 2), and the others (group 3). Then, the passenger-to-deck assignment is determined in each group separately. Since it is not easy to determine the assignment of all passengers in a group at once, we determine the assignment of each passenger one by one using the rule in 3.1. The problem here is the assignment order. If we fix the assignment in the order of call registration time t_i, it makes little difference from the previous rule. An intuitive alternative is the boarding order: the nondecreasing order of O_i if the car direction is upward, and the nonincreasing order if downward. However, it turned out to be ineffective as illustrated in the following example. Suppose that an empty double-deck car ($N_{\mathrm{D}} = 2$) stays at the garage floor and that passengers A and B are waiting at floors 3 and 6, respectively. Suppose further that their destinations are floors 5 and 8, respectively. Since the car can load passenger A earlier than passenger B, passenger A is assigned to deck 1 according to 3 of the rule in 3.1. Then, passenger B is assigned similarly to deck 1. Therefore, (deck 0 of) the car stops at floors 2, 4, 5, and 7. However, the number of stops becomes 3 (at floors 3, 5, and 7) without increasing the total car travel distance if passenger A is assigned to deck 0. This can be achieved by assigning passenger B first and passenger A next, without modifying the rule in 3.1. We observed that such a situation frequently occurs, and hence we determine the assignment in the reverse order of boarding. An example of the assignment order is shown in Figure 3 where that is presented by the numbers in boxes.

We further improve this rule in the following two ways. First, noting that each deck should in general stop at the ground floor where most passengers board or leave the car, we calculate the number of car stops in 2 of the rule in 3.1 by excluding those at the ground floor. It is expected to reduce a negative effect on the performance caused by minimizing the numbers of stops at the ground floor. The second is an improvement by a local search. We apply a simple local search to the assignment obtained by the proposed rule. The neighborhood of an incumbent solution (assignment) is generated by changing the assignment of one passenger.

4. COMPUTATIONAL SIMULATION

We examined the effectiveness of the proposed rule by computational simulation.

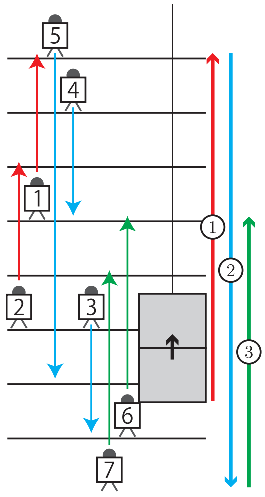


Figure 3: Assignment Order in Proposed Rule

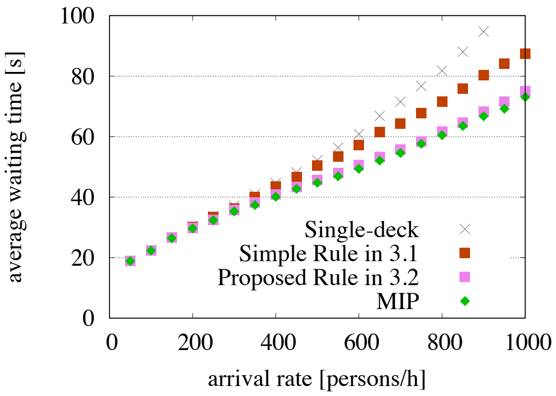
The height of the building is set to $N_{\mathrm{F}} = 30$. For simplicity, we assume that the car speed is constant and is 1floor/s. The door opening time, closing time, passenger boarding time, and leaving time are all assumed to 1s. We generated passenger data (arrival time, origin floor, and destination floor) to emulate uppeak traffic in the morning when most passengers are going upward from the ground floor, and downpeak traffic when most passengers leave the car from the ground floor.

The results are shown in Figures 4 and 5, where the average waiting time of passengers over 30 runs of the simulation is plotted against the passenger arrival rate for the proposed rule in 3.2 as well as the simple rule in 3.1 and the MIP approach in Tanaka (2014). In the MIP approach, the passenger-to-deck assignment problem is formulated as an MIP problem to minimize total car travel time (see Appendix A), and solved by IBM ILOG CPLEX Optimization Studio 12.6.3. For comparison, we also provided the results with $N_{\mathrm{D}} = 1$ (single-deck) in the same graphs.

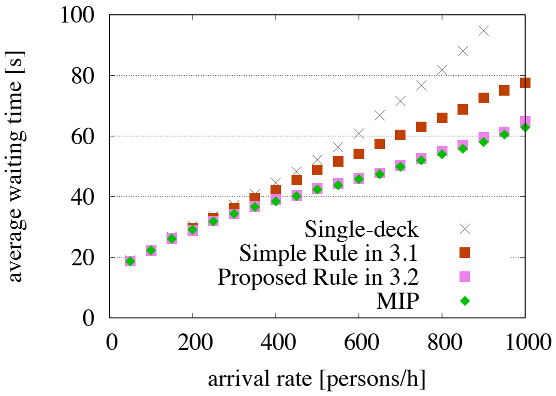
The results indicate that the waiting time by the proposed rule is shorter than that by the rule in 3.1, especially when the passenger traffic is heavy. Moreover, the proposed rule yields as good results as the MIP approach in much shorter computation time (typically, the proposed rule is five times as fast as the MIP approach). The model of the elevator system used in the simulation is so simple that we do not consider the acceleration, deceleration, and capacity of the car. However, even if the acceleration and deceleration are taken into account, the number of car stops remains the primal factor for waiting time. Moreover, it is not so often in practice that the number of passengers exceeds the car capacity. Therefore, we think that the effectiveness of our proposed rule is verified to some extent by these simulation results.

5. CONCLUSION

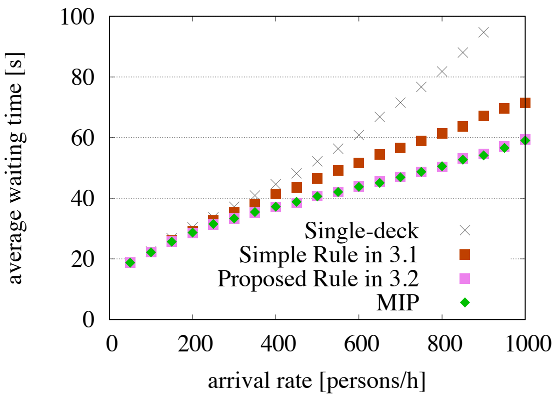
In this study we proposed a new heuristic rule for the passenger-to-deck assignment problem in MDESs.



(a) $N_{\mathrm{D}} = 2$

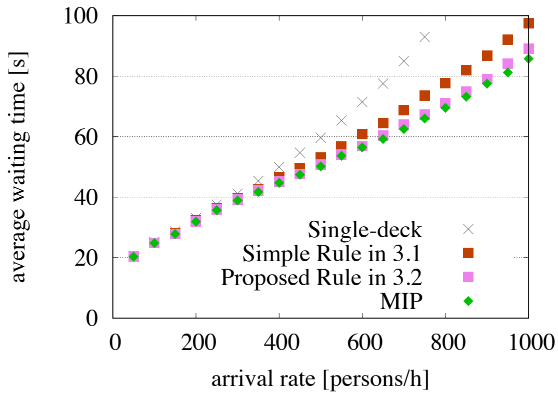


(b) $N_{\mathrm{D}} = 3$

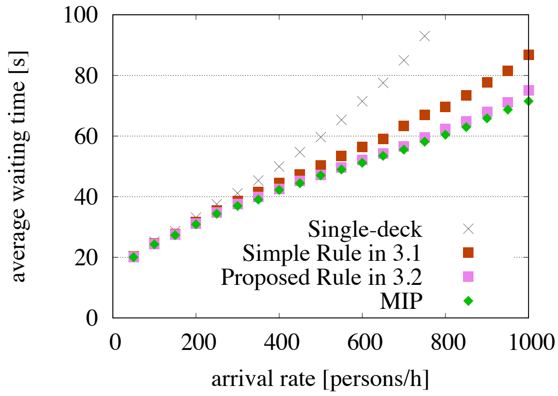


(c) $N_{\mathrm{D}} = 4$

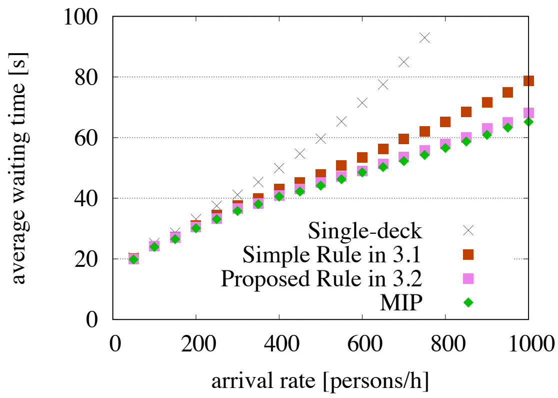
Figure 4: Simulation Results (Uppeak)



(a) $N_{\mathrm{D}} = 2$



(b) $N_{\mathrm{D}} = 3$



(c) $N_{\mathrm{D}} = 4$

Figure 5: Simulation Results (Downpeak)

Its effectiveness was verified by computer simulation. However, the model of an MDES was so simple that no acceleration or deceleration of the car was considered, the car capacity was ignored, and so on. It is desirable to examine the effectiveness of the rule using a more detailed model. It will also be necessary to consider not only the passenger-to-deck assignment problem but also the passenger-to-car assignment problem in MDESs composed of multiple cars. These topics are left for future research.

APPENDIX

A. ASSIGNMENT WITH MIP FORMULATION

In this section we introduce the MIP formulation of the passenger-to-deck assignment problem (Tanaka 2014). Here, we assume that the car is at floor I with the doors closed ($U = \mathtt{CLOSE}$, $H = \mathtt{UP}$, $\mathcal{P}^{\mathrm{on}} = \emptyset$). The problem under the other situations can be formulated similarly, and so is omitted. The following sets define the input of the formulation:

- $\mathcal{P}^s$: the set of the passengers who belong to group $s (\in \{1, 2, 3\})$ in the selective-collective rule
- $\mathcal{C}$: the set of decks ($\mathcal{C} \in \{0, 1, \cdots, N_{\mathrm{D}} - 1\}$)
- $\mathcal{C}_i$: the set of decks to which passenger i can be assigned ($\mathcal{C}_i \subseteq \mathcal{C}$)
- $\mathcal{F}^s$: the set of floors where the car can stop to load or unload some passengers in group s ($\mathcal{F}^s = \bigcup_{d \in \mathcal{C}_i, i \in \mathcal{P}^s} \{O_i - d\} \cup \{D_i - d\}$)
- $\mathcal{H}^s$: the set of floors which can be the highest floor that the car reaches ($\mathcal{H}^s \subseteq \mathcal{F}^s, s \in \{1, 2, 3\}$)
- $\mathcal{L}^s$: the set of floors which can be the lowest floor that the car reaches ($\mathcal{L}^s \subseteq \mathcal{F}^s, s \in \{2, 3\}$)
- $\mathcal{F}_{\mathrm{D}}^1 = \mathcal{F}^1 \setminus (\mathcal{H}^1 \cap \mathcal{H}^2)$, $\mathcal{F}_{\mathrm{D}}^2 = \mathcal{F}^2 \setminus ((\mathcal{H}^1 \cap \mathcal{H}^2) \cup (\mathcal{L}^2 \cap \mathcal{L}^3))$, $\mathcal{F}_{\mathrm{D}}^3 = \mathcal{F}^3 \setminus (\mathcal{L}^2 \cap \mathcal{L}^3)$
- $\mathcal{H}_j^s = \{k \in \mathcal{H}^s \mid k > j\}$
- $\mathcal{L}_j^s = \{k \in \mathcal{L}^s \mid k < j\}$
- $\mathcal{P}_{jd}^s = \{i \in \mathcal{P}^s \mid d \in \mathcal{C}_i \wedge (j = O_i - d \vee j = D_i - d)\}$

The decision variables are as follows:

- x_{id}: 1 if passenger i is assigned to deck d, 0 otherwise ($d \in \mathcal{C}_i, i \in \mathcal{P}^s$).
- y_j^s: 1 if the car stops at floor j, 0 otherwise ($j \in \mathcal{F}^s$).
- h_j^{12}: 1 if the car changes the direction at floor j, 0 otherwise ($j \in \mathcal{H}^1 \cup \mathcal{H}^2$).
- l_j^{23}: 1 if the car changes the direction at floor j, 0 otherwise ($j \in \mathcal{L}^2 \cup \mathcal{L}^3$).
- h_j^3: 1 if the car changes the direction at floor j, 0 otherwise ($j \in \mathcal{H}^3$).
- v_j^{12}: 1 if there are both boarding and leaving passengers at the floor where the car changes the direction (floor j with $h_j^{12} = 1$), 0 otherwise ($j \in \mathcal{H}^1 \cup \mathcal{H}^2$).
- v_j^{23}: 1 if there are both boarding and leaving passengers at the floor where the car changes the direction (floor j with $l_j^{23} = 1$), 0 otherwise ($j \in \mathcal{L}^2 \cup \mathcal{L}^3$).
- z_j^s: The maximum number of passengers over the decks who board or leave the car at floor j ($j \in \mathcal{F}^s$). The stoppage time at floor j is determined by z_j^s.
- z_j^{12}: The maximum number of passengers over the decks who board or leave the car at floor j, if the car changes the direction at floor j and there are both boarding and leaving passengers ($j \in \mathcal{H}^1 \cap \mathcal{H}^2$).
- z_j^{23}: The maximum number of passengers over the decks who board or leave the car at floor j, if the car changes the direction at floor j and there are both boarding and leaving passengers ($j \in \mathcal{L}^2 \cap \mathcal{L}^3$).

The objective function is the total car travel time, which is the sum of stoppage time at each stop floor and travel time between stops:

$$\begin{aligned} \min \sum_{s=1}^{3} \sum_{j \in \mathcal{F}^s} \left(z_j^s + 2y_j^s\right) + 2 \sum_{j \in \mathcal{H}^1 \cup \mathcal{H}^2} j h_j^{12} \\ - 2 \sum_{j \in \mathcal{L}^2 \cup \mathcal{L}^3} j l_j^{23} + \sum_{j \in \mathcal{H}^3} j h_j^3 \\ + \sum_{j \in \mathcal{H}^1 \cap \mathcal{H}^2} \left(z_j^{12} - 2v_j^{12}\right) \\ + \sum_{j \in \mathcal{L}^2 \cap \mathcal{L}^3} \left(z_j^{23} - 2v_j^{23}\right) \end{aligned} \tag{3}$$

The constraints are described as follows.

$$\sum_{d \in \mathcal{C}_i} x_{id} = 1 , \qquad i \in \mathcal{P}^1 \cup \mathcal{P}^2 \cup \mathcal{P}^3 \tag{4}$$

$$y_j^s \geq x_{id} , \qquad i \in \mathcal{P}_{jd}^s, d \in \mathcal{C}_i, j \in \mathcal{F}^s, s \in \{1, 2, 3\} \tag{5}$$

$$\sum_{j \in \mathcal{H}^1 \cup \mathcal{H}^2} h_j^{12} = 1, \sum_{j \in \mathcal{L}^2 \cup \mathcal{L}^3} l_j^{23} = 1, \sum_{j \in \mathcal{H}^3} h_j^3 = 1 \tag{6}$$

$$\left|\mathcal{H}_j^s\right|\left(1 - h_j^{12}\right) \geq \sum_{k \in \mathcal{H}_j^s} y_k^s , \quad j \in \mathcal{H}^1 \cup \mathcal{H}^2, s \in \{1, 2\} \tag{7}$$

$$\left|\mathcal{L}_j^s\right|\left(1 - l_j^{23}\right) \geq \sum_{k \in \mathcal{L}_j^s} y_k^s , \quad j \in \mathcal{L}^2 \cup \mathcal{L}^3, s \in \{2, 3\} \tag{8}$$

$$\left|\mathcal{H}_j^3\right|\left(1 - h_j^3\right) \geq \sum_{k \in \mathcal{H}_j^3} y_k^3 , \qquad j \in \mathcal{H}^3 \tag{9}$$

$$\sum_{j \in \mathcal{H}^1 \cup \mathcal{H}^2} v_j^{12} \leq 1, \sum_{j \in \mathcal{L}^2 \cup \mathcal{L}^3} v_j^{23} \leq 1 \tag{10}$$

$$v_j^{12} \geq y_j^1 + y_j^2 + h_j^{12} - 2 , \qquad j \in \mathcal{H}^1 \cup \mathcal{H}^2 \tag{11}$$

$$v_j^{23} \geq y_j^2 + y_j^3 + l_j^{23} - 2 , \qquad j \in \mathcal{L}^2 \cup \mathcal{L}^3 \tag{12}$$

$$z_j^s \geq \sum_{i \in \mathcal{P}_{jd}^s} x_{id} , \qquad d \in \mathcal{C}, j \in \mathcal{F}_{\mathrm{D}}^s, s \in \{1, 2, 3\} \tag{13}$$

$$z_j^s \geq \sum_{i \in \mathcal{P}_{jd}^s} x_{id} - \left|\mathcal{P}_{jd}^s\right| v_j^{12}, \quad d \in \mathcal{C}, j \in \mathcal{H}^1 \cap \mathcal{H}^2, s \in \{1, 2\} \tag{14}$$

$$z_j^s \geq \sum_{i \in \mathcal{P}_{jd}^s} x_{id} - \left|\mathcal{P}_{jd}^s\right| v_j^{23}, \quad d \in \mathcal{C}, j \in \mathcal{L}^2 \cap \mathcal{L}^3, s \in \{2, 3\} \tag{15}$$

$$z_j^{12} \geq \sum_{s=1}^{2} \sum_{i \in \mathcal{P}_{jd}^s} x_{id} - \left(\left|\mathcal{P}_{jd}^1\right| + \left|\mathcal{P}_{jd}^2\right|\right)\left(1 - v_j^{12}\right), \quad d \in \mathcal{C}, j \in \mathcal{H}^1 \cap \mathcal{H}^2 \tag{16}$$

$$z_j^{23} \geq \sum_{s=2}^{3} \sum_{i \in \mathcal{P}_{jd}^s} x_{id} - \left(\left|\mathcal{P}_{jd}^2\right| + \left|\mathcal{P}_{jd}^3\right|\right)\left(1 - v_j^{23}\right), \quad d \in \mathcal{C}, j \in \mathcal{L}^2 \cap \mathcal{L}^3 \tag{17}$$

$$x_{id} \in \{0,1\}, \qquad d \in \mathcal{C}_i, i \in \mathcal{P}^1 \cup \mathcal{P}^2 \cup \mathcal{P}^3 \quad (18)$$
$$y_j^s \in \{0,1\}, \qquad j \in \mathcal{F}^s, s \in \{1,2,3\} \quad (19)$$
$$h_j^{12}, v_j^{12} \in \{0,1\}, \qquad j \in \mathcal{H}^1 \cup \mathcal{H}^2 \quad (20)$$
$$l_j^{23}, v_j^{23} \in \{0,1\}, \qquad j \in \mathcal{L}^2 \cup \mathcal{L}^3 \quad (21)$$
$$h_j^3 \in \{0,1\}, \qquad j \in \mathcal{H}^3 \quad (22)$$
$$z_j^s \geq 0, \qquad j \in \mathcal{F}^s, s \in \{1,2,3\} \quad (23)$$
$$z_j^{12} \geq 0, \qquad j \in \mathcal{H}^1 \cap \mathcal{H}^2 \quad (24)$$
$$z_j^{23} \geq 0, \qquad j \in \mathcal{L}^2 \cap \mathcal{L}^3 \quad (25)$$

REFERENCES

Tanaka S., Uraguchi Y., and Araki M., 2005a. Dynamic optimization of the operation of single-car elevator systems with destination hall call registration: Part I. Formulation and simulations. European Journal of Operational Research Vol. 167 No. 2: 550-573.

Tanaka S., Uraguchi Y., and Araki M., 2005b. Dynamic optimization of the operation of single-car elevator systems with destination hall call registration: Part II. The solution algorithm. European Journal of Operational Research Vol. 167 No. 2: 574-587.

Strakosch G.R., 1998. The Vertical Transportation Handbook 3rd ed. New York: Wiley.

Hiller B., Klug T., and Tuchscherer A., 2014. An exact reoptimization algorithm for the scheduling of elevator groups. Flexible Services and Manufacturing Journal Vol. 26 No. 4: 585-608.

Ruokokoski M., Ehtamo, H., and Pardalos, P.M., 2015. Elevator dispatching problem: a mixed integer linear programming formulation and polyhedral results. Journal of Combinatorial Optimization Vol. 29: 750-780.

Ruokokoski M. Sorsa J. Siikonen M.-L., Ehtamo, H., 2016. Assignment formulation for the Elevator Dispatching Problem with destination control and its performance analysis. European Journal of Operational Research Vol. 252 Issue 2: 397-406

Sorsa J., Siikonen M-L., Ehtamo H., 2003. Optimal control of double-deck elevator group using genetic algorithm. International Transaction in Operational Research Vol. 10: 103–114.

Hirasawa K., Eguchi T. et al., 2008. A Double-Deck Elevator Group Supervisory Control System Using Genetic Network Programming. IEEE Transactions on Systems, Man, and Cybernetics Part C vol. 38: 535–550.

Mabu S., Yu Lu et al., 2010. A Double-Deck Elevator Systems Controller with Idle Cage Assignment Algorithm Using Genetic Network Programming. Journal of Advanced Computational Intelligence and Intelligent Informatics Vol. 14 No. 5: 487–496.

Tanaka S., 2014. A Study on Deck Assignment for Multi-Deck Elevator Systems. Scheduling Symposium 2014, 231–236. September 29–30, Toyama (Japan). (*in Japanese*)

AN AGENT-BASED MODEL FOR ROBUSTNESS TESTING OF FREIGHT TRAIN SCHEDULES

Matthias Rößler[(a)], Matthias Wastian[(b)], Michael Landsiedl[(c)], Niki Popper[(d)]

[(a),(b),(c),(d)]dwh GmbH, Neustiftgasse 57-59. 1070 Vienna, Austria

[(a)]matthias.roessler@dwh.at, [(b)]matthias.wastian@dwh.at, [(c)]michael.landsiedl@dwh.at
[(d)]niki.popper@dwh.at

ABSTRACT

Optimization for train schedules of freight trains is extremely complex. Having a tight schedule is very efficient, but the impact of delays and cancellation rises and puts an environmentally sustainable and accurately timed supply of goods at risk. In order to address these issues, the Rail Cargo Austria commissioned a proof of concept that incorporates the optimization of train schedules with an agent-based model that tests the robustness of the optimized schedule against delays. Thereby the delays do not have to be incorporated into the optimization directly which would make the optimization problem even more complex and impossible to solve with classic techniques. Additionally, the use of historical data of freight train runs is integrated into the proof of concept and certain parameters for the simulation are gathered by data analysis and machine learning techniques. In this paper the agent-based model and a scenario to show its capabilities are presented.

Keywords: train networks, agent-based models, data analysis, optimization problem

1. INTRODUCTION

The planning of freight transports is a very complex task due to several factors. One is the sheer number of possibilities to schedule routes and locomotives, another the complex interaction with the passenger traffic that plays a very important role in Austria. An optimal planning regarding costs and efficiency leads to a very tight schedule. But for such a schedule the impact of delays and cancellations is very high, especially if there are no back-up traction units available. Thus, a tradeoff between efficiency and minimizing the risk of unserviceability, i.e. maximizing the robustness of the schedules, is necessary.

Currently, a proof of concept (PoC) for the Rail Cargo Austria (RCA) that addresses these issues of the operation in Austria is developed and a follow-up project is already scheduled. The PoC consists of an optimization part, where optimal train schedules are created, and a simulation part, that tests the generated schedules for its robustness against train delays. This paper is focused on the simulation part of the PoC and introduces the corresponding agent-based approach with its entities and their parameters. As not all the important agent parameters are explicitly available (e.g. the delay probability for a certain train), we introduce an ansatz to gain an adequate approximation of these parameters from historical data provided by the RCA.

The final goal is to provide a decision support tool for the set-up of locomotive schedules for large freight railway networks.

2. MODELLING

In order to test the robustness of the train schedules, a Monte Carlo simulation for the impact of different delays is performed. The underlying model is a multi-agent-based model with the following entities:

- Railway stations
- Train tracks
- Traction units

For future applications the list of agent types can easily be expanded, e.g. with freight cars or conductors.

Additionally, the model contains the train schedule which is no agent. In the following subchapters the different elements are described.

2.1. Train Schedule

The train schedule is the core element of the model. It is executed throughout the simulation and according to its entries the traction units are shifted on the network that is built by the tracks and stations. A train that is booked from A to B is typically split into several entries for the train schedule. This makes it possible that different traction units are used for a train along the way.

An entry of the train schedule consists of the following elements:

- **Train ID & name:** The train ID is the mapping to the overlying train and is used for analysis purposes.
- **Train track:** The train track that is occupied if the entry is executed. The train track has among other parameters a certain capacity. Thus, if the track is still taken by other trains

and cannot be accessed by an arriving train, the train schedule has to be re-executed later.

- **Departure time:** The point in time the assigned train track should get occupied. In the case of a previous delay this time is ignored, and the train tries to occupy the track immediately.
- **Arrival time:** The difference between departure and arrival time defines the time the train needs for the assigned track.
- **Assigned traction units:** Typically, one traction unit is assigned to a train, but under certain conditions, for example if the track is too steep, multiple traction units can be assigned to a train.
- **Delay probability:** In order to measure the robustness of a certain train schedule, the trains need delays. The train delay is implemented as a normally distributed random number. The mean and the standard deviation can be influenced by different factors such as the track, time of day or year or the type of the train that occupies the track.

2.2. Railway Stations and Train Tracks

The rail network is a graph, where the train tracks are edges and the stations are nodes. A train has to be somewhere on the network at all times. In the model the tracks and the stations are agents, which allow certain calculations for the track or the station when different events occur (e.g. a train occupies the track or the station). Furthermore, certain parameters can be implemented for the agents. The tracks have the following parameters:

- **Capacity**: Train tracks have a certain capacity of how much trains can occupy them at a certain time. Within the PoC, the capacity will be a fixed number, but it can be modified to incorporate for example train lengths in future works.
  The capacity of the train tracks ensures the first type of delay propagation as a train cannot occupy a track, if its capacity is already reached.
- **Length**: The length of the track is needed to calculate the distance a traction unit has already travelled. This is necessary on the one hand for the analysis of the overall system and on the other hand to ensure that a certain maintenance cycle for the traction unit is fulfilled.

The railway stations have the following parameters:

- **Capacity:** For railway stations the capacity is given by the number of rails they have and how long they are. Since parked trains are not considered in the PoC and it is not clear if the true capacity of a railway station can be determined, the capacity of the railway station is not strictly enforced during simulation.
- **Maintenance points:** Additionally, a railway station can have a maintenance point where traction units can be serviced. Certain inspections and services have to be done in the home base of the traction unit, so the train schedule has to ensure that the traction unit is at its home base for that period of time. Inspections and services are performed periodically depending on the travelling distance of the traction unit.

2.3. Traction Units

The traction units are the active parts of the model. Traction units are assigned to different entries of the train schedule and occupy the associated train track. The distance until the next maintenance and the currently travelled distance are stored as parameters. If the maintenance is not performed in a timely manner, there are no consequences during the simulation, but such issues must be detected in an analysis step afterwards.
The second type of delay propagation happens if a traction unit that is assigned to an entry of the train schedule is not at the station where it should be. If this happens the execution of the train schedule entry has to be halted until the traction unit arrives.

3. HISTORICAL DATA

The parameterization of the model is mainly done using planning data. Especially the train schedule has to be planned in advance and is only analyzed with the simulation.
Other parameters that are not planned come from historical data of trains that utilized the same train network. The main parameters that are gathered through analyzing the historical data are the capacity of the train tracks and the delay probability for each train schedule entry.

3.1. Train Track Capacity

In order to accurately monitor the effects of certain delays on the overall train schedule, it is important for the tracks to have certain capacities.
The capacity of a train track is not easy to define as there are various definitions (Lindfeldt 2015). The easiest regarding the availability in the provided data is the maximum amount of trains on a track at the same time. This definition will be used throughout the PoC. An example for another definition would be the allowed minimum distance between two trains in combination with the maximum allowed velocity on the track. This would yield another form of capacity, but this cannot be derived from the provided data at this point.

3.2. Delay Probability

The main part of the provided historical data consists of past train schedules and their delays for all the stations

the trains passed through. This enables us to analyze the data and to identify certain factors for delays.
As the effects of the delays should be shown by means of the simulation, it is important to decide if a delay that is in the data is a primary delay or a delay that is caused because of another delay, i.e. a follow-up delay. This is the first challenge.
After the data is filtered and only initial delays are left, the next step is to identify influencing factors for the delay that can be used to calculate a delay probability for the train schedule entries. A major factor for this is certainly the train track but also other factors such as the time of day or type of train or traction unit could be of importance. Oneto et al. 2016 also suggest including exogeneous weather data.

4. SIMULATION

In the following chapter first results will be presented. In order to show the feasibility of the model a small scenario based on real timetables from the Austrian Federal Railways was created. For this scenario no traction units were assigned to the train schedule. This means that the only form of delay propagation is by the capacity restrictions of the train tracks.

4.1. Scenario

In order to create a comprehensible scenario a part of the Austrian railway network had to be chosen that is on the one hand somewhat self-contained to ensure readability of the results and on the other hand complex enough to show the capabilities of the modelling approach. To fulfill these two requirements the part of the railway network that is between "*Wien Westbahnhof*" and "*St. Pölten*" was chosen. Between these two stations exist two different routes, an old one that leads through "*Neulengbach*" and a new one that leads through "*Tullnerfeld*". Other stations were chosen in order to sufficiently map the real network, namely "*Wien Hütteldorf*", "*Hadersdorf*", "*Unterpurkersdorf*" and "*Wagram*". Figure 1 shows the routes of all trains that pass at least two of the mentioned stations in black. The routes we are interested in are shown in blue.

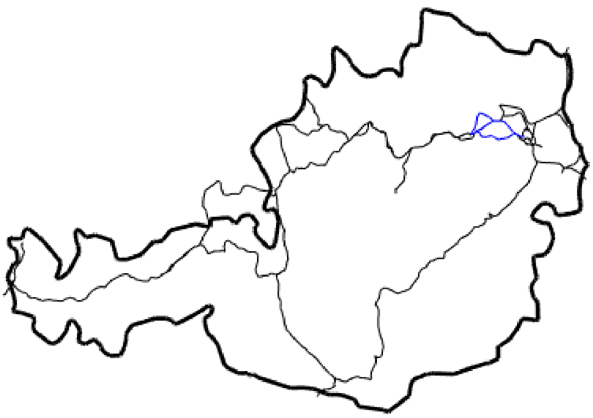


Figure 1: Routes of All Considered Trains (Black) and Region of Interest (Blue)

Figure 2 shows a closeup of the interesting region and includes locations of the stations mentioned above. It can be seen that there is an alternate route between "*Tullnerfeld*" (northernmost point) and "*St. Pölten*" (westernmost point). This route is primarily used by local trains and not by long-distance trains as the other two routes that are considered. Nevertheless, the trains operating on this route are also considered in the simulation. The green route is the one that leads through "*Neulengbach*".

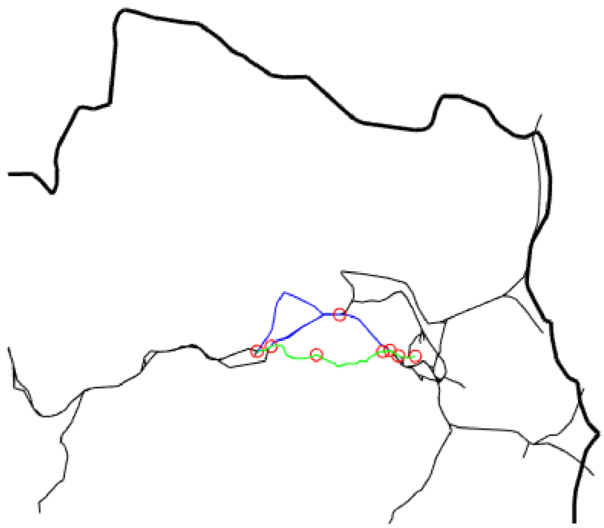


Figure 2: Detailed View of the Considered Routes and Stations

The train tracks considered in the model are the ones between the chosen stations. Additionally, "virtual" tracks that lead to and from every individual station were implemented to depict the trains that also operate outside of the considered region. These additional tracks have unlimited capacity whereas the other tracks have a capacity according to the considered timetable.
The train schedule for the scenario was taken from real timetables provided by the Austrian Federal Railways and consists of one traffic day. For the delay probabilities for the entries of the train schedule a probability of 0.8 and a normal distribution with mean of 60 minutes and standard deviation of 30 was chosen for all entries, except for those entries that are part of trains that go through "*Neulengbach*" (green route in figure 2). These entries have probability 0 for generating a delay. This setup is chosen to show that the effects of delays propagate over the railway network as delays on the green route have to be consequences of delays from other routes.

4.2. Results

The described scenario was simulated 1000 times and the resulting delays were averaged. For every simulation run the primary delay applied during the simulation defined by the random number described above was deducted to get only propagated delays in the simulation results. Figure 3 shows a bar chart of the delays (in minutes) for all considered trains over their respective departure time in minutes after midnight of the observed traffic day.

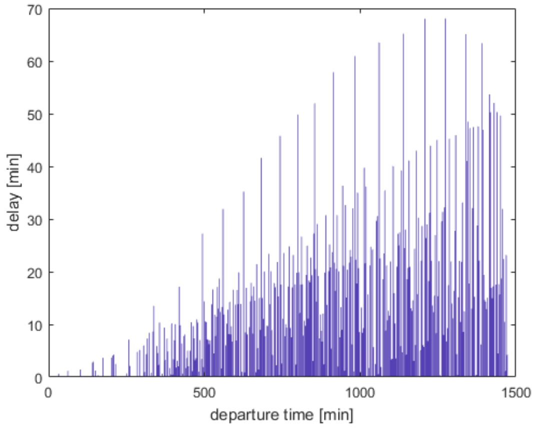


Figure 3: Averaged Propagated Delays of All Considered Trains

As can be seen the delays grow over time as expected.

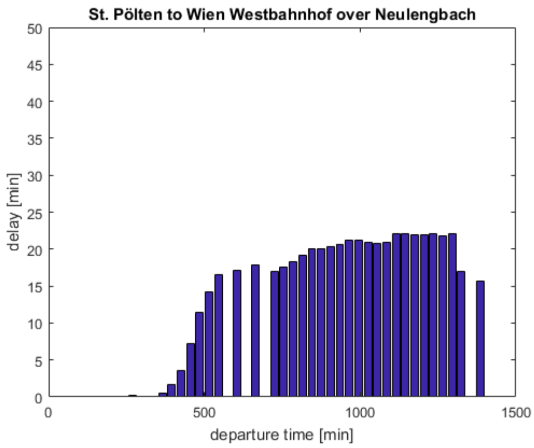


Figure 4: Averaged Propagated Delays of Trains that Operate on the Green Route in the Direction of Wien

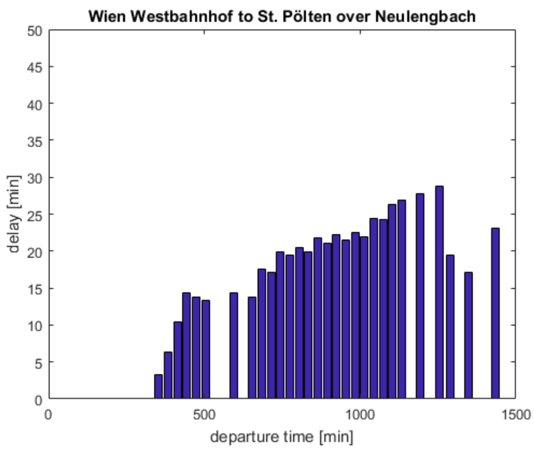


Figure 5: Averaged Propagated Delays of Trains that Operate on the reen route in the direction of St. Pölten

As mentioned above in the scenario description a train that follows the green route from Figure 2 does not generate delays. However, simulation results show that these trains also gather delays by occupied tracks from other delayed trains (figure 4 and figure 5). This shows that delay propagation in the model also works over track limits.

5. CONCLUSION & OUTLOOK

This paper has shown that using agent-based models for delay prediction of train schedules is feasible. A scenario was presented that shows the basic capabilities of the model. For the proof of concept some additional factors have to be considered, that were not part of the presented scenario. This includes the assignment of traction units which is already considered in the modelling approach. Also no findings of the data analysis were considered in the presented scenario. The delay probabilities and distributions were set the same on all train tracks except for the ones where the effects of the delay should be shown, where the probability was set to 0.

As the aim of the proof of concept is to develop optimal train schedules that also satisfy certain robustness criteria, the modelling of the delay propagation is a very important step, because considering the robustness of the schedule during the optimization would increase its complexity drastically.

A proposal for the further development of the PoC was granted by the Austrian Research Promotion Agency (FFG, project number 16767446) and will start soon. In the project the optimization and the simulation are going to be developed further and a closer connection of the two modules, optimization and simulation, is planned.

ACKNOWLEDGMENTS

This paper was written as part of a proof of concept commissioned by Rail Cargo Austria.

REFERENCES

Barthélemy J., Carletti T., 2017. An adaptive agent-based approach to traffic simulation. Transportation Research Procedia, 25: 1238-1248.

Cacchiani V., Caprara A., Galli L., Kroon L., Maróti G., Toth P., 2012. Railway Rolling Stock Planning: Robustness Against Large Disruptions. Transportation Science 46(2): 217-232.

Cerreto F., 2015. Micro-simulation based analysis of railway lines robustness. Proceedings of the 6th International Conference on Railway Operations Modelling and Analysis, pp. 164-1 – 164-13. March 23-26, Narashino, Japan.

Fernandes de Faria C.H., Monteiro da Costa Cruz M., 2017. Railway planning: an integrated approach using discrete event simulation. Proceedings of the International Conference on Harbor, Maritime and Multimodal Logistic Modelling and Simulation, pp. 148-156. September 18-20, Barcelona, Spain.

Flier H.-F., 2011. Optimization of Railway Operations: Algorithms, Complexity, and Models. Dissertation. ETH Zurich.

Geurts P., Ernst D., Wehenkel L., 2006. Extremely randomized trees. Machine Learning, 63(1):3-42.

Guo C., Berkhahn F., 2016. Entity Embeddings of Categorical Variables. CoRR. http://arxiv.org/abs/1604.06737.

Lindfeldt A., 2015. Railway capacity analysis – Methods for simulation and evaluation of timetables, delays and infrastructure. Dissertation. KTH Royal Institute of Technology, Stockholm.

Ljubovic V., 2009. Traffic simulation using agent-based models. Proceedings of the XXII International Symposium on Information, Communication and Automation Technologies, pp. 1-6. October 29-31, Sarajevo, Bosnia and Herzegovina.

Oneto L., Fumeo E., Clerico C., Canepa R., Papa F., Dambra C., Mazzino N., Anguita D., 2018. Train Delay Prediction Systems: A Big Data Analytics Perspective. Big Data Research 11: 54-64.

Oneto L., Fumeo E., Clerico C., Canepa R., Papa F., Dambra C., Mazzino N., Anguita D., 2017. Dynamic Delay Predictions for Large-Scale Railway Networks: Deep and Shallow Extreme Learning Machines Tuned via Thresholdout. IEEE Transactions on Systems, Man, and Cybernetics: Systems 47(10): 2754-2767.

Oneto L., Fumeo E., Clerico C., Canepa R., Papa F., Dambra C., Mazzino N., Anguita D., 2016. Advanced Analytics for Train Delay Prediction Systems by Including Exogenous Weather Data. Proceedings of IEEE International Conference on Data Science and Advanced Analytics. October 17-19, Montreal, Canada.

Pouryousef H., Lautala P., 2015. Hybrid simulation approach for improving railway capacity and train schedules. Journal of Rail Transport Planning & Management 5.4: 211-224.

Salgado D., Jolovic D., Martin P.T., Aldrete R.M.,2016. Traffic Microsimulation Models Assessment – A Case Study of International Land Port of Entry. Procedia Computer Science, 83:441-448.

Wood S., 2012. Traffic microsimulation – dispelling the myths. Traffic Engineering and Control, 53(9): 339-344.

Yaghini M., Khoshraftar M., Seyesabadi M., 2013. Railway passenger train delay prediction via neural network model. In Journal of Advanced Transportation, 47(3): 355-368.

CASH AND VOUCHER IMPACT FACTOR IN HUMANITARIAN AID: A SYSTEM DYNAMIC ANALYSIS

Fahimeh Allahi[(a)], Roberto Revetria[(b)], Roberto Cianci[(c)]

[(a)] Wolfson Department, University of Loughborough, Loughborough, UK
[(b),(c)]DIME, University of Genoa, Genoa, Italy)

[(a)] f.allahi@ lboro.ac.uk, [(b)]roberto.revetria@unige.it, [(c)] cianci@dime.unige.it

ABSTRACT
Natural and man-made disasters seem unpredictable every year, increasing a wide range of universal sufferers. Several people are affected by the direct outcomes of these disasters, and their life depends on disaster relief aid administered by humanitarian organizations. Recently, there has been renewed interest in cash distribution in the humanitarian sector during disaster relief to increase access of vulnerable people to supporting services such as health or education and develop their life's condition while rising the efficiency of humanitarian organizations committed to the program. The research proposes a casual-loop and system dynamic model to assess multi aspects of related impact factors to provide optimal support to beneficiaries. The model provides a decision-making framework with a high-level overview of the interactions between the economy, education, health, and the psychological aspects of recipient's life, provide a system dynamics analysis including relationships that could have led to improve the vulnerable people's condition life.

Keywords: humanitarian relief, casual loop model, system dynamic, decision making framework

1. INTRODUCTION

Natural and human-made disasters cause the disruption of cities and equipment, famine, illnesses, death, human misery and health of people around the world which have grown according to the report by CRED (2014). It is essential to aid recipients immediately after the occurrence of disasters in numerous processes, such as providing meals assistance, cash transferring, and voucher assistance and request from humanitarian organizations to mobilize themselves to support the affected area, reducing the impacts and initiating the improvement of operational skill to better face disasters (Allahi et al. 2018, Tomasini and Wassenhove 2004). Over the past century, humanitarian support has been presented in the class of in-kind foods and services such as food and non-food items, water, medicinal consideration and temporary accommodation. Recently, a considerable literature has grown up around the theme of cash and voucher-based strategies for the procurement of humanitarian aid and are broadly considered a proper relief for in-kind assistance to the vulnerable people (Doocy and Tappis 2016).

Supporters of cash- and voucher-based approaches demonstrate that cash-based programs can be more cost-effective than commodity base alternatives which provide beneficiaries surpassing dignity, greater purchasing power, and high quality of living condition and improve their local economic activity. Although it is becoming extremely difficult to ignore the existence of problems like inflation and additional risks of insecurity and corruption (Harvey 2005).

In the world, near the 90 percent of low-income countries have in-kind transfer programs; however, half of them include a cash transfer program (Gentilini et al 2014). Although cash transfer programs have become a more significant part of social protection programs in the world, a superiority of welfare changes in both developed and developing countries (Tabor 2002, Currie and Gahvari 2008).

Cash and voucher distribution has been preferred for several reasons: this is evident in the case of (Cunha 2014, Currie and Gahvari 2008) which humanitarian organizations could encourage recipients to purchase and consume appropriate food or non-food items. Furthermore, the prices of resources, such as Doocy et al. 2006 and Schule et al. 2017, have declined over the cost-effectiveness of cash and voucher program for the emergency aid of humanitarian organizations. Also, Schule et al. (2017) have shown that vulnerable people have the food security with the cash in hand and could buy food in greater quantity and variety, and people have reported a greater sense of empowerment. In addition, with access to cash, beneficiaries shopped for food more frequently, which increased their access to fresh produce and perishables. Refugees could buy fresher, more healthful foods, increase their diet's variety, and save money.

Regarding numerous aspects of impact factors and the complex relationship among them, the recommended casual loop and system dynamic model illustrate how the methodology of system dynamics can be useful for understanding the behaviors of complex factors and consider them while humanitarian organizations make the decision for the cash transfer program in the disasters or emergencies situation (Briano et al. 2010, Revetria et al. 2008).

2. HUMANITARIAN RELIEF EFFORTS

The humanitarian cash and voucher's aid organizations need to recognize the most proper, effective way of helping disaster-affected families meet their own needs. In doing so, both positive and negative impact factors should be considered.

One of the most important factor that a wide range of papers mentioned is "dignity". There are several classified descriptions for dignity which other research proposed when cash and voucher donated to beneficiaries: financing household decision, greater purchasing power; improved meals both in terms of size and quality, less debt, access to health and to education services, self-sufficiency and resilience statuses, preferences, improvement of household welfare, facilitating individuals to earn and saving cash in bank accounts.

Although cash and voucher transfers often increased dignity impact, the beneficiaries interviewed remarked that in actually hopeless situations, the provided amount of cash and voucher assistance was usually too small for beneficiaries to feel dignity or improve the living condition (Berg 2013). However, based on the interviews, donation organizations were concerned that beneficiaries sometimes spend the transferred cash as "antisocial" expenditure. Antisocial spending indicates expenditures such as cigarettes, alcohol, drugs, prostitutes, and whatever related to cause injury to an individual. Generally, there was insufficient evidence of antisocial spending. Most interviewed beneficiaries spent the cash on items which meet the family requirement, although such spending is really difficult to prove. Nevertheless, in some cases like when men become able to buy their friends drinks, such antisocial expenditure seemed to have a positive psychosocial impact and did not withdraw from support for the family and enhancing their own situation in the community and bring good feeling in hard times. In other hand, women appear self-esteem from the expenditure at the beauty parlour. These findings suggest that overall, women are less likely than men to spend on antisocial activities and perform better in spending the aid cash.

Berg (2013) informed that in cases such as war or natural disaster where the questioned people had been discovered by great suffering, no evidence of dignity was perceived. While cash aid promoted some of the financial troubles, it did not come close to raise beneficiaries dignity or even meeting all their needs. In this situation, emergencies support their short coping strategies and then cash aid would increase recipient's dignity associated with a reduction in the employment of medium-term coping strategies.

Harvey (2005), reviews about the recipient's expenditure and the way of spending money in a wide range of recent cash programmes. It recommended a required attention to the use of cash for consumption and investment. In addition, it reported that some part of the money was spent to purchase necessary items for the beneficiaries' family in conditions with a high risk of inappropriate uses such as citrate or alcohol abuse. However, families were unlikely to report aid organizations staff their anti-social use. Usually, the most monitoring expenditure was on food, debt repayment, school fee and expenses, household goods, basic essential, health expenditure and investment.

This study attempts to address the question of why cash and vouchers are suitable for both donor's organizations and beneficiaries. By reviewing about 34 papers and reports of donor's organizations, most common factors that effect on recipients life and related local area after distributing cash and voucher are presenting in Table 1. There is a need to highlight that both columns review the hypothetical or possible advantage or disadvantage of cash distributing. What can be clearly seen in the table is effectiveness of cash and voucher because of high number of advantages in comparison with disadvantage ones.

The outcomes of evaluations of cash-transfer programs have been widely positive, showing considerable improvements to school enrolment rates and access to education service, decreases in rates of child labor. In addition, it made improvements in health and dietary diversity and in quality of consuming food.

Other impact factors such as social protection, debt repayment, improvement of inter-house relationship and work opportunity may also provide better condition life. It would also increase return rate of recipients to the home. For the aid agencies and government, factors like cost-effectiveness, local rehabilitation improvement of the economy are considerable. Furthermore, it is apparent from this table that very few negative factors like inflation could happen after the cash aid.

Table 1 Positive and negative impact factores of cash distribution

Positive impact factors	Explanation	Negative impact factors	Explanation
Dignity	Meeting people's priority need and improve household welfare, self-sufficiency, resilience statuses, and preferences	Inflation	A large injection of cash may have adverse effects on prices and local markets
Psychosocial well-being	Reducing people's stress levels	Lack of ability of women to control cash	Women typically have more control over food resources than cash in their households, cash could disempower women.
Shelter /accommodation	Living in permanent accommodation	Anti-social spending	Cash will be used for anti-social purposes such as consuming alcohol and …
Social protection	Have work permit and do not have legal issues and can be subdivided into three key component: social insurance, social assistance and standards	Gender based violence or discrimination	Risks of intra-household violence or tension as to who controls the cash
Food security	Dietary diversity or could buy food in greater quantity and variety	Exchange rate	Participants reported significantly higher relative expenditures on snacks/cigarettes for their households than female participants
Living condition	less likely to employ short-term coping strategies (seven days), medium-term coping strategy		
Health expenditure	Cash can increase spending pattern on health		
Education expenditure	Cash can increase spending pattern on education and lead to a reduction in the number of children missing school		
Poverty alleviation	Cash can decrease poverty because of high living condition		
Access to education services	Cash can increase access to services such as education		
Access to health services	Cash can increase access to services such as healthcare		
Investment	Recipient's savings in cash or bank accounts/ gold /other assets included electronics, appliances, and prevention of asset selling.		
Debt repayment	Beneficiaries have used the cash for debt repayment.		
Income	Increasing income security in the long term		
Cost effectiveness	Cash-based programs are likely to have lower transport and logistics costs, but there may be higher administrative costs		
Stimulating the markets	Functioning and rebuild accessible markets and local Rehabilitation for high demands.		
Local Economic	Cash can help to revitalize local economies by increasing the volume of trade and number of traders, thus developing local markets.		
Demands for goods /services	Cash make an increase in consumer demand created by the programme.		

Caring children by wemon	In households that had received cash transfers, mothers have more time to care their children		
Decrease child labour	Cash reduce child labour		
Intra-household roles	Reduce stress between parents and improves relationships within the household		
Beneficiaries' preferences	By improving the impact factors		
Security of cash and the fear of diversion	Security risks for the implementing organization in transporting cash to recipients		
Return rate of recipients	Cash will facilitate the people return to normalcy		
Decreasing of vulnerability	People with no family support, disabled, elderly or orphans, and large families and/or female-headed households.		
Work opportunity	Cash will increase the work opportunity		

The significant factors which evaluated by other researches are illustrated in Figure 1. The x axis shows the most common factors from Table 1 and the y axis related to percentage of research which have mentioned the impact factor on x axis. 34 papers have been studied and the most considered factors extracted. Obviously, the most significant factor is dignity which near the 50 percent of papers suggested. What other factors that can be clearly seen in the bar chart are food security, improvement of economy and better relations between couples. Some other factors such as expenditure on education, debt repayment and anti-social spending are in the middle attention of researchers. Overall, the effectiveness of cash aid can be observed by positive influence of cash aid on society.

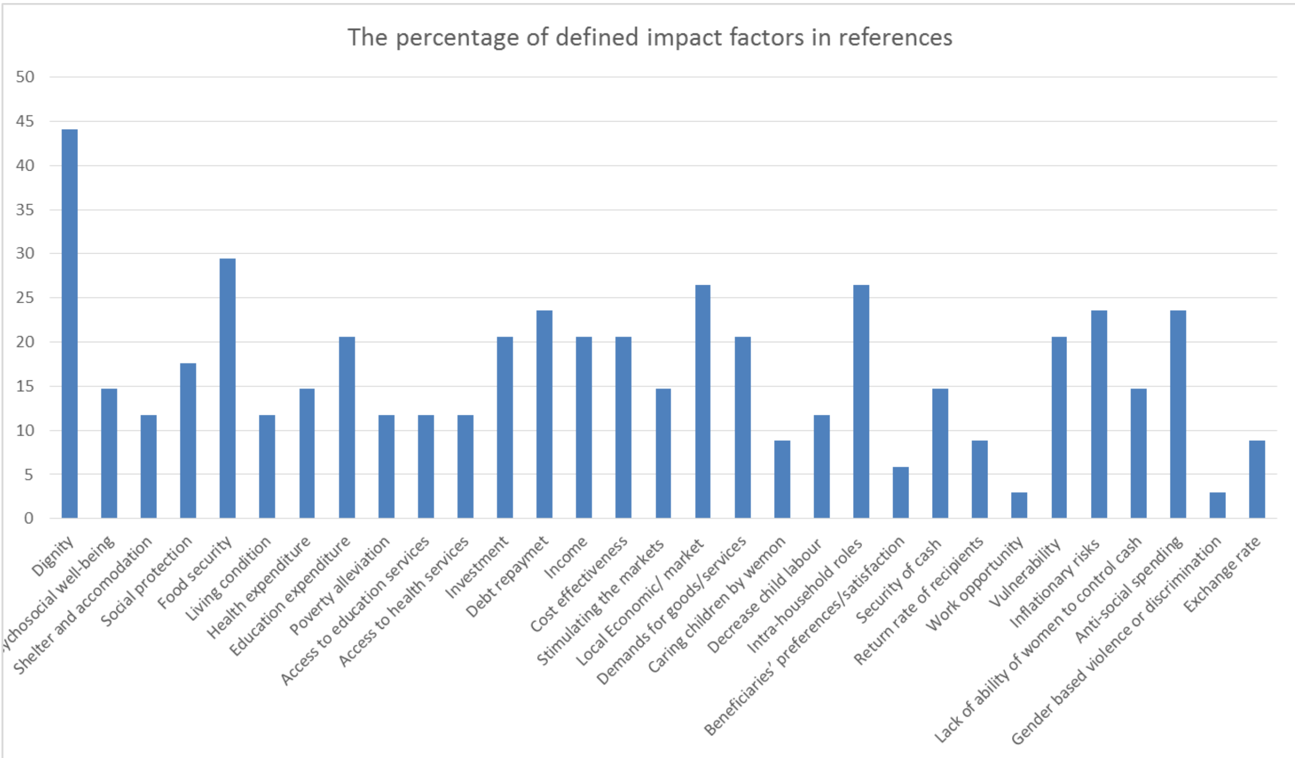


Figure 1 The perecentage of defined impact factors in references

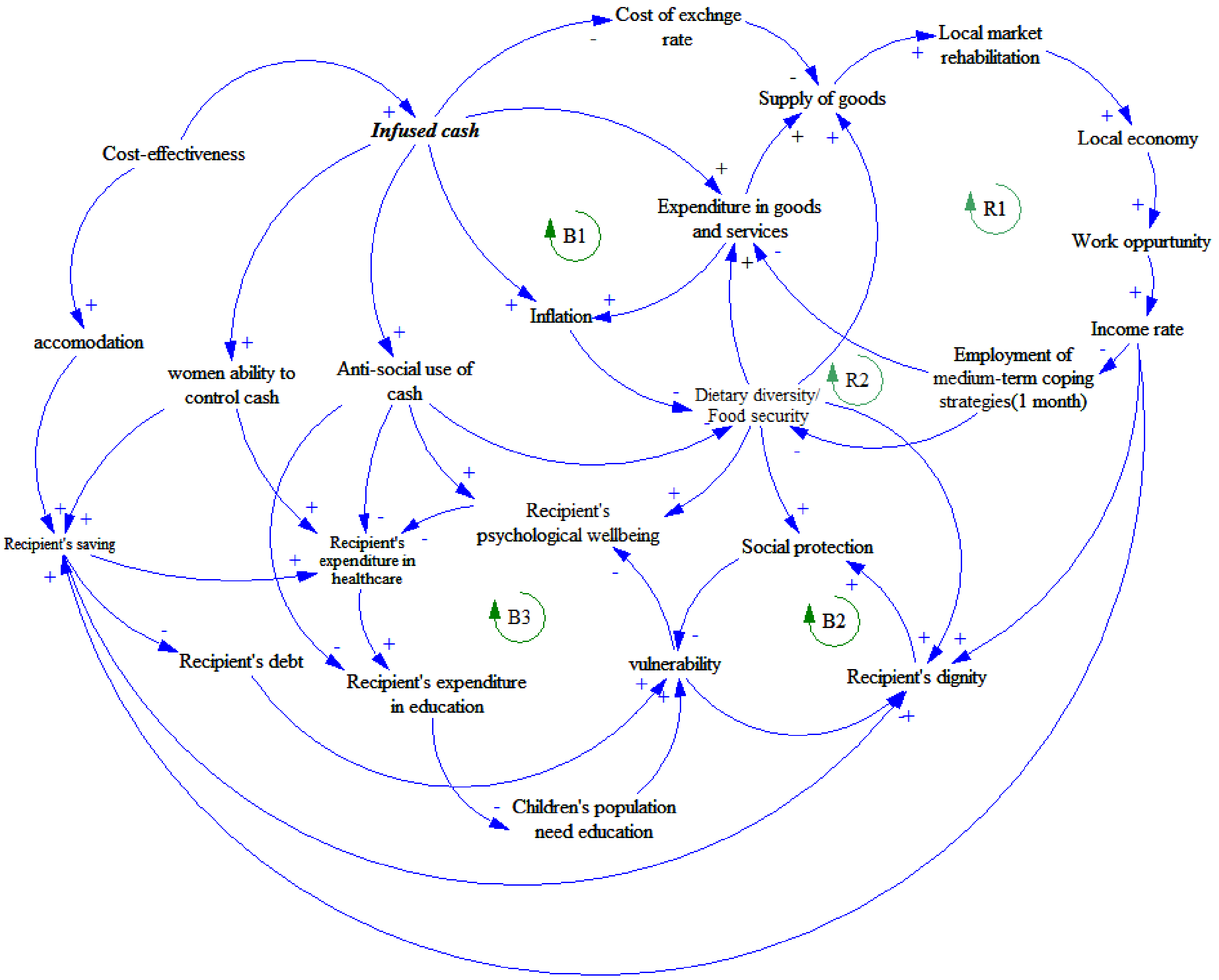


Figure 2 Overview of the humanitarian impact factors of disaster causal structure

3. GENERAL MODELING AND ANALYSIS PROCEDURE GUIDELINES

Causal loop diagrams (CLDs) are an essential tool for interpreting the feedback structure of systems. It has been used in academic achievement for a long time and frequently overused in organizations to quickly capturing assumptions about the causes of dynamics. The presented causal loop diagram in Figure 2, primarily interactions between the most important factors have been regarded, which is largely effect on other variables. It is a good illustration of social behavior in the disaster over time (the dynamics of the system) and can be interpreted by the interaction of positive and negative feedback loops. Five basic structure blocks; positive feedback or reinforcing loops (Positive loops), and balancing loops (negative feedback) are constructed the model. A causal diagram including of variables connected by arrows indicating the causal impacts amongst the variables. The significant feedback loops are also distinguished in the diagram. Link polarities represent the structure of the system and explain what would happen if there was a change. The details and behavior of the variables will not describe (Sterman 2000).

Figure 2 represents the influence of cash on beneficiaries and local area which people live there. Two balancing loops, B1 and B2, sustain the inflation impact when the cash injected to an area. Loop B3 causes expenditure in the health and education in response to changes in recipient's psychological wellbeing and reduction in number of children missing school. Vulnerability factor decrease to increases dignity of recipients as well as increasing in social protection factor that represents loop B2. The economy improvement can be seen in loop R1, which starts with raise in good's supply. Besides, the dietary diversity is influenced by increasing in expenditure in foods and helps to improve social protection. Cash aid can improve the income and payment of debt which cause reducing in the employment of medium term coping strategy (buying food on credit, taking an exploitative or degrading job, or withdrawing children from education). The most significant factors like expenditure on education, health and debt repayment would increase by injecting cash and improve in the live condition of recipients and may improve the level of dignity in suffers in emergency situation. Together these results provide important insights into by using cash aid program, recipient's life would be better and also other beneficiaries such as government and donor's organization can benefit from the cash transferring.

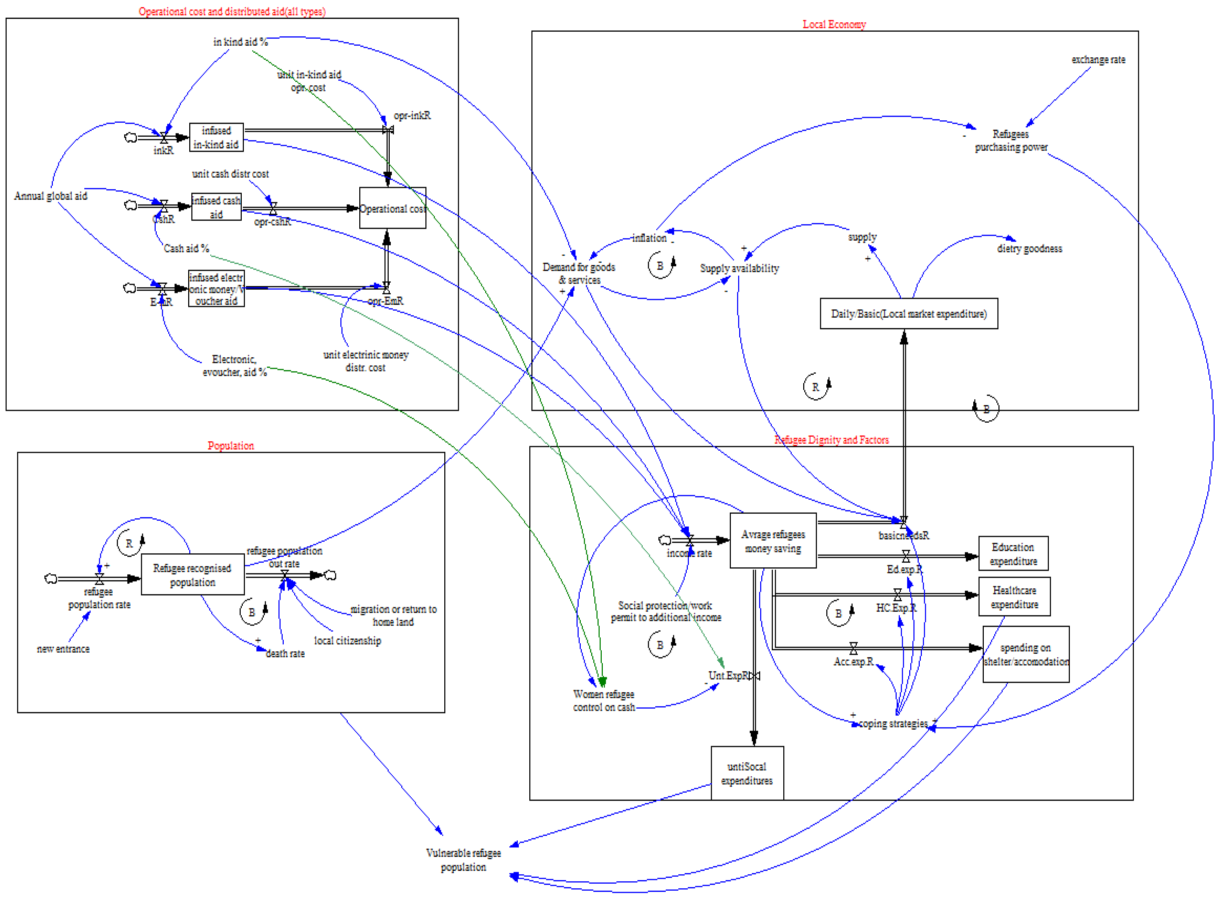


Figure 3 System dynamic model

Figure 3 shows system dynamic model of impact factors which is described in 4 categories; operational cost and distributed aid, local economy, population and refugee dignity and factors. The first category demonstrates the different way of humanitarian aid which is included cash, voucher and in-kind aid. The second one refers to improvement of local economy by cash transfer and the only negative factor can be inflation. The best result of third section would be immigration of people who affects by disaster and represents one of the significant factors of cash aid. The last category mentioned to the impact factors of cash aid on the vulnerable people which is described in casual loop model. For the model we used Vensim®, a modeling tool used to associate values with variables. The stocks are affected by flows, which either increase or decrease the stock, and are also associated with an equation, unit, and definition. The SD model allows us to move from a qualitative model to a quantitative model that can produce results as behaviors and values. In the case of this model, it allows us to see the interdependencies between our four layers.

This modeling technique allows the user to see more aspects of such a complex problem. It also allows the user to see how different layers affect one another.

The model has shown the ability of SD to model across various parameters. It has exemplified that complex problems can be modeled using system of systems modeling techniques. However, there were challenge of finding accurate data to input into the model that in the future would be present in the next paper.

4. CONCLUSION

A powerful collection of research has been provided to develop effects of cash and vouchers in a wide range of emergency situations. It is desirable to understand the outcomes on different kinds of beneficiaries including people, government and aid agencies. The present study was designed to determine the effect of cash and voucher in humanitarian aid with the factors analysis by casual loop and SD model. In summary, for the informants in this study, by cash distributing, people overwhelmingly spend money on basic essentials, health, education and using less coping strategies. Besides, injecting cash stimulates local economies and is usually more cost-effective than other kinds of aid programs.

These findings contribute in several ways to our understanding of important factors and the interventions of them among beneficiaries after the cash aid and provide a basis for further studies.

There is consequently an influential case for funding further in the accurate evaluation of cash and voucher-based programs, in order to be able to obtain the best decisions about their impact.

REFERENCES

Aker, J.C., 2017. Comparing cash and voucher transfers in a humanitarian context: Evidence from the Democratic Republic of Congo. The World Bank Economic Review, 31(1), 44-70.

Allahi F., Leeuw S.D., Sabet E., Kian S., Damiani L., Giribone P., Revetria R., Cianci R., 2018. A Review of System Dynamics Models Applied in Social and Humanitarian Researches. Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering 2018, 4-6 July, London, U.K.

Berg, M.I.C.H.E.L.L.E., Mattinen, H.A.N.N.A. and Pattugaian, G., 2013. Examining Protection and Gender in Cash and Voucher Transfers.

Briano E., Caballini C., Giribone P., Revetria R., 2010. Using a system dynamics approach for designing and simulation of short life-cycle products supply chain. In Proceedings of the 4th WSEAS international conference on Computer engineering and applications, Jan 27, pp. 143-149, World Scientific and Engineering Academy and Society (WSEAS).

Briano E., Caballini C., Mosca R., Revetria R., Testa A., 2010. Proposing a System Dynamic approach to assess and improve Italian ports competitiveness. Proceedings of ACMOS, May 29, pp. 29-31.

Centre for Research on the Epidemiology of Disasters (CRED), 2014. The international disaster database. <http://www.emdat.be/database>.

Cunha, J. 2014. Testing Paternalism: Cash versus In-Kind Transfers. American Economic Journal: Applied Economics.

Currie, J. and Gahvari F., 2008. Transfers in Cash and In-Kind: Theory Meets the Data. Journal of Economic Literature, 46(2), 333-383.

Doocy, S., Gabriel, M., Collins, S., Robinson, C. and Stevenson, P., 2006. Implementing cash for work programmes in post-tsunami Aceh: experiences and lessons learned. Disasters, 30(3), 277-296.

Doocy, S. and Tappis, H., 2016. Cash-based approaches in humanitarian emergencies, a systematic review April 2016.

Gentilini, U., Maddalena H. and Rsulan Y., 2014. The state of social safety nets 2014. Washington, DC: The World Bank Group.

Harvey, P., 2005. Cash and Vouchers in Emergencies. Humanitarian Policy Group Discussion Paper. London: Overseas Development Institute.

Revetria R., Oliva F., Mosca M., 2008. Modelling of Voltri terminal Europe in Genoa using system dynamic model simulation. In Proceedings of the 7th WSEAS international conference on System science and simulation in engineering, Nov 21, pp. 411-417. World Scientific and Engineering Academy and Society (WSEAS).

Schule F., Bitar J., Uekermann F., Taki M., Saidi M., al Omran S., Choufari B., Meerkatt H., 2017. Food-restricted voucher or unrestricted cash? How to best support Syrian refugees in Jordan and Lebanon. Conducted by the Boston Consulting Group.

Sterman J.D., 2000. Business Dynamics: systems thinking and modeling for a complex world, McGraw-Hill, Boston.

Tabor, S., 2002. Assisting the Poor with Cash: Design and Implementation of Social Transfer Programs. World Bank Social Protection Discussion Paper No. 0223.

Tomasini R.M. and Wassenhove L.V., A framework to unravel, prioritize and coordinate vulnerability and complexity factors affecting a humanitarian response operation, 2004. Working Paper, INSEAD, Fontainebleau, France.

AUTHORS BIOGRAPHY

Fahimeh Allahi is a visiting Ph.D. student in Industrial Engineering at the University of Loughborough, department of Wolfson. Her research field is focused on humanitarian supply chain management and business models.

Roberto Revetria is full Professor of the Department of Engineering Mechanics and Energetics of the Università degli Studi di Genova, Genova, Italy.

Roberto Cianci is full Professor of the Department of Engineering Mechanics and Energetics of the Università degli Studi di Genova, Genova, Italy.

A MANAGEMENT FRAMEWORK TO SUPPORT INDUSTRY 4.0 IMPLEMENTATION

Antonella Petrillo[(a)], Fabio De Felice[(b)]

[(a)] Department of Engineering - University of Naples "Parthenope", Italy
[(b)] Department of Civil and Mechanical Engineering - University of Cassino and Southern Lazio, Italy

[(a)]antonella.petrillo@uniparthenope.it, [(b)]defelice@unicas.it

ABSTRACT
Nowadays, Industry 4.0 has given way to exponential advances in technologies. The implementation of these technologies, should in theory, enable firms to improve efficiencies. But "*How to implement them*"? The aim of the present research is to propose an empirically model, based on Prado Maturity Model, to assess the Industry 4.0 maturity of Italian industrial enterprises and to structure a strategic growing business plan within the company. The final aim is to develop a "*user friendly*" but rigorous methodological approach that could be useful both from a scientific and from a practical point of view.

Keywords: Industry 4.0, decision support system, italian industry sectors, maturity model, SMEs

1. INTRODUCTION

Government institutions are promoting the computerization/digitalization of manufacturing process. Industry 4.0 is a name given to the current trend of automation and data exchange in manufacturing technologies (Kagermann *et al*., 2013). The name derived from the term "Industrie 4.0" originates from a project in the high-tech strategy of the German government, which promotes the computerization of manufacturing. Industry 4.0 is a very popular and extensively discussed topic in social groups, industry as well as in the academic field (Berger, 2014). The reason being, is that the potential benefits of Industry 4.0 are widespread, from a reduction in production costs to reductions in time to market introductions of new products (Oesterreich, and Teuteberg, 2016; Ortiz *et al.*, 2018). Furthermore, this "development" provides immense opportunities for the realization of sustainable manufacturing (Mumtaz *et al.*, 2018; Schaltegger and Wagner, 2011). More fundamentally, it will create new business models that will drive both evolutionary and revolutionary changes to the way in which we do business today (Petrillo *et al.*, 2017). Studies conducted by Koch *et al.* (2014) in the year 2014 in Germany show that operational efficiencies will increase by an average of 3.3% annually for the following five years leading to an average annual reduction in costs of 2.6% (McKinsey & Company, 2015). Many companies are performing their own case studies and analyses on the estimated benefits of Industry 4.0 (Wee *et al.*, 2015; Longo *et al.*, 2017, Vignali et al. 2018).

Despite the growing body of economic research on Industry 4.0, little attention has been paid to an examination of opportunities and challenges that are considered relevant for the implementation of Industry 4.0 (Kiel *et al.*, 2017; De Felice *et al.*, 2018). Furthermore, a suitable methodology for implementation of Industry 4.0 is however, less clear. In order, to accomplish this, a framework or implementation methodology of some sort is required to steer implementation efforts. This research was performed to identify and propose a specific framework useful to "measure" the maturity level in a manufacturing process identifying the weaknesses in the management of their strategy. Definitely, the aim of the present research is to propose an empirically model and its implementation to assess the Industry 4.0 maturity of Italian industrial enterprises. The final aim is to develop a methodological approach that could be useful both from a scientific and from a practical point of view. The proposed approach aims to integrates the strengths of existing maturity models into a global model in order to help companies to develop their individual transformation process. In other words, the maturity model aims to support companies in estimating their current situation regarding an examined field of action (Gap Analysis). Evaluation of maturity is carried out through a standardized questionnaire consisting of one closed-ended question per item. The model has been developed using the main principles of the Prado Maturity Model that was launched in December 2002 and reflects forty years of experience on the subject by Darci Prado within IBM and two large, Brazilian international consulting firms (das Leite das Neves *et al.*, 2013; Barber, 2004). The Prado Model was developed to measure the Maturity in Project Management. Thus, our model reflects the principles of the Prado Model but has been modified and adequate to our specific needs.

The rest of the paper is structured as follows. Section 2 discusses the Industry 4.0 paradigm in the Global and Italian context; Section 3 presents the theoretical procedure model; Section 4 introduces a specific case study regarding an Italian company and discusses the main results of the study. Finally, in Section 5 main findings, limitations of the model and future research are analyzed.

2. INDUSTRY 4.0 IN THE GLOBAL AND ITALIAN CONTEXT

The attention towards themes connected to Industry 4.0 has been growing in recent months in Italy as well as in Germany. In fact, Germany and Italy are the two biggest manufacturing countries in Europe.

The 13 countries in which the Industry 4.0 programs are already active are: Austria (Industrie 4.0 Oesterreich); Belgium (Made different - Factories of the future); Czech Republic (Průmysl 4.0); Germany (Industrie 4.0); Denmark (Manufacturing Academy of Denmark - MADE); Spain (Industria Conectada 4.0); France (Alliance pour l'Industrie du Futur); Hungary (IPAR4.0 National Technology Initiative); Italy (Industry 4.0); Holland (Smart Industry); Portugal (Indústria 4.0); Sweden (Smart Industry); Luxembourg (Digital for Industry Luxembourg). The nine in which the works are in the finishing line are Bulgaria, Croatia, Finland, Poland, Romania, Slovakia, Slovenia and the United Kingdom (as shown in Figure1).

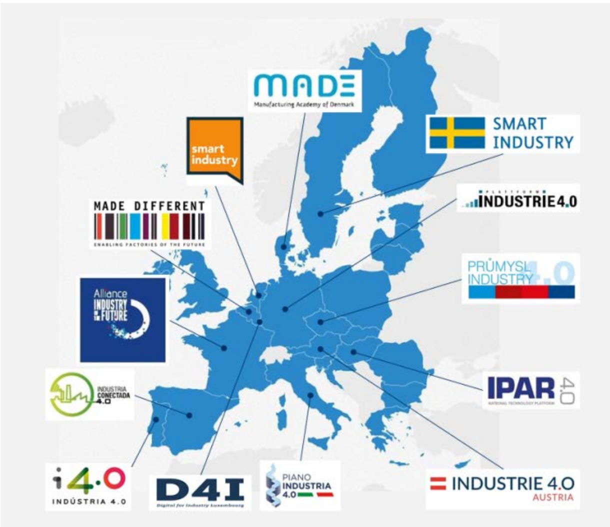


Figure 1: Industry 4.0 strategies in the world

Italy, as a manufacturing country and supplier of various major industries throughout the world, cannot miss the train of the 4th industrial revolution (PricewaterhouseCoopers, 2016). In other words, it is important to exploit the benefits which come from cooperation across European member-states and from new ways of cross-sectoral cooperation. Facing these opportunities and challenges, the Italian Ministry of Economic Development and the German Ministry for Economic Affairs and Energy agree, that Germany and Italy will cooperate in the field of Industry 4.0. The cooperation focuses on three key topics: 1) *Standardisation*; 2) *Support and engagement of SMEs* and 3) *Skills empowerment*. To this end, in Italy, the government launched the so called *Piano Calenda* or Piano Industria 4.0 or in English "National Industry 4.0 Italian Plan" to promote Industry 4.0. It represents a great opportunity for companies operating in Italy. The plan aims to raise tax incentives for investments in goods and technologies that connect physical and digital systems. Therefore, the implementation of the Industry 4.0 program has to be done properly (Erol *et al.*, 2016; Zhou, 2013).

Figure 2 shows that the Italian companies see more space for introduction of Industry 4.0 program.

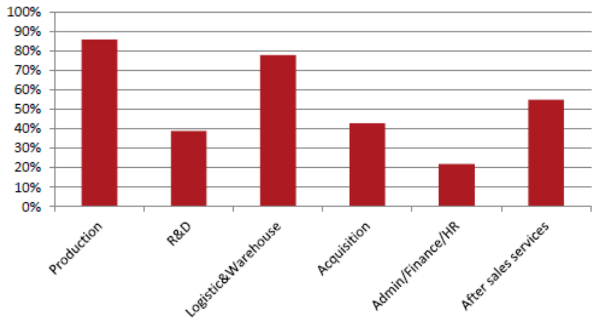


Figure 2: Italian companies see more space for introduction of Industry 4.0 program (Source: Staufer.Italia Study)

While, Figure 3 points out the segmentation of the Italian market of Industry 4.0.

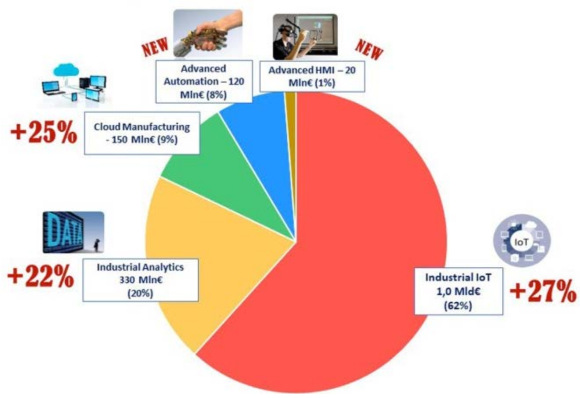


Figure 3: Italian market of Industry 4.0 (Source: Osservatori.net)

In the European context, in June 2017, *Platform Industrie 4.0 in Germany*, *Alliance Industrie du Futur in France* and the *Piano Nazionale I4.0* in Italy, thanks to their leading role in the sector European manufacturing industry, have initiated an intergovernmental collaboration agreement called **Trilateral**, to which voluntary standardization competes, to favor transformation digital sector through a joint action plan and promote a "*Smart Manufacturing*" **Made in Europe**.

However, the concept of "Industry 4.0" is still not very familiar within Italian companies (Rusche *et al.*, 2016).

In detail, two main concerns became apparent. The first being the lack of methodology towards identifying the areas that need to be addressed in implementing Industry 4.0, and the second problem is not knowing

how to practically implement Industry 4.0. These two main concerns can be broken down into three main questions. These questions are: 1) *How can SMEs implement Industry 4.0*? 2) *What does it need to attain them*? And 3) *What does it require to implement it*?
The framework developed in this research attempts to answer the three questions posed above.

3. INDUSTRY 4.0 FRAMEWORK DEVELOPMENT: RI4.0 MODEL

Experiences with various companies have shown that the objectives for the Italian industries are to improve the following areas:

1. **Flexbility:** More flexibility through small batches production with large scale costs;
2. **Speed:** Higher speed from prototype to mass production though innovative technologies;
3. **Productivity:** Increased productivity through reduced setup times, errors and downtime;
4. **Quality**: Better quality and less scrap through sensors to monitor production in real time;
5. **Product competitiveness:** Higher product competitiveness through increased Internet of things functionalities.

However, they are not able to relate each of them to their business Industry 4.0 strategy.
Thus, an Industry 4 readiness assessment tool has been developed in this research. Its purpose is to provide a simple and intuitive way for companies to start to assess their readiness and their potential of industry 4.0 age.
The readiness assessment tool considers the above *5 core dimensions*. The readiness assessment tool has been developed as a *Diagnosis Tool* in order to verify the positioning of the company on "Best Practices" characterizing the process in terms of innovation, digitalization, technologies, tools, and organizational choices. The final out is the elaboration of *2 profiles*: **Current Performance State** and **Desired Performance State**. The comparison of current performance state with desired performance state defines a *Gap Analysis.* Gap analysis provides a foundation for measuring investment of time, money and human resources required to achieve the level of readiness from industry 4.0 perspective (see Figure 4).

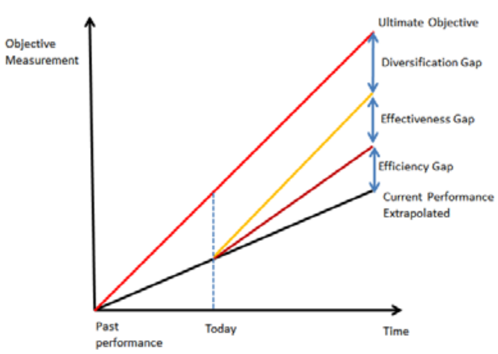


Figure 4: Gap Analysis

The readiness assessment tool is divided into three phases, as described below:

Phase I: **Current State Performance.** The aim of the present phase is to assess the current positioning of the company regarding the Industry 4.0 paradigm or in other words the *"Current State performance"*. At this end, in the present phase a set of "best practices" that if used correctly allow the achievement of performance targets according to Industry 4.0 paradigm was identified. In particular, 10 Best Practices (BPs) have been identified (as explained in the following Section 4). To each BPs have been assigned level of importance using a growing scale from 1 to 4 (Likert Scale), as follows:

- Score 1: there is no strategy dedicated to the implementation of industry 4.0;
- Score 2: the strategy is lacking;
- Score 3: the strategy exists and is evolving;
- Score 4: the strategy exists and is very clear.

Phase II: Readiness State. The measurement of the readiness or in other word of the maturity level of the company regarding market/strategy I4.0 is determined through **Questionnaire A**. The Questionnaire is based on Prado Maturity Model (Prado, 2006) and is comprised of 5 levels of maturity and 6 dimensions (*Strategic Alignment; Behavioral Competence; Organizational Structure; Informatization; Methodology; Technical Competence).* The 6 level are defined as follows*:*

- **Level 1** represents the initial stage, where the department hasn't made any coordinated effort to implement Industry 4.0 strategy;
- **Level 2 "Known"** demonstrates that the organization regularly invests in training and in innovation/digitalization (new manufacturing technologies);
- **Level 3 "Standardized"** has seen the implementation of the Industry 4.0 strategy, which has standardized the use of procedures that require the utilization of planning and innovation processes;
- **Level 4 "Managed"** shows that investments in innovation/digitalization competency are efficient because the managers are better prepared to handle the aspects of their teams, such as Behavioural Competence, Information System, Technical and Contextual Competence etc;.
- **Level 5 "Optimized"** indicates that the company has reached a high level of innovation management understanding.

The Questionnaire consists of 55 statements (the first 15 are referred to the organization in general) for which an assessment is requested addressing four levels of maturity (from 2 to 5). Each level has 10 questions with 5 choices (A, B, C, D, and E), which weigh 10, 7, 4, 2, and 0, respectively.

Proc. of the Int. Conference on Modeling and Applied Simulation 2018,
ISBN 978-88-85741-07-2; Bruzzone, De Felice, Frydman, Longo, Massei and Solis, Eds.

The final score of the maturity assessment is obtained from the answers and data of the **Questionnaire A**, according to Equation (1)

Final evaluation = (100 + Total_Points) / 100 (1)

This final score is given on a scale of 1 through 5, which can be interpreted according to illustration in Figure 5 and as follows:

- Up to 1.60 Very Weak;
- Between 1.60 and 2.60 Weak;
- Between 2.60 and 3.20 Average;
- Between 3.20 and 4.00 Good;
- Between 4.00 and 4.60 Very Good;
- Above 4.60 Excellent.

The correlation between the 6 dimensions and the 5 levels shows how mature the project management of an institution is. Standards for **Level Adherence Assessment** are:

- Up to 20% Weak Adherence;
- Up to 40% Average Adherence
- Up to 70% Good Adherence;
- Up to 90% Very Good Adherence;
- Up to 100% Excellent Adherence.

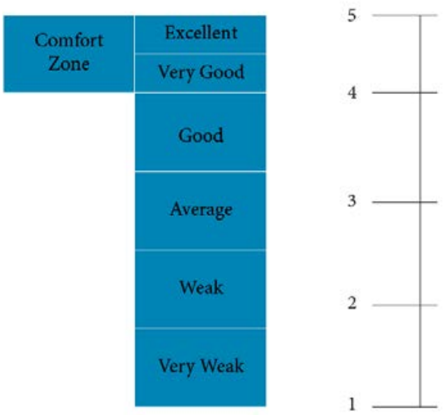


Figure 5: Prado Scale

Phase III: Desired State Performance

From the comparison between Current State Performance and Desired State Performance differences arise between the two profiles or the Gap Analysis.

4. CASE-STUDY: AN ITALIAN INDUSTRIAL ENTERPRISE

In this paragraph the methodological approach is applied to a real case study. In detail, in the following, results obtained from a case-study with a Southern Italian manufacturing enterprise (Company X) with around 200 employees operating in an aerospace sector are presented.

Phase I: Current State Performance

Phase I is the most critical phase because in the present phase it was necessary to define the BPs. At this end, we identified them through the National Industry 4.0 Italian Plan. According to the Italian Government, the Industry 4.0 paradigm is defined as follows "*Production process able to manage and circulate information related to the generation of added value between the various components of the production system - machines, human beings, products, IT systems – interconnected*".

Thus, a specific set of 10 BPs have been identified, as follows:

- BP#1: Incentive investments in technologies and goods I4.0;
- BP#2: Spreading the culture I4.0;
- BP#3: Developing skills I4.0;
- BP#4: Funding R&D I4.0;
- BP#5: Strengthen and innovate the presence of international markets;
- BP#6: Participation in the Competence Center and Digital Innovation Hub;
- BP#7: Use IoT interoperability standards;
- BP#8: Launch of innovative projects and technological development;
- BP#9: Support for experimentation and "in vivo" production of new I4.0 technologies;
- BP#10: Legal model and adequate managerial skills.

An experts team of academic and company managers composed of 1 academic, expert on Piano Nazionale I4.0; 1 Chief Executive Officer (CEO); 1 ITC Manager; 1 Process Manager; 1 Administration Manager assessed the current state of implementation of the10 BPs using a 1- 4 scale. Table 1 show the result of the assessment (see Appendix). The final score is the geometric mean of the judgments of the 5 experts.

Phase II: Readiness State

The problem that arises now is to identify the Readiness State for each of the 10 BPs. The questionnaire was administered to Chief Executive Officer (CEO), in order to define the maturity level of the market/strategy I4.0. Table 2 shows a Sample of questionnaire for level 2 KNOWN (see Appendix). Similarly, were developed for Questionnaire per each level. Adherence to the four levels proposed by the model used in the research is shown in Figure 6 (see Appendix).

An analysis of Figure 6 shows that Level 2 called "Known", shows the most adherence. According to Prado, it demonstrates isolated attempts to standardize procedures. The result indicates that the company needs to improve its strategy.

Figure 7 (see Appendix) represents a comparison of the level of maturity among the company X, the National target and the Regional target. Comparing the score obtained through this case study with the overall average scores of Italian companies (national and regional context) it is possible to note that the company is more in line with the regional average of in the same category.

Phase II: Desired State Performance.

Gap Analysis is shown in Figure 8. The Gap Analysis is the result of the comparison of Current State Performance and Desired State Performance. The Desired State performance is the result of the interrogation of the 5 experts after 6 months from the evaluation of the Current State and after the evaluation of the Readiness State.

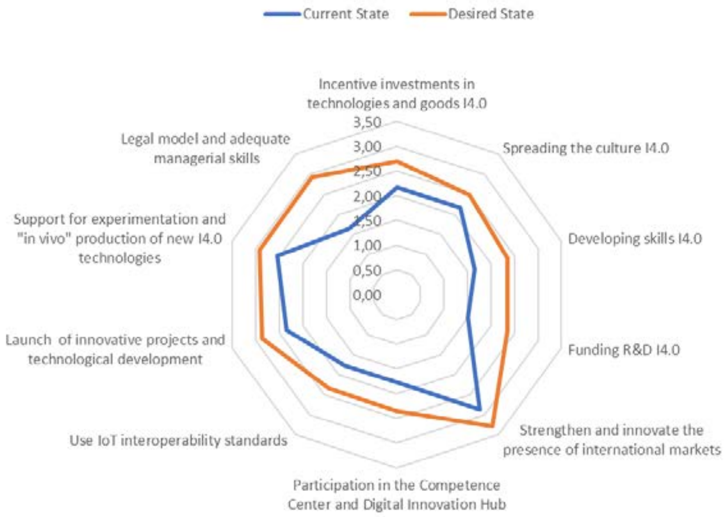


Figure 8: Gap Analysis

In summary, in our opinion to assure that the strategy of companies can evolves towards Industry 4.0 principles it is necessary: 1) use best practices for I4.0; 2) ensure a continuous improvement and 3) invest in technological and processes innovation.

5. CONCLUSION

Italian companies typically machinery providers and system integrators are facing the risks associated with an increasing complexity of the systems that combine the physical world and digital. The digital transformation, once completed, will result in complete automation and interconnection towards a more technological reality; it is not simply a question revolution that involves companies, but the everyday life of all of us. One of the contributing factors to this transformation is the lack of specific tools that can support the enterprises to pursue the advantages of this “new” revolution. Thus, the research work presented here aimed for the development of a maturity model and a related tool for assessing the Industry 4.0 maturity of manufacturing enterprises. The main goal of the research was to create a simple and easy to use model that provides reliable results. We believe that the implementation of an efficient management approach is capable to reach “solid” results. The scores obtained through this research, based on the Prado Model, allowed for the identification of the level of maturity of the company in comparison of national and regional context. Future research activities will mainly aim to expand the sample of interviewed companies and to improve the accuracy of the maturity items. Furthermore, based on the findings of this “pilot” maturity model, our aim is to define a more domain specific model for the assessment of Industry 4.0 maturity in different sectors.

ACKNOWLEDGMENTS

This research represents a result of research activity carried out with the financial support of University of Naples Parthenope, namely “Bando di sostegno alla ricerca individuale per il triennio 2015-2017. Annualità 2017”.

REFERENCES

Barber, E., 2004. Benchmarking the Management of Projects: a Review of Current Thinking”, International Journal of Project Management, Vol. 22, No. 4, pp. 301-307.

Berger, R., 2014. Industry 4.0 - The new industrial revolution – How Europe will succeed. Roland Berger Strategy Consultants.

De Felice, F., Petrillo, A., Zomparelli, F., 2018. Prospective design of smart manufacturing: An Italian pilot case study. Manufacturing Letters, 15, pp. 81-85.

Erol, S., Schumacher, A., Sihn, W., 2016. Strategic guidance towards Industry 4.0 – a three-stage process model, in Proc. of Interna-tional Conference on Competitive Manufacturing 2016 (COMA16), Stellenbosch, South Africa.

Kagermann, H., W. Wahlster and J. Helbig, eds., 2013: Recommendations for implementing the strategic initiative Industrie 4.0: Final report of the Industrie 4.0 Working Group.

Koch, V., Kuge, S., Geissbauer, R. Schrauf, S., 2014. Industry 4.0: Opportunities and challenges of the industrial internet. Tech. Rep. TR 2014-2, PWC Strategy GmbH, United States, New York City, New York (NY).

Kiel, D., Müller, J., Arnold, C., Voigt, K.-I., 2017. Sustainable Industrial Value Creation: Benefits and Challenges of Industry 4.0. Int. J. Innov. Manag., 21, 1740015.

Longo, F., Nicoletti, L., Padovano, A., 2017. Smart operators in industry 4.0: A human-centered approach to enhance operators' capabilities and competencies within the new smart factory context. Computers and Industrial Engineering, vol. 113, pp. 144-159.

McKinsey & Company, 2015. Industry 4.0: How to navigate digitization of the manufacturing sector. Tech. rep., McKinsey and Company, New York City, New York (NY).

Leite das Neves, L.A., Nicolini do Patrocínio Nunes, L.E., Alves Corrêa, V., Cerqueira Rezende, M., 2013. Application of the Prado - Project Management Maturity Model at a R&D Institution of the Brazilian Federal Government. J. Aerosp. Technol. Manag., São José dos Campos, Vol.5, No 4, pp.459-465.

Mumtaz, U., Ali, Y., Petrillo, A., 2018. A linear regression approach to evaluate the green supply chain management impact on industrial organizational performance. Science of the Total Environment 624, pp. 162-169.

Oesterreich, T.D., Teuteberg, F., 2016. Understanding the implications of digitisation and automation in the context of Industry 4.0: A triangulation approach and elements of a research agenda for the construction industry. Comput. Ind., 83, 121–139.

Ortíz, M.A., Betancourt, L.E., Negrete, K.P., De Felice, F., Petrillo, A., 2018. Dispatching algorithm for production programming of flexible job-shop systems in the smart factory industry. Annals of Operations Research, 264(1-2), pp. 409-433.

Petrillo, A., De Felice, F., Zomparelli, F., 2017. Performance measurement for world-class manufacturing: a model for the Italian automotive industry. Total Quality Management and Business Excellence pp. 1-28.

Prado, D., 2006. Project Management Maturity Model (Prado-PMMM): INDG-Tecs.

PricewaterhouseCoopers, 2016. The Industry 4.0 / Digital Operations Self Assessment.

Rusch, M., Treusch, O., David, U., Seiter, M., 2016, Industrie 4.0 – Controllers Aufgaben. In: Controller Magazin, 3, pp. 70-79.

Schaltegger, S., Wagner, M, 2011. Sustainable entrepreneurship and sustainability innovation: categories and interactions. Business Strategy and the Environment, 20(4), 222-237.

Vignali G., Bertolini M., Bottani E., Di Donato L., Ferraro A., Longo F., 2018. Design and testing of an augmented reality solution to enhance operator safety in the food industry. International Journal of Food Engineering, vol. 14, no. 2.

Wee, D., Kelly, R., Cattel, J., Breunig, M., 2015. Industry 4.—How to Navigate Digitization of the Manufacturing Sector; McKinsey & Company: Dusseldorf, Germany; Berlin, Germany; Munich, Germany.

Zhou, J., 2013. Digitalization and intelligentization of manufacturing industry," Adv. Manuf., vol. 1, no. 1, pp. 1–7, Mar. 2013.

AUTHORS BIOGRAPHY

Fabio De Felice, Professor at the University of Cassino and Southern Lazio, board member of several international organizations. 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 multi-criteria techniques to decisions (Analytic Hierarchy Process, Analytic Network Process), to RAMS Analysis and Human Reliability Analysis.

Antonella Petrillo, Degree in Mechanical Engineering, Phd at the University of Cassino and Southern Lazio. Now Researcher at the University of Naples "Parthenope" (Department of Engineering) where she conducts research activities on Multi-criteria decision analysis (MCDA), Industrial Plant, Logistic and Safety.

APPENDIX 1

Table 1: 10 Best Practices – Current State

BP	Best Practices	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Final Score
1	Incentive investments in technologies and goods I4.0	2	1	3	4	2	**2,17**
2	Spreading the culture I4.0	1	2	3	2	4	**2,17**
3	Developing skills I4.0	3	1	2	1	2	**1,64**
4	Funding R&D I4.0	1	1	2	2	2	**1,52**
5	Strengthen and innovate the presence of international markets	2	2	3	4	4	**2,86**
6	Participation in the Competence Center and Digital Innovation Hub	3	2	1	1	3	**1,78**
7	Use IoT interoperability standards	2	3	3	1	1	**1,78**
8	Launch of innovative projects and technological development	3	2	1	3	4	**2,35**
9	Support for experimentation and "in vivo" production of new I4.0 technologies	1	3	4	3	3	**2,55**
10	Legal model and adequate managerial skills	2	3	2	1	1	**1,64**

Table 2: Sample of questionnaire for level 2 KNOWN (10 items)

LEVEL 2 – KNOWN - Question 1		
Regarding internal and external training occurred in the last 12 months related to basic aspects of Industry 4.0, select the most appropriate option	**Answer**	**Points**
Many staff participated in training in the last 12 months	A	10
The situation is slightly inferior than that described in option A	B	7
The situation is significantly inferior than that described in option A	C	4
There is some effort in this direction	D	2
There is no effort in this direction	E	0

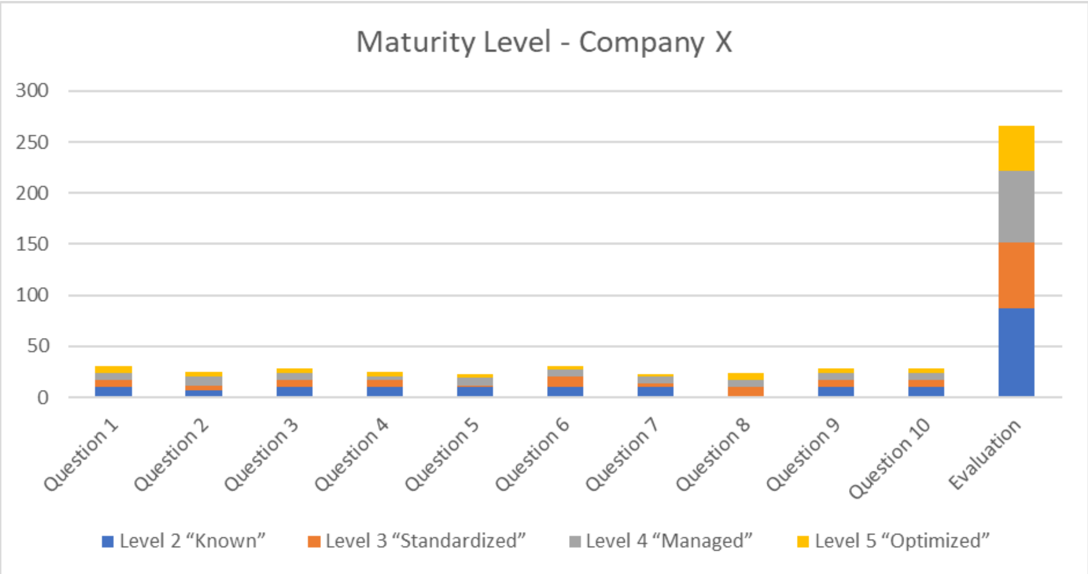


Figure 6: Maturity Level for Company X

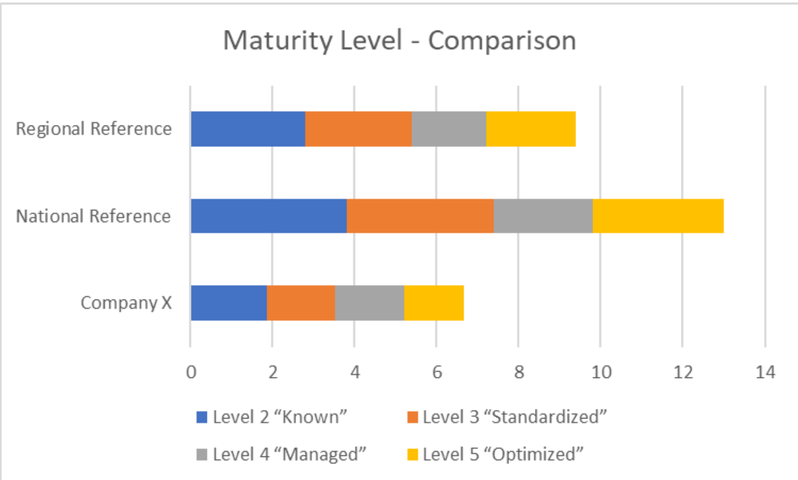


Figure 7: Maturity Level Comparison

RULE-BASED MODELING OF SUPPLY CHAIN QUALITY MANAGEMENT

Juan Miguel Cogollo Flórez[(a)], Alexander Alberto Correa Espinal[(b)]

[(a)]Instituto Tecnológico Metropolitano-ITM, Medellín, Colombia
[(b)]Universidad Nacional de Colombia, Medellín, Colombia

[(a)]juancogollo@itm.edu.co, [(b)]alcorrea@unal.edu.co

ABSTRACT
Analytical modeling of Supply Chain Quality Management (SCQM) is one of the main research avenues in both Quality Management (QM) and Supply Chain Management (SCM). Therefore, the purpose of this paper is to analyze rule-based modeling for SCQM coordination and integration and to develop a model for SCQM integration using Fuzzy Cognitive Maps. The model contains nine SCQM concepts and their relations. The convergence analysis of the inference process validated the initial selection of the SCQM concepts and the values of their relations. The results of the model allowed identifying the concepts rejections, returns, and defective product in production, as the most important decision making variables for improvement quality management in the supply chain studied.

Keywords: Modeling, Supply Chain Quality Management, rule-based models, Fuzzy Cognitive Map.

1. INTRODUCTION

Supply Chain Quality Management (SCQM) refers to the coordination and/or integration strategies along the supply chain, in order to improve quality and overall performance (Kuei and Madu 2001; Robinson and Malhotra 2005; Foster 2008; Mellat-Parast 2013; Flynn and Zhao 2015).
Evans, Foster and Linderman (2014) carried out a study on the content and research trends in quality management and established that current avenues of research are focused on development of mathematical models, global contingency analysis, supply chain quality, information technologies and strategic benefits. In addition, the authors stated that few works have been developed on quality performance in the supply chain and how cooperation can lead a better performance.
Although products quality and safety would be considered as overcome issues by applying the current standardization systems, in recent years there have been frequent events of recalls due to safety and/or quality failures in automotive, toy, computer and electronic equipment manufacturers (Flynn and Zhao 2015). So, the need for researching mechanisms and strategies of SCQM is a current issue.
Therefore, it is necessary development and implementation of rule-based models for SCQM coordination and integration. In order to contribute to generation and dissemination of knowledge in this area, this paper describes a rule-based model for SCQM integration using FCM.
This paper is structured as follows: The next section provides a general definition of SCQM coordination and integration. Then, the results of a literature review on SCQM rule-based modeling are shown. After, there is a brief description on modeling with Fuzzy Cognitive Maps. Lastly, the results of developing the model for SCQM integration using FCM are presented and analyzed as part of the conclusions.

2. SUPPLY CHAIN QUALITY MANAGEMENT COORDINATION AND INTEGRATION

In SCM, Coordination is carried out by sending the correct signals or sharing the correct information and the same policies (Montoya-Torres and Ortiz-Vargas 2014). Likewise, Integration refers to "making a whole, gathering its constituent parts", that is, synchronizing requirements, concepts and flows of the chain members aimed to maximize competitive advantages at strategic, tactical and operative levels (Bautista-Santos et al. 2015).
Based on above, SCQM coordination may be considered as the Quality Management practices aimed to improve supply chain overall performance and the SCQM integration as development of a single point of view (a common language) in all supply chain members to synchronously manage issues related to Quality Management.

3. SCQM RULE-BASED MODELING: A REVIEW

Formal modeling (analytical or mathematical) of a system may be basically using two approaches: Descriptive modeling and Rule-based modeling. Descriptive modeling is static and consists of describing the real state of a system at any time using quantitative methods; it is usually used in the initial stages of research of a complex system. Rule-based modeling consists of formulating dynamic rules that explain the behavior observed in a system and allows to predict future states of the system (Sayama 2015).
A comprehensive Literature Review was carried out, in order to identify the most important contributions that have been made concerning to SCQM rule-based

modeling, over the last 18 years. The used methodology allowed us to seek, identify, interpret and synthesize the documented evidence in the time period (2000 to 2018) and to identify research topics.
In order to identify how SCQM may be coordinated and integrated through rule-based models, the following research question was formulated: What rule-based models for SCQM coordination and/or integration have been proposed in the period between 2000 and 2018?
Then, the search was carried out considering a compendium of the following databases: Science Direct, IEEE, Springer, Taylor & Francis and Emerald. By applying the search criteria and considering the selected databases, 357 publications were obtained.
The selection of the documented evidence concerning to the research question was carried out considering two inclusion criteria: (i) It presents a rule-based model for SCQM analysis, and (ii) It presents a rule-based model for SCQM coordination and/or integration. Moreover, the papers that met any of the following characteristics (exclusion criteria) were excluded: (i) It presents theoretical development on SCQM, and (ii) It presents a descriptive model for SCQM coordination and/or integration.
The 357 publications obtained in the search process were analyzed by application of the above inclusion and exclusion criteria and, then, 15 papers were selected. Based on analysis of the selected papers, a profile of the SCQM Rule-Based Modeling was made (Figure 1), according to applied tool, SCQM objective (coordination and/or integration) and authors.

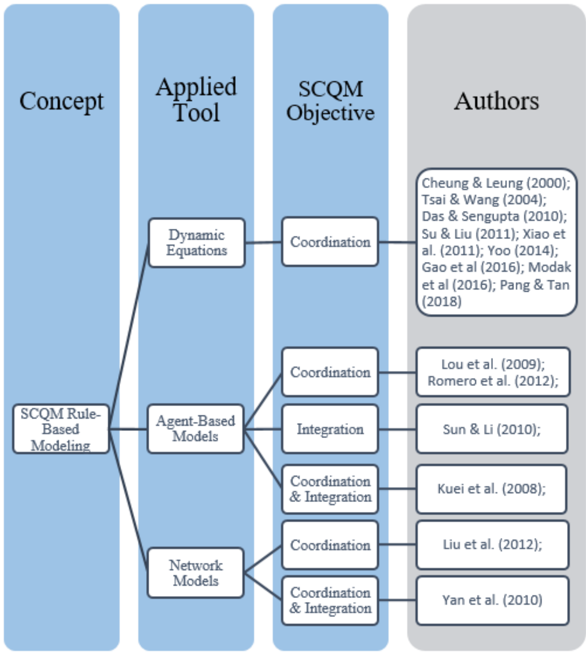


Figure 1: Profile of the SCQM Rule-Based Modeling according to the selected papers.

The 15 articles selected and confirmed as relevant as per the literature review are visually represented in Venn Diagram form in Figure 2 in line with the SCQM objective of the models proposed in them. Most of the rule-based models were for SCQM coordination (80%). The 13% of the models were for SCQM coordination and integration. On the other hand, only 7% of the models were for SCQM integration. Based on the above, the need for developing rule-based models for SCQM coordination and integration is evident.

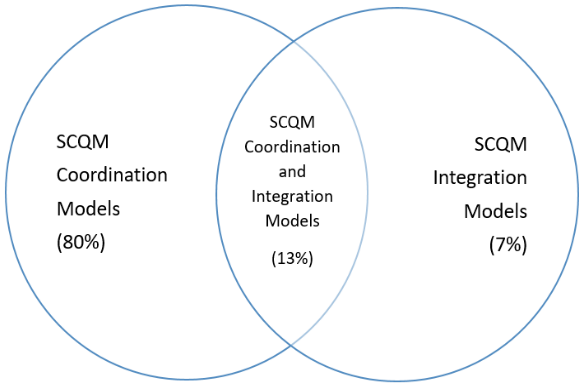


Figure 2: Representation of percentage of articles identified proposing ruled-based models for SCQM coordination and/or integration.

The following paragraphs detail the main contributions identified in the literature in SCQM rule-based modeling, based on the profile of the Figure 1.

3.1. Dynamical Equations Models

The Dynamical Equations models consider how the states of components change over time through their interactions with other nodes that are connected to them. Cheung and Leung (2000) developed a coordinated replenishment model considering inventory control and quality control problems, with a two-item inventory and quality costs depending on the costs acceptance sampling plans. Tsai and Wang (2004) proposed a model of collaborative quality control in the semiconductor industry, considering practices and protocols to ensure quality of outsourced and offshored processes.
Das and Sengupta (2010) developed a mathematical model for supply chain design considering suppliers recruitment with appropriate contractual agreements and quality assurance systems. Su and Liu (2011) developed a SCQM coordination model by contracts and information exchange, considering external failure sharing in a supplier-manufacturer supply chain.
Xiao et al (2011) proposed mechanisms for coordinating quality assurance policies in Make To Order (MTO) environments, considering one manufacturer and one retailer via a revenue-sharing contract. Yoo (2014) developed a model for coordinating return policy and product quality decisions in supply chains under risk aversion of a supplier.
Gao et al (2016) proposed a coordination model for quality improvement in a two-stage decentralised supply chain with a partial cost allocation contract. Modak et al (2016) proposed a hybrid contract mechanism for the coordination of a manufacturer–distributer–duopolistic

retailers supply chain for a product, where the manufacturer supplies lotsize with a random portion of imperfect quality. Pang and Tan (2018) proposed a model for making quality decisions of a single product in a supply chain with a supplier and two competing manufacturers. The model examines the optimum quality strategies under different cooperative mechanisms and investigate its effects on channel members' profits.

3.2. Agent-Based Models

Agent-based models are computational simulation models that involve many discrete agents, where each agent behavioral rules are described in an algorithmic fashion rather than a purely mathematical way. This allows implementing complex internal properties of agents and their nontrivial behavioral rules (Sayama 2015).

Lou et al (2009) proposed a cooperative quality management architecture in supply chains, with production and outsourcing, based on a multi-agent system with information exchange. Romero et al (2012) developed a collaborative model for solving problem adapted to SCQM, considering strategic decisions in product development.

Sun and Li (2010) developed a SCQM model based on immune theory by analyzing the mechanisms and functions from four aspects: recognition, learning, memory, effects. Kuei et al (2008) developed a simulation study of SCQM considering the effectiveness of supply chain operations, demand uncertainty, supply chain speed, and quality and distribution issues.

3.3. Network Models

Network models are based on a set of network layers and how they interact. They are one of the most recent developments of complex systems science. Their historical roots are discrete mathematics and statistical physics (Sayama 2015).

In this area, Liu et al (2012) developed a Petri net model for SCQM conflict resolution of a complex product. It allows making decisions to select the appropriate activities when a quality conflict has happened. Yan et al (2010) proposed an ontology of collaborative supply chain. It provides a foundation for the coordination and integration of the business process to measure, analyze, and continually improve the quality of products, services, and process.

4. MODELING WITH FUZZY COGNITIVE MAPS

Fuzzy Cognitive Map (FCM) is a tool for modeling complex systems utilizing existence knowledge and human experience and/or available knowledge from existing databases in the form of rules. FCMs are a combination of fuzzy set theory with heuristic learning of neural networks (León et al. 2010).

Figure 3 illustrates a FCM representation and its elements. A FCM consists of concepts (C) which represent the studied system, and directed arcs which represent the casual relationships between the concepts (Kosko 1986). The concepts are indexed by subscripts i (cause node) and j (effect node). Each directed arc is labeled with a fuzzy value (e) in the interval [-1, +1] that represents the strength of impact between the concepts. The weights assigned (e_{ij}) to the pairs of concepts (C_{ij}) are stored in an adjacency matrix E (Papageorgiou 2014).

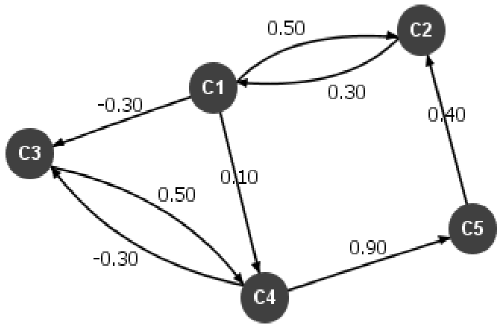


Figure 3: Example of a FCM representation with five concepts.

The mathematical modeling is carried out after the FCM is constructed and data from the input concepts are feed. Values of the concept C_i in time t are represented by the state vector $A_i(k)$. The state of the FCM is described by the state vector $A(k) = [A_i(k), \dots, A_n(k)]$. The value A_i in a moment k+1 is calculated by the sum of the previous value of A_i in a prior moment k with the product of the value A_j of the cause node C_j in precedent moment k and the value of the cause-effect link e_{ij}:

$$A_i(k+1) = f\left(A_i(k) + \sum_{j=1}^{N} A_j(k) \cdot e_{ji}\right) \quad (1)$$

where f (·) is an activation function which gives values of concepts in the range [0, 1] and is formulated as follows:

$$f(x) = \frac{1}{1+e^{-mx}} \quad (2)$$

where m is a real positive number and x is the value $A_i^{(k)}$ on the equilibrium point.

FCMs have been applied in recent years in the analysis of complex systems, through simulation, modeling and decision analysis (Rezaee et al. 2017), since they allow predicting and analyzing influences and performances of complex event using analysis of causal relationships that change over time (Dickerson and Kosko 1994; Jetter and Kok 2014).

FCMs have been used, for example, to model the causal relationships of a stakeholder relations management system (Susniene et al. 2014), to estimate crop yield (Mourhir et al. 2017), to model the total energy behavior of buildings (Mpelogianni et al. 2015), among other applications.

5. DEVELOPMENT OF A SCQM INTEGRATION MODEL USING FUZZY COGNITIVE MAP

The development of the SCQM model using FCM was carried out through three steps (Kandasamy and Smarandache 2003): Selection of concepts and relations (adjacency matrix), Performance of the inference process and Convergence analysis.

5.1. Selection of concepts and relations

The first stage is the development of the SCQM concepts and their relations. The relations values were established through analysis of historical data in the supply chain of a plastics sector company, whose name is kept in reserve due to commitments of confidentiality of the information supplied. Table 1 shows the adjacency matrix of SCQM concepts. Figure 4 shows the FCM corresponding to the relationships established in Table 1.

Table 1: Adjacency matrix for SCQM modeling using FCM

SCQM Concepts		C1	C2	C3	C4	C5	C6	C7	C8	C9
		Logistics costs (%)	Quality costs (%)	Return on Assets (ROA)	Defective product in production (%)	Emissions, effluent and waste (%)	Rejections and Returns (%)	Customer Satisfaction (OTIF deliveries)	Employee Satisfaction (%)	Suppliers Development (%)
C1	Logistics costs (%)	0	0.2	-0.7	0	0.3	0.2	-0.5	0	-0.4
C2	Quality costs (%)	0.5	0	-0.8	1	0.75	1	0.9	0.3	0
C3	Return on Assets (ROA)	0	0	0	0	0	0	0	0	0
C4	Defective product in production (%)	-1	-1	-1	0	-0.8	-1	-0.8	-0.25	0
C5	Emissions, effluent and waste	-1	-0.9	-0.8	-0.6	0	0	-0.6	-0.4	0
C6	Rejections and Returns (%)	-1	-1	-1	0	-1	0	-1	-0.25	0
C7	Customer Satisfaction (OTIF	-1	-0.5	0.7	0	0	1	0	0.8	0
C8	Employee Satisfaction (%)	1	1	1	1	0.6	1	1	0	0
C9	Suppliers Development (%)	1	1	1	1	1	1	1	0.6	0

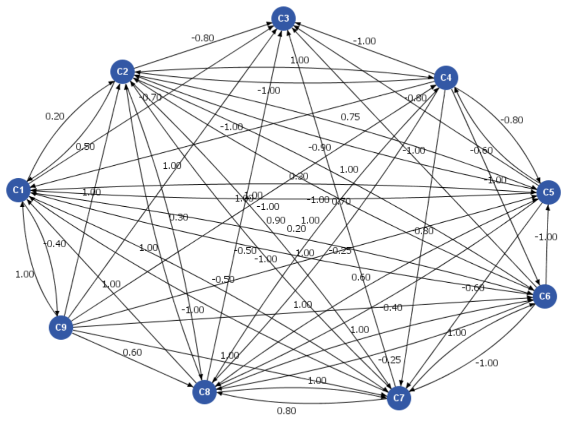


Figure 4: The FCM developed for SCQM modeling.

5.2. Performance of the inference process

The parameters related to the inference process are the inference rule, the transfer function and the stopping criterion. In this model, the inference process was carried out by performing WHAT-IF simulations and the Kosko´s activation rule with self-memory (equation 1) was used. The stopping criterion was established when a fixed-point attractor is reached.

The WHAT-IF simulation is performed by multiplying the input configuration vector representing the state of each node with the adjacency matrix. Thus, the input for iteration *k*+1 is the sum of output from iteration *k* and input of iteration *k*.

The Kosko´s activation rule with self-memory was used because, unlike the standard Kosko´s activation rule, it takes into consideration the previous state of the current node in addition to the other, linked nodes and the weight of those connections (Obiedat and Samarasinghe 2016). Figure 5 shows the results of the inference process, demonstrating the convergence of model concepts.

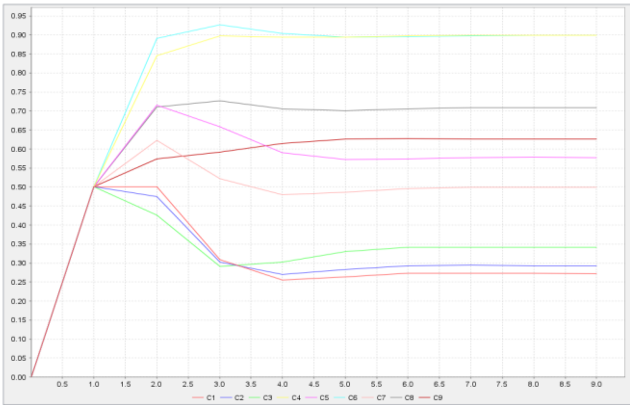


Figure 5: Results of the inference process.

5.3. Convergence analysis

After performing the inference process, a FCM has three behavior options: (i) The successive state vectors converge quickly to a stable state (a fixed-point attractor), (ii) A series of state vectors appear over and over periodically (a limit cycle), or (iii) The state vector changes "randomly" after every cycle (a chaotic behavior) (Tsadiras 2008).

Because the studied model characteristics, we established the stopping criterion when a fixed-point attractor is reached. In SCQM modeling is desired that outcomes converge to a fixed value because this way the future stable state of the system can be predicted and right decisions may be made.

Figure 5 shows that concept values converged fast to their final, stable values. Table 2 shows the values of concepts in the last simulation step and their order. These results allow us to identify three decision-making ranges and to prioritize the SCQM concepts most important.

The first one consists of the highest concept values: rejections, returns, and defective product in production. The second one consists of employee satisfaction, suppliers development, emissions, effluent and waste and customer Satisfaction. The third one consists of ROA, quality costs and logistics costs. This is because the latter are output vectors whose results depend on the decisions made in the concepts of the first two ranges.

Table 2: The order and final values of vectors for SCQM modeling

SCQM CONCEPT		ORDER	VALUE
C6	Rejections and Returns (%)	1	0.8987
C4	Defective product in production (%)	2	0.8984
C8	Employee Satisfaction (%)	3	0.7094
C9	Suppliers Development (%)	4	0.6265
C5	Emissions, effluent and waste (%)	5	0.5779
C7	Customer Satisfaction (OTIF deliveries)	6	0.4996
C3	Return on Assets (ROA)	7	0.3411
C2	Quality costs (%)	8	0.2923
C1	Logistics costs (%)	9	0.2721

6. CONCLUSIONS

In this paper, we analyzed rule-based modeling for SCQM coordination and integration by a comprehensive literature review on studies published from 2000 to 2018, in order to identify the main contributions, the analytical modelling techniques used and the main research avenues. Most of rule-based models for SCQM were for coordination and there is a lack of models for integration and joint coordination and integration.

We developed a SCQM model using FCM. The results obtained by simulation and the convergence analysis of the inference process validated the initial selection of the SCQM concepts and the values of their relations.

This paper is the result of the initial step of an ongoing research aimed to develop an analytical model for Supply Chain Quality Management coordination and integration using multi-stage modelling approach, considering multiple products and supply chain levels. The next steps will be focused on determination of a greater number of levels and agents, their coordination and integration strategies and refinement of inference process.

REFERENCES

Bautista-Santos H. et al., 2015. Integration model of collaborative supply chain. Dyna, 82(193), 145–154.

Cheung K.L. and Leung, K.F., 2000. Coordinating replenishments in a supply chain with quality control considerations. Production Planing and Control, 11(7), 697–705.

Das K. and Sengupta S., 2010. Modelling supply chain network: a quality-oriented approach. International Journal of Quality & Reliability Management, 27(5), 506–526.

Dickerson J.A. and Kosko B., 1994. Virtual Worlds as Fuzzy Cognitive Maps. Presence: Teleoperators and Virtual Environments, 3(2), 173–189.

Evans J.R., Foster S.T. and Linderman K., 2014. A Content Analysis of Research in Quality Management and a Proposed Agenda for Future Research. Quality Management Journal, 21(2), 17–44.

Flynn B. and Zhao X., 2015. Global Supply Chain Quality Management: Product Recalls and Their Impact. Boca Raton: CRC Press.

Foster S.T., 2008. Towards an understanding of supply chain quality management. Journal of Operations Management, 26(4), 461–467.

Gao C. et al., 2016. Incentives for quality improvement efforts coordination in supply chains with partial cost allocation contract. International Journal of Production Research, 54(20), 6213–6231.

Jetter A.J. and Kok K., 2014. Fuzzy Cognitive Maps for futures studies-A methodological assessment of concepts and methods. Futures, 61, 45–57.

Kandasamy V. and Smarandache F., 2003. Fuzzy Cognitive Maps and Neutrosophic Cognitive Maps. Phoenix: Xiquan.

Kosko B., 1986. Fuzzy cognitive maps. Int. J. Man-Machine Studies, 24, 65–75.

Kuei C.-H. and Madu C.N., 2001. Identifying critical success factors for supply chain quality management (SCQM). Asia Pacific Management Review, 6(4), 409–423.

Kuei C.-H., Madu C.N. and Winch J.K., 2008. Supply chain quality management: a simulation study. Information and Management Sciences, 19(1), 131–151.

León M. et al., 2010. Fuzzy Cognitive Maps for Modeling Complex Systems. Proceedings of 9th Mexican International Conference on Artificial Intelligence, MICAI 2010, pp. 166–174.

Liu Y. et al., 2012. Petri net model for supply-chain quality conflict resolution of a complex product. Kybernetes, 41(7/8), 920–928.

Lou P. et al., 2009. Production-Outsourcing Supply Chain Quality Management Based on Multi-Agent System. Proceedings of The 16th International Conference on Industrial Engineering and Engineering Management, 2009. IE&EM '09, pp. 1555–1559.

Mellat-Parast M., 2013. Supply chain quality management: An inter-organizational learning perspective. International Journal of Quality & Reliability Management, 30(5), 511–529.

Modak N.M., Panda S. and Sana S.S., 2016. Three-echelon supply chain coordination considering duopolistic retailers with perfect quality products. International Journal of Production Economics, 182, 564–578.

Montoya-Torres J.R. and Ortiz-Vargas D.A., 2014. Collaboration and information sharing in dyadic supply chains: A literature review over the period 2000–2012. Estudios Gerenciales, 30, 343–354.

Mourhir A. et al., 2017. Exploring Precision Farming

Scenarios Using Fuzzy Cognitive Maps. Sustainability, 9(7), 1241-1263.
Mpelogianni V., Marnetta P. and Groumpos P.P., 2015. Fuzzy Cognitive Maps in the Service of Energy Efficiency. IFAC-PapersOnLine, 48(24), 1–6.
Obiedat M. and Samarasinghe S., 2016. A novel semi-quantitative Fuzzy Cognitive Map model for complex systems for addressing challenging participatory real life problems. Applied Soft Computing Journal, 48, 91–110.
Pang J. and Tan K.H., 2018. Supply chain quality and pricing decisions under multi-manufacturer competition. Industrial Management & Data Systems, 118(1), 164–187.
Papageorgiou E., 2014. Fuzzy Cognitive Maps for Applied Sciences and Engineering: From Fundamentals to Extensions and Learning Algorithms. Berlin: Springer.
Rezaee M.J., Yousefi S. and Babaei M., 2017. Multi-stage cognitive map for failures assessment of production processes: An extension in structure and algorithm. Neurocomputing, 232, 69–82.
Robinson C.J. and Malhotra M.K., 2005. Defining the concept of supply chain quality management and its relevance to academic and industrial practice. International Journal of Production Economics, 96(3), 315–337.
Romero J.C. et al., 2012. Collaborative methodology for supply chain quality management: Framework and integration with strategic decision processes in product development. Proceedings of the 6th European Conference on Information Management and Evaluation, ECIME 2012, pp. 418–427.
Sayama H., 2015. Introduction to the Modeling and Analysis of Complex Systems. New York: Open SUNY Textbooks.
Su Q. and Liu Q., 2011. Supply Chain Quality Management by Contract Design. In D. Önkal & E. Aktas, eds. Supply Chain Management - Pathways for Research and Practice. Rijeka: InTech, 57–74.
Sun P. and Li Q., 2010. Study on Supply Chain Quality Management Model Based on Immune Theory. Proceedings of the 2010 International Conference on Management and Service Science, pp.1–4.
Susniene D. et al., 2014. Using Fuzzy Cognitive Map Approach to model the casual relationships in stakeholder management at companies. Proceedings of the 5th IEEE International Conference on Cognitive Infocommunications, pp. 121–124. Vietri sul Mare, Italy.
Tsadiras A.K., 2008. Comparing the inference capabilities of binary, trivalent and sigmoid fuzzy cognitive maps. Information Sciences, 178(20), 3880–3894.
Tsai T.P. and Wang F.-C., 2004. Improving Supply Chain Management: A Model for Collaborative Quality Control. In Advanced Semiconductor Manufacturing, 2004, ASMC'04. IEEE Conference and Workshop, pp. 36–42.
Xiao T., Yang D. and Shen H., 2011. Coordinating a supply chain with a quality assurance policy via a revenue-sharing contract. International Journal of Production Research, 49(1), 99–120.
Yan J. et al., 2010. Ontology of Collaborative Supply Chain for Quality Management. World Academy of Science, Engineering and Technology, 4(4), 365–370.
Yoo S.H., 2014. Product quality and return policy in a supply chain under risk aversion of a supplier. International Journal of Production Economics, 154, 146–155.

AUTHORS BIOGRAPHY

Juan M. Cogollo Flórez is a Professor at the Department of Quality and Production, Instituto Tecnológico Metropolitano-ITM, Medellín, Colombia. He received a MSc in Management Engineering in 2011, from the Universidad Nacional de Colombia. He is currently a doctoral student in Engineering - Industry and Organizations at Universidad Nacional de Colombia. His current research interests are Performance Measurement, Supply Chain Quality Management and Advanced Statistical Quality Control.

Alexander A. Correa Espinal is a Full Professor at the Department of Organizational Engineering, Universidad Nacional de Colombia, Medellín, Colombia. He received the PhD in Statistics and Operational Research in 2007 from the Universitat Politècnica de Catalunya (Spain) and the MSc in Industrial Engineering in 1999 from the Universidad de los Andes (Colombia). His current research interests are Advanced Design of Experiments, Advanced Statistical Quality Control and Total Quality Management.

TOWARDS RETOOLING THE MICROSOFT HOLOLENS AS OUTDOOR AR AND MR DEVICE

Christoph Praschl[(a)], Oliver Krauss[(b)], Gerald Zwettler[(c)]

[(a),(b),(c)]Research Group for Advanced Information Systems and Technology (AIST), Research and Development Department, University of Applied Sciences Upper Austria, Softwarepark 11, 4232 Hagenberg, AUSTRIA
[(c)]School of Informatics, Communications and Media, University of Applied Sciences Upper Austria, Softwarepark 11, 4232 Hagenberg, AUSTRIA

[(a)]christoph.praschl@fh-hagenberg.at, [(b)]oliver.krauss@fh-hagenberg.at, [(c)]gerald.zwettler@fh-hagenberg.at

ABSTRACT
This research work covers generic approaches for the determination of the outdoor position and orientation of an augmented reality device due to the lack of outdoor capability of depth-sensor based devices currently available on the market.
The determination of the orientation is primarily achieved with an attitude heading reference system (AHRS) for a rough estimation. Based on a connected/inbuild video camera the accuracy at minor changes of the orientation is enhanced applying registration to assess the differences in orientation between two video frames, compensating gyroscope drift errors. The position determination is achieved using GPS with a rover- and base station real time kinematic beacon system to achieve enhanced precision. Results show that due to sensor application AR hardware considered for indoor use can be retooled to properly work outdoors, at large distances and even inside running vehicles. Thus, future implementation of applications in various domains is facilitated.

Keywords: Augmented Reality, Orientation, Positioning, Image registration, modelling AR training scenarios

1. INTRODUCTION

Electronic devices are getting more and more ubiquitous and developments over the last years in the field of mobile devices led to more flexible applications which are independent from a static working place and enable the virtual and the real world to merge. This flexibility and mixture results in the prominence of the field of Augmented Reality (AR) which originally came up in 1968 (Sutherland, 1968). In particular, developments in terms of specialized devices for this field of application, as the Microsoft HoloLens (Microsoft, 2018a) or the Magic Leap One (Magic Leap, 2018), combined with software engineering advancements, as Apple´s ARKit (Apple, 2018), Google´s ARCore (Google, 2018) or the Vuforia framework (PTC Inc., 2018a), have contributed to the increase of interest in AR. Especially the software aspect allows AR to become more ubiquitously because first steps in this field can already be done using some smartphone instead of expensive specialized data glasses (Papagiannakis, Singh and Magnenat-Thalmann, 2008).
In case of data glasses, the position of the AR-device can be equated to the user´s position and the orientation relative to the viewing direction because the association is directly created through the manner the device is used. For smartphones the position can be roughly approximated, because a user holds the device in well-known poses, close to their person as to see the display. This leads to a different association of the real world to the virtual one, because it is not directly depending on the user but on the alignment of the device.

1.1. State of the art

From a state-of-the-art point of view, several generic and AR independent but suitable approaches for position and orientation determination exist. These techniques can be distinguished in the area of application between indoor and outdoor. The following explanations refer only to outdoor approaches but can also be partly used for indoor usage under certain circumstances.

Generic orientation approaches

Image registration as well as similarity-metrics based pattern matching approaches can be utilized for image based orientation determination. The former of those is an image processing technique in which two images are compared according to quantitative metrics and transformed in an affine way utilizing optimization approaches to maximize the level of congruence. Therefore, one reference image is translated, scaled and/or rotated and is compared to the original image in iterative steps (Goshtasby, 2012). The comparison is based on affinities represented by a distance metric which can be calculated for example with the sum of square errors, cross correlations, Euclidean distance maps on extracted borders or mutual information (Melbourne, Ridgway, and Hawkes, 2010).

Computer vision frameworks like Open Source Computer Vision Library (OpenCV) (OpenCV, 2018a), Insight Segmentation and Registration Toolkit (ITK) (ITK, 2018), or ImageJ (ImageJ, 2018) offer robust functionality for calculating the shift between two images, which can be used to determine changes in orientation. Therefore, different algorithms are implemented in those frameworks which use unique features to calculate the translation changes of two images. In OpenCV a phase correlation function (OpenCV, 2018b) is available which is based on the Fourier shift theorem to detect frequency shifts and so to quickly conclude the image transformation.

Generic positioning approaches

In the field of position determination, different sensor-based approaches can be used. Inside-out tracking methods are distinguished from outside-in ones (Boger, 2014). While the first ones are techniques which are able to determine the device`s position in a passive way by analyzing the collected data of internal measurement units requiring no external devices or targets, the second variant uses external systems which provide information about the own localization from an outer point of view.

Internal algorithms like the Simultaneous Localization and Mapping (SLAM) algorithm (Aulinas, Petillot, Salvi and Lladó, 2008) or sensor fusion techniques (Elmenreich, 2002) can be used to draw conclusions about the device´s position in the world using sensors like accelerometers, gyroscopes and magnetometers – for example utilizing an inertial measurement unit (IMU) or an AHRS sensor platform – as well as positioning systems as the Global Positioning System (GPS) (U.S. Air Force, 2008), Galileo (European Communities, 2007) or GLONASS (Polischuk, et al., 2002). The position determination can also be achieved base on acceleration. The data of an accelerometer can be used for calculating the velocity and in a further step approximate the position (Qiu, Wang, Zhao, Qin, Li, Hu 2018). Other techniques can be used for an internal identification of the position as well, for example using detected Bluetooth beacons (Ehrenborg, 2015) or wireless local area networks (WLAN) (Zaidi, Tourki, Ouni, 2010). By recognizing such wireless emitters and based on the connection strength, geometric attempts can be used to draw conclusions about the position when the absolute position of those devices is known. Alternatively, external approaches can be based on monitoring techniques in the form of calibrated video cameras using a grid to identify the real word position (Buschmann, Müller, Fischer, 2004).

Approaches for Augmented Reality

In terms of AR, different user tracking methods (Bostanci, Kanwal, Ehsan, Clark, 2013) for indoor as well as outdoor applications have been researched in the last years. For example, studies were done using infrared markers and a head-mounted camera to determine the user´s position and orientation (Maeda, Ogawa, Kiyokawa, Takemura, 2004). A similar technique is available through AR frameworks like Vuforia or Wikitude (Wikitude, 2018), which allows to associate the real world´s position and orientation to the virtual ones using pattern matching algorithms for marker-based approaches (PTC Inc., 2018b).

1.2. Status Quo of Position and Orientation Determination for Augmented Reality devices

Despite the technical evolution, current hardware solutions in the field of AR are not mature enough to enable outdoor usage, because orientation and positioning doesn't work well enough to create interactive user experiences that require exact placement in the real world and specialized data glasses as the Microsoft HoloLens are not recommended to be used in outdoor scenarios at all. The reason for this is that in case of data glasses, the AR-experience is primarily based on the device´s feature detection, respectively it´s room recognition which is achieved by the combination of several sensors and is limited due to those building blocks. In case of Microsoft´s HoloLens the restriction are the depth sensors, which are basically infrared sensors, with a range up to almost three meters (Microsoft, 2018b). These sensors allow the AR-device to place virtual objects into the real world and to fix them stably in place using unique features from the environmental information. Because of the limitation of three meters the Microsoft HoloLens is not able to find features at greater distances to real world objects and so is not able to position and associate the virtual ones with the real environment in outdoor scenarios. Additionally, such infrared depth perception cameras are incapable of outdoor positioning due to ambient infrared during the day.

Another example is Metavision´s Meta2 (Metavision, 2018) which is not a standalone device as the Microsoft HoloLens and so can only be used in combination with a powerful graphics-card / computer, which is a problem in terms of outdoor usage.

This lack in possibilities to determine the position in real world outdoor scenarios and the orientation of current AR-devices results in the need for some supplementary approaches.

2. METHODOLOGY

The approaches for sensor- and image-based orientation and for GPS-RTK positioning were implemented for the usage on a Microsoft HoloLens to evaluate the outdoor ability. Therefore, the existing system of the AR-device including the system software with its indoor orientation and positioning components as well as the built in sensors, were extended by additional sensors and software approaches, see Fig. 1. To achieve this, sensors were evaluated and selected based on quality attributes like accuracy, size, weight, communication protocol and energy consumption, because the extensions should not impair the functionality and usability of the AR-device.

To facilitate orientation determination an image based (see section 2.1) and a sensor based (see section 2.2) approach were combined to a hybrid outdoor-orientation component (see section 2.3). For the localization part two approaches were evaluated. On the one hand positional changes are calculated using information of an accelerometer (see section 2.4) and on the other hand a GPS-RTK base- and master station system (see section 2.5) was included to determine the position in outdoor scenarios. Those approaches are not combined because of the restriction to an axis-parallel movement for the accelerometer-based approach and the already sufficient accuracy of the GPS-RTK position determination.

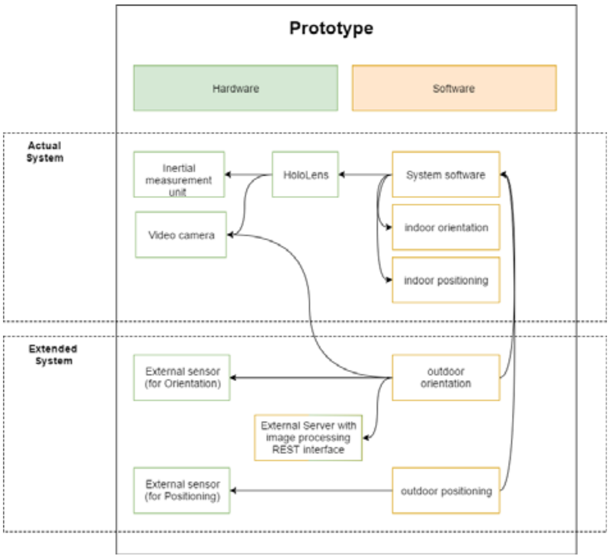


Figure 1: Prototype architecture showing the Microsoft HoloLens architecture (actual system) and the necessary hardware and software extensions (extended system).

2.1. IMAGE BASED ORIENTATION

Changes in orientation can be calculated based on the presumption that the orientation around an axis can be simplified to a translation along an axis. Therefore, the prerequisite that homogenous object distances in the images need to exist, must be fulfilled (Praschl, 2017). Thus, real objects and especially their visual features are present at homogenous distances, see Fig. 2.

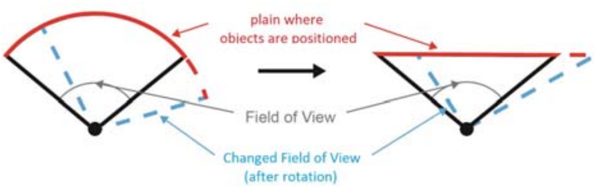
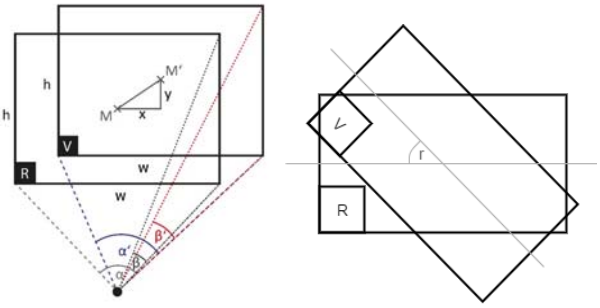


Figure 2 Modified from (Praschl, 2017): Orientation simplification.

Based on this simplification, changes in orientation can be calculated from two images R and V created in a staggered sequence from the ratio between translation and the associated image-dimension to rotation and the associated camera´s field of view. This also requires an image registration technique to calculate the shift and rotation differences of the compared images. Therefore, the translation on x-axis is defined as x, the translation on y-axis as y and the rotational offset as r, as can be seen in Fig. 3 (Praschl, 2017).



(a) (b)

Figure 3 Modified from (Praschl, 2017): (a) Translation of image R to image V and (b) Rotation of image R to image V.

The rotational changes (r) can be transferred directly to the change in orientation around the z-axis (yaw). In turn the difference in orientation around the x-axis (roll) can be calculated with the translation on the x-axis, the image´s width (x) and the horizontal camera´s field of view (α), see Equ. 1. Similarly, to the roll calculation, the difference in orientation on the y-axis (pitch) can be calculated with the translation on the y-axis, the image´s height (y) and the vertical camera´s field of view (β), see Equ. 2 (Praschl, 2017).

$$roll = \frac{x}{w} * \alpha \qquad (1)$$

$$pitch = \frac{y}{h} * \beta \qquad (2)$$

2.2. SENSOR BASED ORIENTATION

One possibility for the orientation determination can be a combined hardware and software solution in form of an AHRS sensor platform which is based on building blocks for measuring magnetic flux density as well as angular and linear acceleration on three axes. What distinguishes this to an IMU is that AHRS systems do not only provide raw data but also contain on-board processing units for combining the data of the single sensors and for calculating for example the orientation angles (Tomaszewski, Rapinski, Pelc-Mieczkowska, 2017). There are different approaches which can be used for this kind of filter. The combination is usually done with non-linear estimations using filters like an extended Kalman filter (Kalman, 1960) or one of the below described ones.
(Pedley, 2013) describes a solution which uses the axis data of an accelerometer (acc_x, acc_y and acc_z) to calculate the orientation angles pitch and roll, see Equ. 3, respectively Equ. 4. Based on those calculations

in combination with a magnetometer (mag_x, mag_y and mag_z) the third orientation angle yaw y can be calculated, see Equ. 5 (Caruso, 2000).

$$p = \tan^{-1}\left(\frac{acc_x,}{\sqrt{acc_y^{\,2} + acc_z^{\,2}}}\right) \tag{3}$$

$$r = \tan^{-1}\left(\frac{acc_y,}{\sqrt{acc_x^{\,2} + acc_z^{\,2}}}\right) \tag{4}$$

$$\begin{aligned} &\quad mag_y * \cos(r) + mag_z * \sin(r), \\ y = \tan^{-1}&(mag_x * \cos(p) + mag_y * \sin(r) *) \\ &\quad \sin(p) - mag_z * \cos(r) * \sin(p) \end{aligned} \tag{5}$$

Another orientation filter is described in (Madgwick, 2010), using the data of a three-axis accelerometer and a three-axis gyroscope with the possibility to improve the results by adding the information of a three-axis magnetometer. This filter is based on a gradient-descent algorithm and achieves an accuracy of < *0.6°* in static and respectively < *0.8°* in dynamic usages.

2.3. HYBRID ORIENTATION

To achieve better results, different approaches can be combined for the orientation determination. Therefore, the single calculations can be merged using a weighted averaging method, see Equ. 6, but this could lead to inaccuracies in certain circumstances where one single solution results in a completely incorrect outcome.

$$\overline{x_w} = \frac{\sum_{i=1}^{n} x_i * w_i}{\sum_{i=1}^{n} w_i} \tag{6}$$

To avoid such a behavior, the single approaches are differentiated at several ranges of orientation changes. On the one hand, sensor-based approaches can be used for a fast and initial estimation of the orientation but because of errors in the different building blocks, this method is too inaccurate in scenarios showing no or only minor changes, due to limitations in the sensors resolution as well as jitter (Hesse, Schnell, 2014). On the other hand, image-based approaches can only be used in scenarios showing small changes because robust level of matching features in both images needs to be reached to allow robust results, e.g. incorporating a lot of reference frames at cost of performance with respect to both, computational complexity and required frame rate (Nestares, Gat, Haussecker, Kozintsev, 2010). One approach to handle those problems is to weight the image-based approach stronger up to a certain threshold and to rely more to a sensor-based alternative above this.

2.4. ACCELERATION BASED POSITIONING

One possible process which can be used for calculating the position of a device is to use the so called dead reckoning. This method uses a previously determined position in combination with the current velocity to estimate the current position. This navigation technique can be done using a single accelerometer and based on the acceleration data, velocity can be calculated with an approximation of the integral, see Equ. 7. The current velocity v_n is calculated as the sum of accelerations using the average of the current acceleration a_n value and the previous one a_{n-1} multiplied with the time delta of the two measurements. Analogous, the position p_n is calculated using the current velocity and the previous velocity, see Equ 8. This calculation can be performed for each axis separately and works in that form for linear and axis-parallel movements but needs adaptations in cases where next to the translation also rotations are present (Keyglove, 2011).

$$v_n = \sum_{i=1}^{n} \frac{a_i + a_{i-1}}{2} * (t_i - t_{i-1}) \tag{7}$$

$$p_n = \sum_{i=1}^{n} \frac{v_i + v_{i-1}}{2} * (t_i - t_{i-1}) \tag{8}$$

2.5. GPS BASED POSITIONING

GPS is a global satellite system for position determination and was developed in the 1970s by the US Department of Defense. This system provides unidirectional localization data up to an accuracy of about 5 meters which can be received by specialized GPS receiver nearly anywhere on earth. Due to those characteristics this system is suitable for outdoor applications in nearly any form (Tsui, 2005).

ROVER AND BASE STATION APPROACH

The accuracy of GPS systems can be improved with methods like real time kinematic (RTK). RTK enhances the precision of the position data with a non-moving reference station, the so called base station, which allows to correct the data of the moving counterpart, called the rover. For this adjustment the base station needs the undisturbed signal carrier wave and the data of five or more satellites to calculate its own average position and increase the precision of the correction factor to reduce the rover´s error in terms of its position. The usage of such a rover and base station approach allows to improve the accuracy of GPS up to centimeter-level (Xu, 2012). Different hardware solutions in the field of GPS-RTK are available for example in form of specialized beacons which are suitable for usage in combination with an AR-device because of its characteristics as size and weight (Emlid, 2018).

3. IMPLEMENTATION

The presented techniques for outdoor orientation and positioning were implemented in C# for the Microsoft HoloLens as Universal Windows Platform (UWP) (Microsoft, 2018c) application using the UNITY (Unity Technologies, 2018a) game development platform. To avoid high computational loadings on the AR-device the used computer vision methods were outsourced to an external server with a representational state transfer (REST) interface but also with the possibility of communicating via web sockets using binary messages, see Fig. 5. This external part for the orientation calculation was implemented in a separate computer vision framework based on the OpenCV Java wrapper by OpenPnP (OpenPnP, 2017) in combination with the Spring Framework (Pivotal Software, 2018) and allows amongst other things to send images from any device as the Microsoft HoloLens and to response with the image-based orientation results.

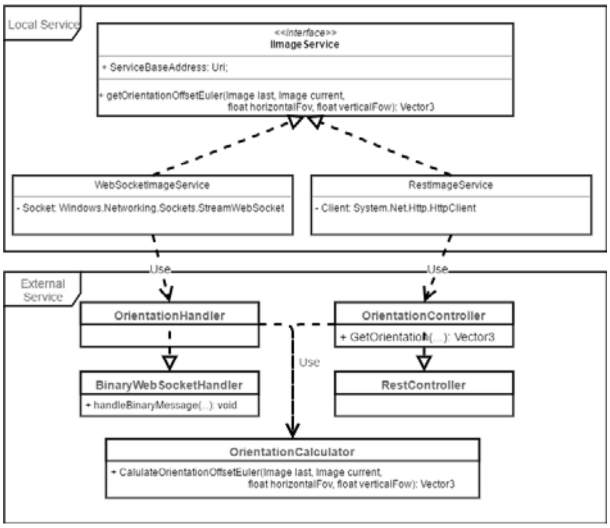


Figure 4: External image based orientation service.

In terms of software architecture, the orientation as well as the position determination are defined as single components which are combined in a summarizing position and orientation module for the separation of concerns. This combined module serves the purposes to adapt the internal position and orientation determinations of the Microsoft HoloLens by additional calculated ones, more on that below. Both components are defined as interfaces for which different implementations are provided and can be exchanged, see Fig. 6.

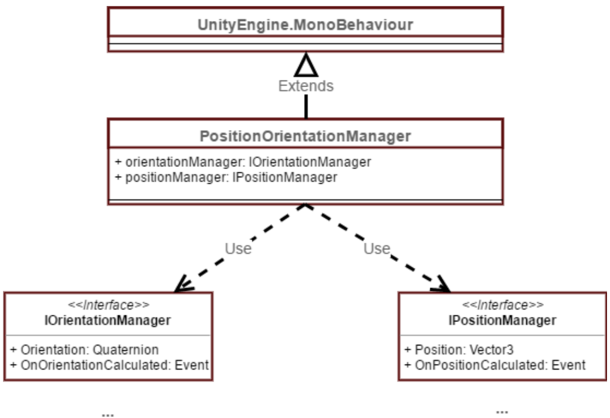


Figure 5: Position and Orientation rough Architecture.

This allows to swap for example the acceleration-based positioning approach with the GPS counterpart, see Fig. 7, but also to exclusively use either the image-based orientation module, the sensor-based one or a hybrid implementation, see Fig. 8. Furthermore, the image-based orientation implementation uses any wrapper of a camera access service. This makes it possible to dynamically decide based on the application which approach suits better to the power requirements, by neither taking a single photo in a defined time interval, which results at that certain moment in a higher computation workload, or to use each n^{th} frame from a continuous video recording where the average workload is a bit higher, but the access of the single frame does not affect the system a lot. Furthermore, the utilized communication approach is abstracted to allow the external access with either REST calls or a defined web socket interface, see Fig. 5.

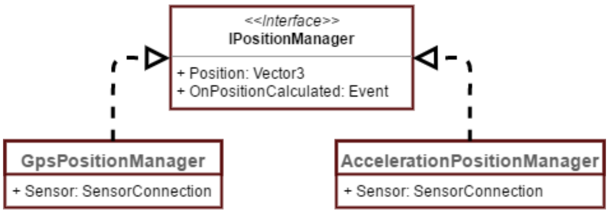


Figure 6: PositionManager.

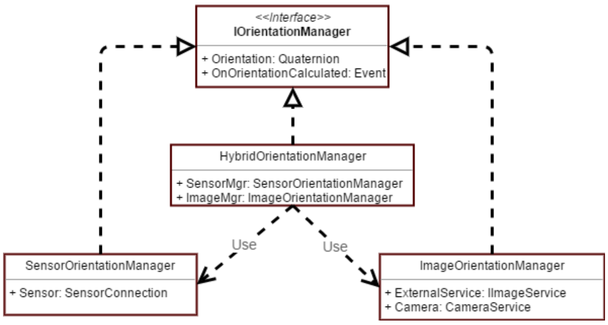


Figure 7: OrientationManager

The approach described in section 2.1 is implemented using the `PhaseCorrelate` method with a focus on quick search for the translation between the images. For this the orientation determination is limited to changes on the x- and the y-axis because of the algorithm´s lack on detecting rotational changes. In addition, a sensor-based approach is implemented too using an AHRS platform in form of a x-io Technologies “Next Generation IMU” (NGIMU) which provides pre-processed orientation data as quaternions or Euler angles based on the different building blocks (x-io Technologies, 2016).
The acceleration-based positioning approach utilizes the NGIMU´s accelerometer. Therefore, the sensor system is set to an accelerometer data send rate of *60Hz* and the determination is done using the dead reckoning process. To improve the calculation an initial measurement of the noise is done to remove it from the further data readings and a threshold is defined as minimal

acceleration to initialize the continuous velocity and position calculation. In the case of negative acceleration measurements over several frames the stop of the calculation is initiated and is canceled at a velocity of about zero. This approach allows the position determination on a straight path with a single acceleration followed by continuing movement and a deceleration. A possible field of application of this implementation in terms of AR is the usage in a car on a straight road for example for risk simulations.

Additional to the acceleration-based positioning approach, also a GPS-RTK based one is implemented using the Emlid Reach RTK kit (Emlid, 2018). Those beacons are able to be set up as a rover and base station system and are able to communicate via a wireless network using the Transmission Control Protocol/Internet Protocol (TCP). This way it is possible to receive the position messages on the Microsoft HoloLens and to use this information for the outdoor positioning.

The development of AR-devices in UNITY follows the approach that the device determines the orientation and position by itself. For example, the Microsoft HoloLens uses its sensors to draw conclusions about its localization and adapts those data in every frame onto the UNITY main camera (Unity Technologies, 2018b) object which is equal to the user´s view of the virtual world. This works fine as long as it is for indoor applications but in terms of outdoor applications this behaviour results in the problem that there is no possibility to directly manipulate the camera´s position and orientation from outside. A solution for this is to wrap the camera object with another UNITY GameObject (Unity Technologies, 2018c) and modify this one, what happens as mentioned before in the combined position and orientation module. This results in the position and orientation modification of all child objects of the wrapper because they are associated relatively to the wrapper´s local coordinates system. Utilizing this approach, the wrapper´s position (p_w) has to be adapted by the additional calculated localization (p_a) minus the internal one calculated by the Microsoft HoloLens (p_h), see Equ. 9. Similarly, the orientation of the wrapper (o_w) has to be adapted by multiplying the quaternion which represents the additional calculated orientation (o_a) with the inversed absolute orientation quaternion received from the Microsoft HoloLens (o_h), see Equ. 10

$$pos_w = pos_a - pos_h \qquad (9)$$

$$o_w = o_a * o_h^{-1} \qquad (10)$$

4. RESULTS

The following results were measured from tests utilizing different hardware components. The main component thereby is the Microsoft HoloLens, which is assembled into a helmet and adapted by a sun protection film to improve the visual outdoor experience to allow a significantly improved outdoor visibility of the virtual objects, as seen in Fig. 8. For the wireless communication between the Microsoft HoloLens and the different sensors two Huawei E5573Cs-322 4G mobile WIFI (Huawei, 2016) routers are used. The orientation approach was tested using two different sensors. On the one hand a WIT JY901 (WIT Motion, 2018) was utilized for initial tests, but was exchanged by a x-io Technologies NGIMU (x-io Technologies, 2016) AHRS sensor platform, which was also used for the acceleration-based positioning approach. The GPS-RTK positioning was done using an EMLID Reach RTK kit (Emlid, 2018) consisting of two RTK modules.

Figure 8: Microsoft HoloLens in Helmet.

Orientation determination results

The sensor-based orientation approach is tested using a WIT JY901 (WIT Motion, 2018) sensor in combination with the algorithmic calculation described by (Pedley, 2013) but does not result in usable outcomes because of the sensor´s axis limitation and the not sufficient accuracy. Due to that, orientation filter by (Madgwick, 2010) is used as on-board algorithm of the NGIMU AHRS sensor platform. This method resulted in more accurate and usable orientation data but because of the different axis definitions of the sensor to those used in UNITY a transformation of the orientation data was necessary, which is described in (x-io Technologies, 2017).

The sensor-based orientation approach was tested multiple times in a static scenario set-up to evaluate the stability using a degree circle printed underlay, as it can be seen in Fig. 9. Over a period of 5 minutes this approach achieved in average stable results with a small inaccuracy due to the sensor´s observational errors and was able to hold a virtual object on the same position northbound, which can be seen exemplary in Fig. 10 with the single images and in Fig. 11 with those images overlaid.

It was also evaluated in a dynamic test scenario with two test persons in a fixed position, wearing the helmet with the Microsoft HoloLens. For this test the test person looked around and also turned around multiple times and moved their head back to the starting position afterwards, see Fig. 12 showing exemplary such a movement. Despite of some delay in which the sensor had to settle again also in this scenario the sensor-based approach resulted in valid outcomes in average. Fig. 13

shows that the virtual object is placed stable to the north over the period of the test, but it also indicates some discrepancy which is due to the lack of opportunity to stay exactly on the same position after multiple times turning around.

Figure 9: degree circle underlay.

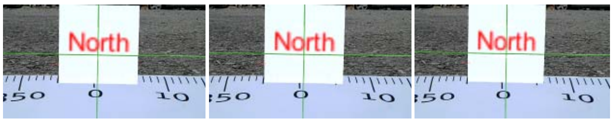


Figure 10: Test images over a time period of 5 minutes; left: at the beginning; middle: after 2,5min; right after 5min.

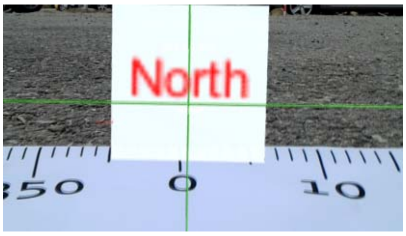


Figure 11: The three single images overlaid.

Figure 12: image sequence with head movement and rotation around own axis between the single images.

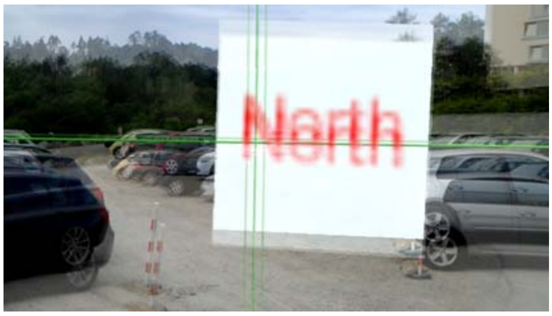


Figure 13: Three single northbound images overlaid.

The image-based orientation approach achieves in average an accuracy of about *0.05°* but is strongly recommended for only small changes from around *5°* between the compared images and depends heavily on the accuracy of the used registration method (Praschl, 2017).

Position determination results

The position determination approach based on the dead reckoning process using the accelerometer data of a NGIMU (x-io Technologies, 2016) AHRS sensor platform was successfully tested in cooperation with the Austrian Automobile Club (ÖAMTC) with about 100 test persons in a driver safety training simulation. This simulation was tested with the test persons as car passenger, see Fig. 14. The test was designed so that a driver accelerated the car up to about 40km/h and after the defined distance of 50 meters the virtual representation of a child, which was hidden before by a virtual truck, runs over a virtual pedestrian crosswalk and the passengers had to explain what is displayed during the drive, see Fig. 15 and Fig. 16.

Figure 14: Dead reckoning test setup.

Figure 15: Dead reckoning test live view.

Figure 16: Dead reckoning setup of virtual scene.

The GPS-RTK based positioning approach was tested in several passes in a scenario with the initial position as zero position. During the test a test person with the Microsoft HoloLens and the rover beacon moved around to several defined places and went back to the initial position. The distance between the stay places was measured roughly, due to a three dimensional movement, using a tape measure, see exemplary Table 1 and Fig. 17. In average the approach achieved an accuracy of about *30 cm*.

Table 1: GPS-RTK position results in meters as x (forward), y (upward) and z (sideways)

GPS-RTK result	Measured result
(-0.1, 0.1, 0.0)	(0.0, ?, 0.0)
(-0.5, 0.1, 0.1)	(-0.4, ? , 0.0)
(9.1, 0.1, 0.3)	(8.8, ?, 0.45)
(14.5, 0.1, 2.3)	(14.3, ?, 2.5)
(20.1, 0.1, -7.5)	(20, ?, -7)
(0.4, 0.1, 0.0)	(0.0, ?, 0.0)

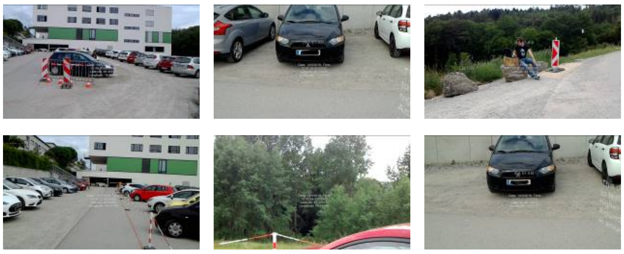


Figure 17: Snapshots of positioning stay places.

Figure 18: Snapshots of combined test stay places.

Position and Orientation determination results

The interaction of the orientation and the positioning module was tested based on the GPS-RTK positioning test scenario in combination with the sensor-based orientation approach. In this test case a virtual cross was positioned northern to the initial position and a test person again moved between several positions. Arrived at the respective positions the test persons turned in the direction of the virtual object to evaluate if it is still in place, see Fig. 18.

5. DISCUSSION AND CONCLUSIONS

In this paper strategies for retooling the Microsoft HoloLens as outdoor AR and MR device have been presented. For all relevant issues, like inaccuracies of positioning and orientation, appropriate concepts have been developed and validated via prototypes. Thus, infrared surrounding tracking, GPS inaccuracies and direct solar radiation are no longer limitation to development of outdoor AR applications.

Future development steps should also deal with the stabilization of the positioning as well as the orientation to get to even more accurate results. For example, next to the combined orientation approach also a hybrid implementation for positioning should be discussed to create a solution which is not depending on only one single approach.

6. OUTLOOK

The described approach has been tested in general with no explicit use case, so future practical use could take place in various predefined scenarios. Such application scenarios can be set in the area of outdoor training simulations for example for civil driver safety training but also for the education of drivers of emergency services as police, rescue or fire department in the situation of emergencies as well as for military vehicle driving training. Furthermore, also driving-independent outdoor training simulations can be implemented based on the tested orientation and positioning approaches, so for example military or police Special Forces are able to virtually train hostage taking situations without to endanger anyone. Additionally, such an AR simulation approach also allows to educate firefighters to handle house fire without setting the house on fire. Finally, a highly relevant field of application is the military sector, allowing training of officers in complex troop maneuvers and tactics at minimal costs and high flexibility.

ACKNOWLEDGMENTS

This research was funded by the Austrian Research Promotion Agency (*FFG*), program track *General Program* and was done in cooperation with the company REALSIM, especially Thomas Peterseil.

REFERENCES

Apple, 2018. *ARKit*. Available from https://developer.apple.com/arkit/ [accessed 9th May 2018].

Aulinas J., Petillot Y., Savli J., Lladó X., 2008. *The SLAM problem: a survey*. In Proc. of the 2008 Conference on Artificial Intelligence Research and Development, 363-371.

Boger Y., 2014. *Positional tracking: "Outside-in" vs. "Inside-out"*. Available from https://vrguy.blogspot.co.at/2014/08/positional-tracking-outside-in-vs.html [accessed 9th May 2018].
Bostanci, E., Kanwal, N., Ehsan, S., and Clark, A., 2013. User Tracking Methods for Augmented Reality. *International Journal of Computer Theory and Engineering* Vol. 5: 93-98.
Buschmann, C., Müller, F., and Fischer, S., 2004. *Grid-Based Navigation for Autonomous, Mobile Robots*. Technical University of Braunschweig, Braunschweig, Germany.
Caruos, M., 2000. Applications of magnetic sensors for low cost compass systems. *IEEE 2000 Position Location and Navigation Symposium*, 177-184. 13. – 16. March 2000, San Diego, USA.
Ehrenborg, G., and Fransson, H., 2015. *Sensor Fusion based Indoor Positioning with iBeacons*. Master´s Thesis. Chalmers University of Technology, Gothenburg.
Elmenreich, W., 2002. *Sensor Fusion in Time-Triggered Systems*. Doctor´s Thesis. Vienna University of Technology.
Emlid, 2018. Reach RTK kit. Available from https://docs.emlid.com/reach/ [accessed 9th May 2018].
European Communities, 2007. *Co-operation Agreement On a Civil Global navigation Satellite System (GNSS) between the European Community and its Member States and Ukraine Overview*. Available from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/243190/7199.pdf [accessed 9th May 2018].
Google, 2018. *ARCore Overview*. Available from https://developers.google.com/ar/discover/ [accessed 9th May 2018].
Goshtasby, A. Ardeshir, 2012. *Image Registration - Principles, Tools and Methods*. Springer, London.
Hesse, S., and Schnell, G., 2014. *Sensoren für die Prozess- und Fabrikautomation*. Translation: Sensors for process and factory automation. Wiesbaden, Germany: Springer Vieweg.
Huawei, 2018. *E5573*. Available from https://consumer.huawei.com/lk/support/mobile-broadband/e5573/ [accessed 9th May 2018].
ITK, 2018. *Insight Segmentation and Registration Toolkit (ITK)*. Available from https://itk.org/ [accessed 9th May 2018].
ImageJ, 2018. *ImageJ*. Available from https://imagej.net/ [accessed 9th May 2018].
Kalman R., 1960. A new Approach to Linear Filtering and Prediction. *Journal of Basic Engineering*: 35-45.
Keyglove, 2011. *Acceleration, Velocity, Position and Dead Reckoning*. Available from http://www.keyglove.net/2011/10/12/acceleration-velocity-position-and-dead-reckoning/ [accessed 9th May 2018].
Madgwick S., 2010. *An efficient orientation filter for inertial and intertial/magnetic sensor arrays*. Available from http://x-io.co.uk/res/doc/madgwick_internal_report.pdf [accessed 9th May 2018].
Maeda, M., Ogawa, T., Kiyokawa, K., and Takemura, H., 2004. *Tracking of user position and orientation by stereo measurement of infrared markers and orientation sensing*. 8th International Symposium on Wearable Computers, 77-84. 31. October – 03 November 2004, Washington D.C., USA.
Magic Leap, 2018. *Magic Leap One*. Available from https://www.magicleap.com/ [accessed 9th May 2018].
Melbourne, A., Ridgway, G., and Hawkes, D.J., 2010. *Image Similarity Metrics*. In Proc. of the SPIE.
Metavision, 2018. *Meta 2 Augmented Reality Development Kit*. Available from https://meta-eu.myshopify.com/products/meta2 [accessed 9th May 2018].
Microsoft, 2018a. *Microsoft HoloLens*. Available from https://www.microsoft.com/en-us/HoloLens [accessed 9th May 2018].
Microsoft, 2018b. *Spatial mapping design*. Available from https://docs.microsoft.com/en-us/windows/mixed-reality/spatial-mapping-design [accessed 9th May 2018].
Microsoft, 2018c. *What´s a Universal Windows Platform (UWP) app?*. Available from https://docs.microsoft.com/en-us/windows/uwp/get-started/universal-application-platform-guide [accessed 9th May 2018].
Nestares, O., Gat, Y., Haussecker, H., and Kozintsev, I., 2010. Video stabilization to a global 3D frame of reference by fusing orientation sensor and image alignment data. In *IEEE International Symposium on Mixed and Augmented Reality*, Seoul, 2010, pp. 257-258.
Qiu S., Wang Z., Zhao H., Qin K., Li Z., Hu H., 2008. Inertial/magnetic sensors based pedestrian dead reckoning by means of multi-sensor fusion. *Information Fusion*, Vol. 39: 108-119.
OpenCV, 2018a. OpenCV. Available from https://opencv.org/ [accessed 9th May 2018].
OpenCV, 2018b. Motion Analysis and Object Tracking. .Available from https://docs.opencv.org/2.4/modules/imgproc/doc/motion_analysis_and_object_tracking.html [accessed 9th May 2018].
OpenPnP, 2017. *OpenCV 3.2.0 (packaged by OpenPnP)*. Available from https://github.com/openpnp/opencv [accessed 9th May 2018].
Papagiannakis G., Singh G., Magnenat-Thalmann N., 2008. A survey of mobile and wireless technologies for augmented reality systems. *Computer Animation and Virtual Worlds*, Vol. 19: 3-22.
Pedley M., 2013. *Tilt sensing using a three-axis accelerometer*. Available from: https://cache.

freescale.com/files/sensors/doc/app_note/AN3461.pdf [accessed 9th May 2018]
Polischuk G., Kozlov V., Ilitchov V., Kozlov A., Bartenev V., Kossenko V., Anphimov N., Revnivykh S., Pisarev S., Tyulyakov A., Shebshaevitch B., Basevitch A., Vorokhovsky Y., 2002. *The global navigation satellitesystem GLONASS: development and usage in the 21st century*. 34th Annual Precise Time and Time Interval (PTTI) Meeting: 151-160, The Institute of Navigation Inc, Manassas, USA.
PTC Inc., 2018a. *Vuforia*. Available from https://www.vuforia.com/ [accessed 9th May 2018].
PTC Inc., 2018b. *Image Targets*. Available from https://library.vuforia.com/articles/Training/Image-Target-Guide [accessed 9th May 2018].
Praschl C., 2017. *Bildbasierte Orientierung im Outdoorbereich*. Translation: Image-based Outdoor Orientation. Bachelor´s Thesis. University of Applied Sciences Upper Austria.
Pivotal Software, 2018. *Spring*. Available from https://spring.io/ [accessed 9th May 2018].
Sutherland I., 1968. *A head-mounted three-dimensional display*. Fall Joint Computer Conference, 757-764. 09 – 11 December 1968, Salt Lake City, USA.
Tomaszewski, D., Rapinski J., and Pelc-Mieczkowska R. I., 2017. Concept of AHRS algorithm designed for platform independent IMU attitude alignment. *Reports on Geodesy and Geoinformatics*, Vol 104: 33-47.
Tsui J., 2005. *Fundamentals of Global Positioning System Receivers: A Software Approach*. New Jersey, USA: Wiley.
U.S. Air Force, 2008. *Global Positioning System Wide Area Augmentation System Performance Standard*. Available from https://www.gps.gov/technical/ps/2008-WAAS-performance-standard.pdf [accessed 9th May 2018].
Unity Technologies, 2018a. *Unity*. Available from https://unity3d.com/ [accessed 9th May 2018].
Unity Technologies, 2018b. *Camera.main*. Available from https://docs.unity3d.com/ScriptReference/Camera-main.html [accessed 9th May 2018].
Unity Technologies, 2018c. *GameObject*. Available from https://docs.unity3d.com/ScriptReference/GameObject.html [accessed 9th May 2018].
Wikitude, 2018. *Wikitude*. Available from https://www.wikitude.com/ [accessed 9th May 2018].
WIT Motion, 2018. *Attitude Sensor*. Available. from http://www.wit-motion.com/english.php?m=goods&a=details&content_id=96 [accessed 9th May 2018].
x-io Technologies, 2016. *Next Generation IMU*. Available from http://x-io.co.uk/ngimu/ [accessed 9th May 2018].
x-io Technologies, 2017. *NGIMU-Unity-Example*. Available from https://github.com/xioTechnologies/NGIMU-Unity-Example/blob/master/Simple%20Demo/Assets/Code/UnityHelper.cs [accessed 9th May 2018].
Xu H., 2012. *Application of GPS-RTK Technology in the Land Change Survey*. International Workshop on Information and Electronics Engineering, 3454 – 3459. 10. -11. March 2012, Harbin, China.
Zaidi, M., Rached, T., and Ouni, R., 2010. *A new geometric approach to mobile position in wireless LAN reducing complex computations*. 5th International Conference on Design & Technology of Integrated Systems in Nanoscale Era, 1-7. 23-25 March 2010, Hammamet, Tunisia.

AUTHORS BIOGRAPHY

Christoph Praschl, born in St. Pölten, Austria attended the University of Applied Sciences Upper Austria, Campus Hagenberg where he studied software engineering and graduated BSc. in 2017. In 2017 he started the master's degrees studies of software engineering and since then he is a research assistant in the research group AIST.

Oliver Krauss, born in Salzburg, Austria attended the University of Applied Sciences Upper Austria in the Software Engineering bachelor and master programs. He joined the research group AIST in 2012 as a research associate, and has been a lecturer in the Software Engineering and Medical and Bio-informatics programs since 2014. His lectures include teaching AR/VR in medical and industrial applications. He is currently enrolled in the PhD program at the Faculty of Engineering & Natural Sciences at the Johannes Kepler University Linz, in the field of genetic improvement of performance in interpreters and compilers.

Gerald A. Zwettler was born in Wels, Austria and attended the University of Applied Sciences Upper Austria, Campus Hagenberg where he studied software engineering for medicine and graduated Dipl.-Ing.(FH) in 2005 and the follow up master studies in software engineering in 2009. In 2010 he has started his PhD studies at the University of Vienna at the Institute of Scientific Computing, where he received his degree in December 2014. Since 2005 he is working as researcher and lecturer at the Upper Austrian University of Applied Sciences at the school of informatics, communications and media at the Campus Hagenberg in the field of medical image analysis and software engineering with focus on computer-based diagnostics support and medical applications. Since 2017 he is associate professor at the software engineering department and since 2018 head of research group AIST in the field of advanced information systems and technology.

BIM-ENABLED HEALTH & SAFETY ANALYSIS OF CROSS LAMINATED TIMBER ONSITE ASSEMBLY PROCESS

Tsvetomila Duncheva[a], Samer BuHamdan[b], Robert Hairstans[c], Mohamed Al-Hussein[d]

[a] [c] Edinburgh Napier University, Edinburgh, Scotland, UK
[b][d] University of Alberta, Edmonton, Canada

[a] m.duncheva@napier.ac.uk, [b]buhamdan@ualberta.ca [c]r.hairstans@napier.ac.uk
[d]malhussein@ualberta.ca

ABSTRACT
There is a global need for sustainable urban housing and offsite timber systems such as Cross Laminated Timber construction can be part of the solution to this need. However, the health and safety (H&S) impacts of CLT installation technologies have not yet been investigated utilising BIM enabled. This research project used a case study method to analyse the constructability of CLT panels installation and a System Dynamics (SD) simulation model was construed to analyse workers' time spent using hand tools for prolonged time, with potential health impacts. The results demonstrated that CLT installation H&S impacts could be correlated to CLT connections specification. These findings shine new light on Design for Safety (DfS) considerations for CLT medium rise construction with regards to social sustainability, in the context of urban residential buildings.

Keywords: Building Information Modelling/Management (BIM), Cross Laminated Timber (CLT), Social sustainability, Design for Safety

1. INTRODUCTION

Across the globe there is a race to provide housing for the enlarging population, the pressures for which are especially strong in the hotspots of economic growth, namely large cities. Urban centres of the likes of London, Boston, Bombay, Sydney and Beijing are all facing the challenges of population increase with higher rates than housing can be provided. In major cities in the southern hemisphere this trend has portrayed by the enlargement of informal settlements (King et al. 2017). One of the key strategies for coping with this urbanisation challenge is medium-density residential construction which needs to encompass all three sustainability branches; environmental, social and economic (UN 2016).

Offsite construction has been suggested as part of the solution to the housing crisis (Miles & Whitehouse 2013; Smith 2014). Mass timber construction in particular has the highest environmental sustainability credentials because of the high utilisation of a renewable resource, the production of which transforms CO_2 into O_2, namely wood. In addition, in offsite timber construction, a part of the building process is transferred from the building site to the factory environment, where there are opportunities for application of resource-efficient production methods developed in the manufacturing and automotive industries (Hairstans 2015; Womack & Jones 2003).

2. LITERATURE REVIEW

2.1. Offsite timber construction

With the change in production building method there are increased opportunities for automation, digitisation and control. As a result of these innovations, often cited offsite construction attributes are speed onsite, predictable costs and energy-efficient building fabric performance (Kamali & Hewage 2016; Dodoo et al. 2014). Moreover, an advantage of offsite construction with a high percentage of prefabrication, is improved safety, due to reduced need for working at heights and use of equipment for strenuous manufacturing and assembly tasks (Schoenborn 2012). However, the H&S aspects of the onsite assembly processes have been little discussed in previous research.

2.1.1. CLT for tall residential construction

Cross-laminated timber is a type of mass offsite timber system, in which lamellae are glued in perpendicular grain direction to each other (Hairstans 2018). CLT is an engineered-timber product, whose higher strength and stiffness properties allow for utilisation of engineered timber as the main superstructure material in increasingly tall buildings. For example, the 18-storey 53m-tall student accommodation building at the University of British Columbia, Vancouver, Canada combined Building Information Modelling/Management (BIM) with CLT construction to deliver the tallest timber building in the world at the time of writing (Fallahi et al. 2016). Furthermore, the construction of CLT buildings up to 150 meters in combination with a concrete core has been theorized for high-density urban environments (Van De Kuilen et al. 2011).

2.1.2. CLT modelling and optimization

There are many research studies on CLT production, and structural optimization. For example, Crawford and colleagues (2015) investigated the potential to produce

CLT from home-grown timber resources in Scotland, whereas Izzi and colleagues (2016) calculated strength factors of nailed CLT connections. In addition, the integration of shear tests for the lamination of CLT panels has been investigated by comparison of test results with desktop study calculation results, to propose practical testing methods and their specimen size considerations (Betti et al. 2016).

Optimization studies have also been conducted on CLT for economic factors. Composite structures with CLT panels and supporting timber ribs have been demonstrated to minimise structural volume of CLT material for compliance with Eurocode 5 (EC5) (Stanić et al. 2016). Best-practice production methods have been outlined in previous research, including finger-jointing, adhesive application and hydraulic or vacuum pressing, with emphasis on quality control procedures for guaranteed product speciation (Brandner 2014). Moreover, increases in the level of prefabrication of CLT panels by inclusion of façade elements in the factory manufacturing process has been shown to result in construction programme acceleration (Gasparri et al. 2015).

However, fewer studies have investigated the health and safety (H & S) aspects of CLT construction. Fire safety is one of the main H & S areas, which are limiting the opportunities for tall timber building construction. Specifically, more knowledge needs to be contributed to the areas of system-level fire performance, interfaces between CLT and composite construction systems and fire-stopping of openings for services within the CLT panels (Barber & Gerard 2015). Yet the impacts of CLT construction and specifically different connection options on onsite installers' health have not been the subject of previous studies.

2.2. Building Information Modelling/ Management

BIM interfaces with offsite construction methods through the increase in digitisation, automation and manufacturing in construction (Vernikos et al. 2014). BIM offers opportunities for increased understanding of site conditions during construction through analysis of site environmental factors and visualisation of the project, with risk levels as an overlay to 3D virtual models or 4D construction schedules (Hardin & McCol 2015).

For example, opportunities have been investigated to apply BIM technologies and process to automate fall prevention from slab edges using guard rails installation (Zhang et al. 2015). Moreover, regarding the question of how such technologies can be practically applied in construction management, previous research has found that onsite simulations have potential to be integrated with 'toolbox' meetings, at which teams are gathered to discuss the health and safety requirements prior to task start (Ganah & John 2015).

Further digitisation advances for BIM safety offer opportunities for improved safety management on construction sites online databases, virtual reality, overlaid 4D schedules and active instead of passive PPE enabled by sensing and warning technologies (Zhou et al. 2012). However, digital tools for the reduction of H&S risks through design are not as sophisticated as those for application during the construction stage. Process flowcharts have been suggested as simple yet effective tool to identify the drivers for potential safety hazards and remove them through Design for Safety (Malekitabar et al. 2016).

2.3. Construction safety simulation

The research studies outlined above investigated the optimization of CLT as a product or the integration of BIM practices with Safety management, however further opportunities for offsite systems H&S optimization lie in research of construction processes with the use of simulation models.

Although not focused on offsite systems, several research studies have investigated simulation models which aimed to capture the complexities of workers' safety behaviour on site (Goh & Askar Ali 2016; Guo et al. 2016; Mohammadfam et al. 2017). Because of the universal time over-runs of construction programmes, a significant factor which affects safety behaviour is the resulting pressure on workers to expedite their tasks and the co-relation between production pressure and accident occurrence has been proven through a System Dynamics (SD) model (Han et al. 2014).

2.4. The gap in knowledge

Overall, as Table 1 shows, there is gap in knowledge on the integration of the four outlined themes of CLT construction, BIM, Simulation Research Methods (SRM) and occupational health and safety.

Table 1: Literature review themes and gap analysis

Reference	CLT	BIM	SRM	H & S
(Fallahi et al. 2016)	✓	✓		
(Izzi et al. 2016)	✓			
(Betti et al. 2016)	✓		✓	
(Stanić et al. 2016)	✓		✓	
(Brandner 2014)	✓			
(Crawford et al. 2015)	✓			
(Barber & Gerard 2015)	✓			✓
(Gasparri et al. 2015)	✓		✓	
(Van De Kuilen et al. 2011)	✓		✓	
(Schoenborn 2012)	✓			✓
(Ganah & John 2015)		✓		✓
(Malekitabar et al. 2016)		✓		✓
(Zhou et al. 2012)		✓		✓
(Zhang et al. 2015)		✓		✓
(Goh & Askar Ali 2016)			✓	✓
(Guo et al. 2016)			✓	✓
(Han et al. 2014)			✓	✓
(Mohammadfam et al. 2017)			✓	✓

3. METHODOLOGY

The objective of the present research is to develop a decision support model to assist designers and project management evaluating the occupational hazards associated with CLT panels installation due to nailing and screwing efforts associated with a certain design of the structural connections between the panels.

The information required for the simulation process are fed from two input points; a database that is linked with a BIM model that contains geometrical information pertaining to the design and the user who provides information related to the design of the structural connections. The principles of this paper were validated during an in-depth qualitative interview with an industry expert in DfMA design and project management of CLT residential buildings

The overall methodology utilized in this research is presented in Figure 1.

Based on the provided information the simulation model emulates the entire installation process of the building, so the potential exposure time to occupational hazards can be evaluated and, therefore, the associated risk can be assessed. Once this information is known the designers and/or project management team can evaluate their practices accordingly to reduce the correspondent hazards.

3.1. The architecture of the simulation model

The simulation model uses discrete simulation in Simphony.Net environment to model the construction operations related to CLT panels installation. The model consisted of interconnected sub-modules, where each sub-module was responsible for a certain phase of the construction process.

3.1.1. Weather conditions sub-nodule (WCSM)

Craning operations are indispensable in offsite construction. These operations are sensitive to weather conditions, chiefly, wind speed and gusts that halt craning work when in excess of safe working limits. Accounting for weather conditions within the simulation engine is essential for the reliability of outcomes. For this purpose, the weather sub-module generates discreate events that follows the wind patterns prevailing in the area where the construction takes place (wind speed is modelled as a statistical distribution that is obtained from fitting the meteorological data). Once the wind speed of a generated incident exceeds the maximum allowable limit for craning it triggers the construction sub-module to stop craning operations until the wind speed goes back to the working limits. Figure 2 shows the weather sub-module used in the simulation model.

Figure 2: Weather conditions sub-nodule

3.1.2. Construction operations sub-module (COSM)

The core of the simulation engine is the construction operations sub-module, where the major effort of simulation is happening. As the purpose of developing the simulation model is to evaluate the health hazards associated with CLT installations, the focus of the COSM is on the installation task.

The simulation process begins once the panels are delivered on-site. Panels are usually delivered to the site following the installation sequence, except for some discrepancy that impose storing the panels until it is time for them to be installed. The COMS addresses this issue

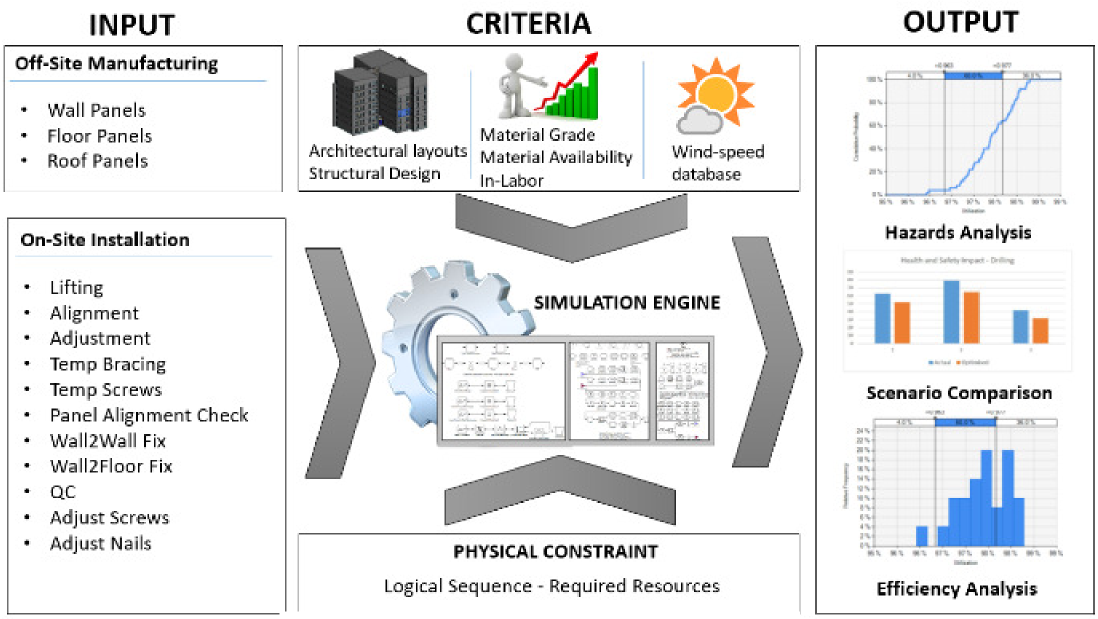


Figure 1: Research Methodology

using a probabilistic composition that assesses the likelihood of the wrong delivery of the panels and incorporates it to the correspondent sub-module. The next task for the COSM is to mimic the lifting process, where it interacts with the WCSM for safe-working conditions. Once the panels are in the designated place the next step is a nailing and/or screwing process to be fixed in place. The model simulates both process independently to allow for collecting more customized data. To simulate the installation process, the COSM requires the following input:

- the vertical (floor number) and horizontal (floor plan location) locations of each wall;
- the function of each wall (e.g. stability load bearing wall, non-load bearing wall, etc.)
- wall connection design per the function and location of the wall;
- wall geometry; and,
- productivity information for installation tasks.

Figure 3 shows the COSM in Simphony.Net Environment. The input for the simulation model, which was described previously, comes from two sources; a BIM model and the user. The following sections will discuss the information exchange with data sources.

3.2. Information exchange with data sources

The efficiency of utilizing simulation as a decision support tool is proportional to the ease of feeding the information pertaining to the various scenarios needed to be evaluated. As demonstrated in 2.1, except of the wind speed database and productivity information, the information required to simulate the installation process is related to the building. The size of the data pertaining to a certain building is proportional to the size and the complexity of that building, and the building grows in size and complexity, the manual acquisition of the required information by the simulation model gets harder. Therefore, the simulation engine is designed to retrieve the building-related infatuation from a BIM model, while productive and wind speed patters are input manually. Section 2.2.1 elaborates more on the link between the simulation model and BIM models. While Section 2.2.2 discuss user-fed information.

3.2.1. Linking BIM models with the simulation engine

BIM models contain accurate data about the building categorized per element. BIM authoring tools support export building data to databases management environments such as MS Access. On the other hand, Simphony.Net support generating entity with properties obtained from an external database. Therefore, MS Access is used as a medium between BIM authoring tools and the simulation engine, where BIM authoring tools export required information into a MS Access database, which Simphony.Net will use to generate entities with the building properties. The following information types are stored in the database:

- ID: a unique identifier from the BIM model, which is also used for horizontal location of the wall;
- type: indicates the function of the wall, which will further define the type of connections needed;
- length, where the generic information is stored; and,
- floor: to define the vertical location of the element.

This information is acquired for wall and floor panels. This link allows processing large amount of information pertaining to the building in almost no time, which

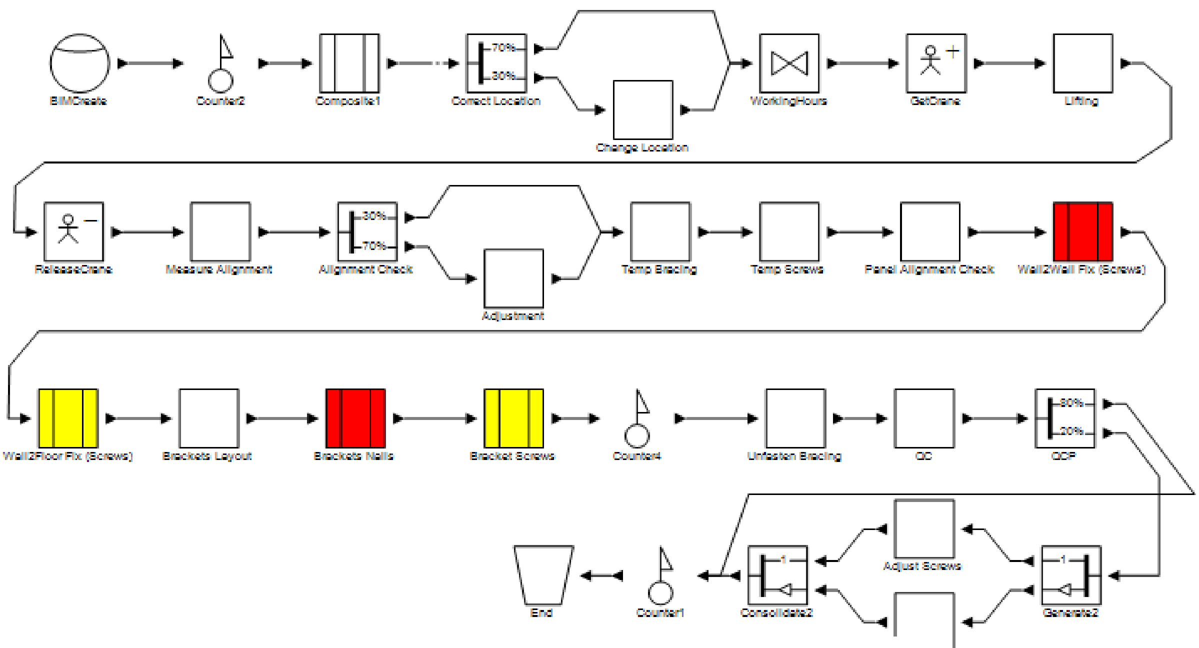


Figure 3: Construction operations sub-module

Proc. of the Int. Conference on Modeling and Applied Simulation 2018,
ISBN 978-88-85741-07-2; Bruzzone, De Felice, Frydman, Longo, Massei and Solis, Eds.

expedite the simulation process as it automates the input process. Furthermore, design nature in the construction industry has an iterative process, which in case of manual feed of information will burden the implantation of the developed model during assessment. Hence, the developed link with the BIM models relive this burden.

3.2.2. User-fed information

In cases where the information is required by the simulation engine is global (i.e. applies to all entities in the model), the user can manually input this information. The information that falls within the global nature is as follows:

- Productivity (labour-minutes per connection);
- connection design per panel type;
- distribution functions that predict wind speed patterns.

This information is input once per scenario and therefore the need for automation is not pressing and won't affect the ease of implementing the developed model. The statistical distribution of wind data generation is an area for further work in the research, which will be included in a subsequent journal publication.

The last section of the methodology deals with analysing simulation generated data.

3.3. Analysing simulation data for occupational health and safety by design

Research into hand-arm vibration syndrome (HAVS) has revealed the clear connection between increased exposure to vibration tools among joiners and construction workers and increased risk of HAVS (Palmer et al. 1999). This '99 report produced for the Health and Safety Executive (HSE) indicated that 4.2 million male and 667,000 female workers were exposed to hand-transmitted vibration at work, of which carpenters and joiners were the second-largest male group, after welders. The report findings demonstrated that there was a nearly 2 times higher rate of cold-induced finger blanching in carpenters and joiners than in workers not exposed to vibration tools. In carpenters and joiners the rate was especially high, between 1.89 and 3.37.

According to the HSE, more than 600 new people were reported to suffer from vibration 'white finger', cumulatively more than 7,500 people in the ten years between 2006 and 2015, shown in Figure 4 (HSE 2016). Other health problems associated with exposure to vibration are carpal tunnel syndrome and occupational deafness (ICE 2015). According the ICE, the only way designers can prevent Occupational Health and Safety Hazard such as HAVS, is by eliminating or reducing the 'need for workers to use hand-held vibrating tools or machinery. Specific to CLT projects this would translate as optimizing the screw connections in the superstructure.

It is speculated that the risk of HAVS and other occupational health and safety hazards associated with use of vibration tools in carpentry increases incrementally with the time spent using vibration tools, which can be measured as a percentage of overall labour-hours, as outlined in Equation 1.

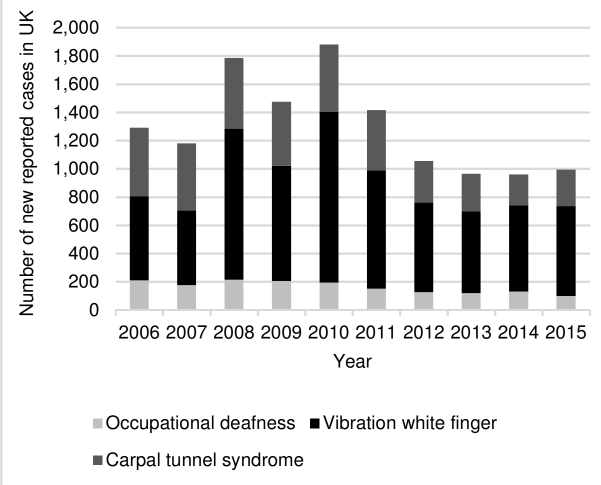


Figure 4: Health impacts associated with vibration tools. Authors' own graph based on data from (HSE 2016).

$$\%\ \textit{time spent using vibration tools} = \%\ \textit{increase in chance of HAVS symptoms} \quad (1)$$

The simulation model provides time information pertaining to the performed tasks based on input building data. The hazard associated with the screwing and nailing tasks are assessed based on (1) exposure times to nailing and/or screwing tasks and (2) number of the workers who are exposed to the mentioned tasks.

3.3.1 Limitations

To speculate more accurately how much the chance of HAVS increases with use of vibration tools, further research is needed in establishing the correlation between time spent using equipment and the likelihood of HAVS symptoms. For example, the number of carpenters and joiners working in construction in the latest year available year 2015 in the UK, needs to be compared to the number of reported HAVS cases. Moreover, the average time spent using vibration tools of carpenters and joiners on construction sites should be compared to the time spent using vibration tools in the case study building. In addition, not all connections will have an equal impact on the health & safety of workers, from the observations in this project it can be stated that connections at angles other than perpendicular to the surface are more strenuous. This approach will be the object of further study.

4. CASE STUDY

An innovative CLT urban residential building was the case study selected to demonstrate the simulation methodology. This was the Yoker project, the tallest CLT building in Scotland at the time of writing at 7 storeys tall. The building site is situated on the north bank of the river Clyde and a small ferry is in operation, which takes passengers across to the opposite bank. The building is oriented along the north-south axis with a T shape, which maximises views over the river Clyde. The area of Yoker is mainly residential and its high street feature the characteristic of Glasgow red-stone tenement buildings as shown in Figure 5.

Figure 5: Yoker project in context. Courtesy of CCG.

4.1. Input data

4.1.1. Tasks duration

To account for the stochastic nature of the construction activity, a triangular distribution was used for the installation tasks durations. Table 2 shows the parameter of the triangular destructions used for each of the tasks.

Table 2: Duration of task used in the simulation model

Task name	Duration (Triangular Distribution Parameters in min)		
	Low	average	High
Correct location	0.5	1	3
Change location	2	5	8
Lifting	1	2.5	4
Measure alignment	0.25	0.5	1
Adjustment	0.5	1	2
Temp Bracing	1	2	4
Temp Screws	1	1.5	2
Panel Alignment Check	0.25	0.5	1
Wall2Wall Fix	NS*0.25	NS*0.5	NS*0.75
Wall2Floor Fix	NS*0.75	NS * 1	NS*1.25
Brackets Layout	NB*0.25	NB*0.5	NB*1
Brackets Nails	NN*0.25	NN*0.5	NN* 0.75
Brackets Screws	NS*0.25	NS*0.5	NS*0.75
Unfasten Bracing	1	1.5	3
QC	120	240	360
Adjust Screws	NS*0.5	NS * 1	NS*1.5
Adjust Nails	NS*0.5	NS*1	NS*1.5

Where;

- NS is the number of screws in one connection plate.
- NN is the number of nails in one connection plate.
- NB is the number of brackets to be installed on one panel.

4.1.2. BIM model

The BIM model of the building, shown in Figure 6, was prepared by the architects and was used to overlay the engineers' CLT model with the architectural model. As the model was delivered in IFC format the authors had to modify the definition of the element and added information to enable export to the developed database.

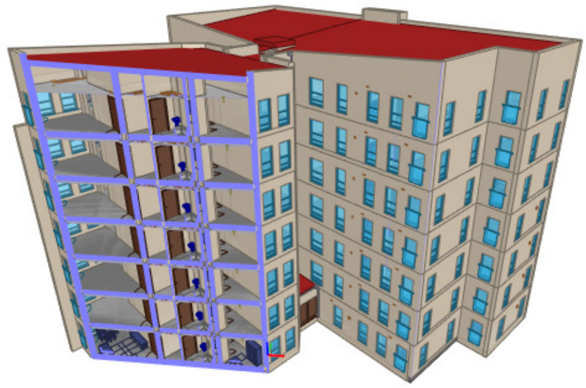


Figure 6: The BIM model of the case study building

4.1.3. Wind speed patterns

The site is located in the west outskirts of Glasgow and is to the south-west is exposed to the river. This is reflected in the wind rose diagram shown in Figure 7. The image shows that higher frequency and velocity winds are oriented from the SW, the same direction as the central axis of the building, facing the river. Wind data was sourced from an online weather database in the public domain (MeteoBlue 2018). Typically, in the area there are 34 days per year with wind speeds above 30 m/h, at which crane operations needs to stop, concentrated between November and March (5 months), with much fewer high-wind days in the spring and summer months.

During the installation of the CLT panels in the case study building 5 days were lost due to high winds. This is less than the typical 8 high wind days in the months March to May, and could be explained by some high wind days falling on non-working days, and/or a spring season with lower than usual wind frequency and velocity in 2017.

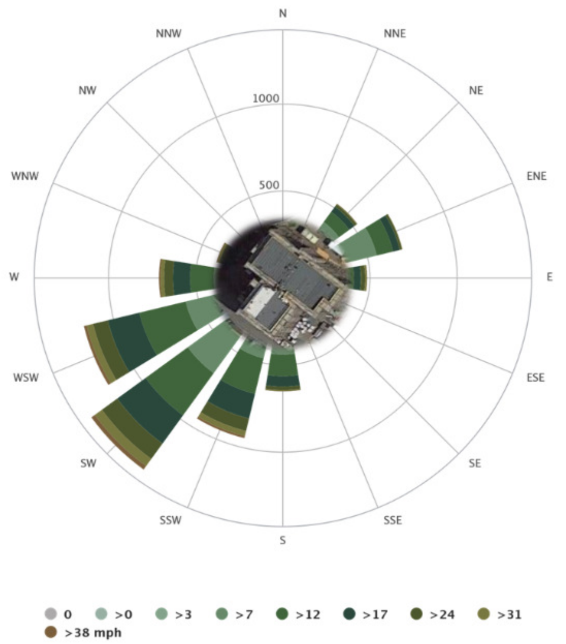


Figure 7: Case study building in wind rose context

4.2. Simulation output

The model is set to run 100 times, where Figure 8 shows finish times of the project through the multiple runs. The installation duration is 52,711min with a 95% confidence interval of [52471,52951]. On the other hand, Table 3 demonstrates the percentage of the total project duration that was spent on screwing and nailing of panels. As per Equation 1, the increase in the HAVS symptoms can be estimated as per Table 3. Considering the information pertaining to the increase in the HAVS symptoms, decision makers can adjust the design of the connection between the panels to reduce the likelihood of having HAVS symptoms.

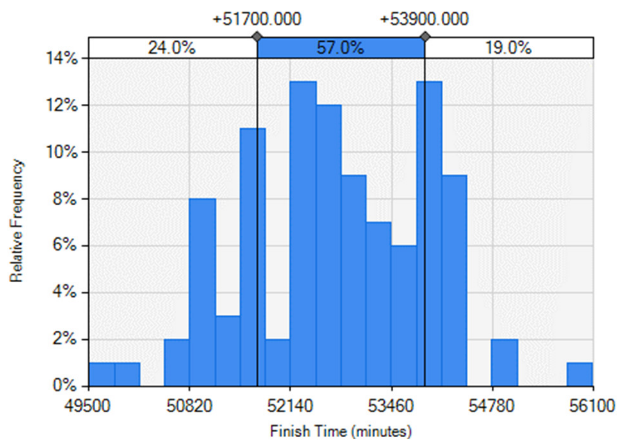


Figure 8: Finish times of the project

Table 3: Presentation of screwing and nailing tasks

Task	Average Utilization	Standard Deviation	Maximum Utilization
Screw Floor 0	2.10%	0.10%	2.30%
Screw Floors 1,2,3	92.10%	1.20%	95.10%
Screw Floors 4,5,6	37.80%	1.40%	41.20%
Nailing Floor 0	31.60%	1.20%	34.40%
Nailing Floors 1,2,3	23.10%	1.00%	26.40%
Nailing Floors 4,5,6	24.00%	1.10%	26.30%

Table 4: the increase in the HAVS symptoms

Task	Increase Percentage
Screw Floor 0	2.10%
Screw Floors 1,2,3	92.10%
Screw Floors 4,5,6	37.80%
Nailing Floor 0	31.60%
Nailing Floors 1,2,3	23.10%
Nailing Floors 4,5,6	24.00%

ANALYSIS AND DISCUSCUSSION

To interpret the meaning of these results, it should be noted that the CLT installation was sequenced in to phase, first the panels were installed using temporary connectors by half the team, and then connections were finished by the rest of the team members. Therefore, only the CLT panels assembly was on the critical path and has a higher rate of completion compared to the connectors finishing. With view of the differences in skill requirements and the criticality of the tasks, the panel assemblers were senior in expertise, whereas the second connections crew were junior and were in training. These associations were revealed during a qualitative validation interview with an industry expert.

Therefore, most of the effect of these results would have been experienced by the most junior staff, subjected to the highest number of repetitive tasks. However, further research would be needed to decipher the correlation between individuals' use of vibration equipment and their respective likelihood of HAVS symptoms development. Such studies could perhaps be undertaken using emerging technologies such as smart watches, smart safety helmet and others (Robinson et al. 2016). With view of the urgent need for recruitment of new talent in the construction industry underpinned by the skills gap in construction, this approach seems anti-productive. If those training to become construction professionals are exposed to monotonous tasks which potentially increase the likelihood of H&S impacts, even if substantial dampeners are used to mitigate the H&S impacts, the nature of the task at hand may discourage the new entrants in the construction industry. A recommendation is made that the necessary practice in precise, accurate and quick connections installation is complemented with utilisation of technology such as Augmented Reality BIM to overlay (Hardin & McCol 2015). Their responsibilities can moreover be enhanced by providing young trainees with tablets and training on how to input daily component-based progress statuses, reflected in 5D BIM models, showing 3D + Time + Completion Status, which could be reflected using different colour schemes. A similar visualisation approach has already been proposed in the context of energy utilisation visualisation.

Importantly, the results suggested that the screws on Floors 1,2 and 3 could have imposed the highest H&S risks to the CLT installers and a correlation is hypothesised between the high number of complex connectors on those floors and their H&S impact. However, a careful examination of the differences in detailing with consideration for structural requirements will be the object of further study by the authors.

5. CONCLUSION

Using cross limited timber as construction material is gaining increasing recognition from construction practitioners and researchers due to its low environmental impact and improved levels of constructability, and with the increasing demand on CLT in the construction project it is important to assess the H&S aspects associated with CLT installation. Therefore, this paper utilizes Building Information Modelling with discrete event simulation to develop a decision support model that assist designers and project management teams to evaluate the potential H&S hazards during CLT installation that are associated with a certain design. With the help of the developed model, designers can test the compliance of their design with health and safety regulations and modify according to the findings, and as a result the implementation of CLT becomes safer and more appealing for a wider range of contractors and owners, improving the sustainable practice in the construction industry.

ACKNOWLEDGMENTS

We sincerely thank the Built Environment Exchange (beX) programme for facilitating the knowledge exchange between Scottish Offsite Timber and Canadian Building Information Modelling/Management (BIM) proficiency. This research project was possible thanks to the construction of an innovative CLT storey medium-rise building in Scotland and we greatly appreciate the co-operation of the contractor's R&D lead and site manager. We are moreover sincerely grateful to Fiona Bradley from University of Glasgow for supervision of Mila Duncheva's PhD. Another critical contribution we are grateful for is the contribution of the architects, engineers and installers of the project, who shared their invaluable knowledge and provided design, specification and inspections documents and a building information model. We would especially like to acknowledge the drive and ambition of the contractor, who pushed the boundaries of tall CLT construction in a Scottish context and provided the basis for this research.

REFERENCES

Barber, D. & Gerard, R., 2015. Summary of the fire protection foundation report - fire safety challenges of tall wood buildings. *Fire Science Reviews*, 4(1), p.5. Available at: http://www.firesciencereviews.com/content/4/1/5.

Betti, M. et al., 2016. Comparison of newly proposed test methods to evaluate the bonding quality of Cross-Laminated Timber (CLT) panels by means of experimental data and finite element (FE) analysis. *Construction and Building Materials*, 125, pp.952–963. Available at: http://dx.doi.org/10.1016/j.conbuildmat.2016.08.113.

Brandner, R., 2014. Production and Technology of Cross Laminated Timber (CLT): A state-of-the-art Report. *Focus Solid Timber Solutions - European Conference on Cross Laminated Timber (CLT)*, (May 2013), pp.3–36.

Crawford, D. et al., 2015. Viability of cross-laminated timber from UK resources. *Proceedings of the ICE - Construction Materials*, 168(CM3), pp.110–120. Available at: http://www.icevirtuallibrary.com/content/serial/coma.

Dodoo, A., Gustavsson, L. & Sathre, R., 2014. Lifecycle primary energy analysis of low-energy timber building systems for multi-storey residential buildings. *Energy and Buildings*, 81, pp.84–97. Available at: http://dx.doi.org/10.1016/j.enbuild.2014.06.003.

Fallahi, A. et al., 2016. Designing for Pre-Fabrication and Assembly in the Construction of UBC's Tall Wood Building. In *2016 Modular and Offsite Construction (MOC) Summit Edmonton, Alberta, Canada, September 29 - October 01, 2016*. Edmonton, Alberta, Canada.

Ganah, A. & John, G.A., 2015. Integrating building information modeling and health and safety for onsite construction. *Safety and Health at Work*, 6(1), pp.39–45. Available at: http://dx.doi.org/10.1016/j.shaw.2014.10.002.

Gasparri, E. et al., 2015. Construction management for tall CLT buildings: From partial to total prefabrication of fa??ade elements. *Wood Material Science and Engineering*, 10(3), pp.256–275.

Goh, Y.M. & Askar Ali, M.J., 2016. A hybrid simulation approach for integrating safety behavior into construction planning: An earthmoving case study. *Accident Analysis and Prevention*, 93, pp.310–318. Available at: http://dx.doi.org/10.1016/j.aap.2015.09.015.

Guo, B.H.W., Yiu, T.W. & González, V.A., 2016. Predicting safety behavior in the construction industry: Development and test of an integrative model. *Safety Science*, 84, pp.1–11. Available at: http://dx.doi.org/10.1016/j.ssci.2015.11.020.

Hairstans, R., 2015. *Building Offsite. An introduction*, Edinburgh. Available at: http://www.buildoffsite.com/outputs/publications/.

Hairstans, R., 2018. *Mass Timber - An Introduction to Solid Timber Laminate Systems*, Edinburgh: Arcamedia. Available at: https://www.arcamedia.co.uk/product-page/mass-timber-an-introduction-to-solid-timber-laminate-systems.

Han, S. et al., 2014. Toward an understanding of the impact of production pressure on safety performance in construction operations. *Accident Analysis and Prevention*, 68, pp.106–116.

Hardin, B. & McCol, D., 2015. *BIM and Construction Management. Proven Tools, methods and workflows* 2nd ed., Indianapolis: John Wiley & Sons Inc.

HSE, 2016. IIDB02 - New Cases of non-lung diseases in England, Wales and Scotland, 2006 to latest year. *IIDB - Industrial Injuries Disablement Benefit Scheme*. Available at: http://www.hse.gov.uk/statistics/tables/index.htm#iidb [Accessed August 6, 2017].

ICE, 2015. *ICE manual of health and safety in construction* 2nd ed. C. McAleenan & D. Oloke, eds., London: ICE Publishing.

Izzi, M. et al., 2016. Experimental investigations and design provisions of steel-to-timber joints with annular-ringed shank nails for Cross-Laminated Timber structures. *Construction and Building Materials*, 122, pp.446–457. Available at: http://dx.doi.org/10.1016/j.conbuildmat.2016.06.072.

Kamali, M. & Hewage, K., 2016. Life cycle performance of modular buildings: A critical review. *Renewable and Sustainable Energy Reviews*, 62, pp.1171–1183. Available at: http://dx.doi.org/10.1016/j.rser.2016.05.031.

King, R. et al., 2017. *Confronting the Urban Housing Crisis in the Global South: Adequate, Secure, and Affordable*, Washington D.C. Available at: https://www.wri.org/sites/default/files/towards-more-equal-city-confronting-urban-housing-crisis-global-south.pdf%0Ahttps://www.wri.org/sites/default/files/Confronting_the_Urban_Housing_Crisis.pdf%0Ahttps://www.wri.org/sites/default/files/towards-more-equ.

Van De Kuilen, J.W.G. et al., 2011. Very tall wooden buildings with Cross Laminated Timber. *Procedia Engineering*, 14, pp.1621–1628. Available at: http://dx.doi.org/10.1016/j.proeng.2011.07.204.

Malekitabar, H. et al., 2016. Construction safety risk drivers: A BIM approach. *Safety Science*, 82, pp.445–455. Available at: http://dx.doi.org/10.1016/j.ssci.2015.11.002.

MeteoBlue, 2018. Climate Yoker. *Scotland, United Kingdom*. Available at: https://www.meteoblue.com/en/weather/forecast/modelclimate/yoker_united-kingdom_11550824 [Accessed May 5, 2018].

Miles, J. & Whitehouse, N., 2013. *Offsite Housing Review*, London. Available at: http://cic.org.uk/news/article.php?s=2013-02-28-cic-presents-housing-minister-with-offsite-housing-review-report.

Mohammadfam, I. et al., 2017. Constructing a Bayesian network model for improving safety behavior of employees at workplaces. *Applied Ergonomics*, 58, pp.35–47. Available at: http://dx.doi.org/10.1016/j.apergo.2016.05.006.

Palmer, K.T. et al., 1999. *Hand-transmitted vibration: Occupational exposures and their health effects in Great Britain*, Sudbury.

Robinson, W.G., Chan, P.W. & Lau, T., 2016. Sensors and sensibility: examining the role of technological features in servitizing construction towards greater sustainability. *Construction Management and Economics*, 34(1), pp.4–20. Available at: http://www.tandfonline.com/doi/full/10.1080/01446193.2016.1139146.

Schoenborn, J.M., 2012. A Case Study Approach to Identifying the Constraints and Barriers to Design Innovation for Modular Construction. , April, pp.2–162.

Smith, R.E., 2014. Off-site and Modular Construction Explained. *National Institute of Building Sciences*, p.6. Available at: https://www.wbdg.org/resources/offsiteconstructionexplained.php.

Stanić, A., Hudobivnik, B. & Brank, B., 2016. Economic-design optimization of cross laminated timber plates with ribs. *Composite Structures*, 154, pp.527–537.

UN, 2016. *World Cities Report 2016*, Nairobi. Available at: http://wcr.unhabitat.org/main-report/.

Vernikos, V.K. et al., 2014. Building information modelling and its effect on off-site construction in UK civil engineering. *Institution of Civil Engineering Proceedings*, 167(MP3), pp.152–159.

Womack, J. & Jones, D., 2003. *Lean Thinking* 2nd ed., London: Simon and Schuster.

Zhang, S. et al., 2015. BIM-based fall hazard identification and prevention in construction safety planning. *Safety Science*, 72, pp.31–45. Available at: http://dx.doi.org/10.1016/j.ssci.2014.08.001.

Zhou, W., Whyte, J. & Sacks, R., 2012. Construction safety and digital design: A review. *Automation in Construction*, 22, pp.102–111. Available at: http://dx.doi.org/10.1016/j.autcon.2011.07.005.

AUTHORS BIOGRAPHY

Tsvetomila Duncheva is a Research Assistant at the Centre for Offsite Construction + Innovative Structure, working on multiple projects, whose main aims are to mitigate the skills shortage in the UK construction industry and to facilitate collaboration between offsite timber companies in Scotland. Mila is now completing her PhD on measuring and comparing the productivity of panelised, mass timber and volumetric offsite timber systems. To achieve this Mila extracted best practice information from international leaders in volumetric in the UK, mainland Europe and Canada. This PhD study is funded through an 'Engineering the Future' scholarship from the University of Strathclyde and supervision is provided through collaboration with the University of Glasgow. Beforehand Mila completed a BA (Hons) Architecture degree with distinction at the University of Kent, Canterbury. Mila undertook two architectural assistant internships in her home country Bulgaria on residential architecture design and BIM integration, respectively. Through the combination of education and practice Mila completed her RIBA Part I training in 2013.

Samer BuHamdan is Post-Doctoral Fellow at the University of Alberta as part of NSERC Industrial Research Chair in the Industrialization of Building Construction. Samer's research utilizes building information modelling (BIM) to evaluate the interactions and inter-relationships among components of a project. This deep study of system components serves as the foundation for the development of a dynamic approach to project planning considering time and costs as targets. The result is the development of a decision making model in the context of complexity theory that captures the long-term quantitative effects of engineering decisions as well as the qualitative aspects of human behaviour. It is expected that the principles underlying this research can be extrapolated to explain and predict the behaviour of any component in a complex system in which there are many competitors and limited resources.

Dr. Robert Hairstans is an Associate Professor and head of the Centre for Offsite Construction + Innovative Structures (COCIS), a team with specialist knowledge in the field of timber technologies and engineering adding value to the supply chain with an emphasis on offsite (modular) prefabrication. His research interests include wood innovation, structural timber systems, offsite manufacture and modular construction with an emphasis on delivering more sustainable communities. Dr Hairstans has delivered over £1M of research, commercial and knowledge exchange activities working in collaboration with industry, academia and external organisations. Dr Hairstans is Co-Chair of the international Modular and Offsite Construction Summit and academic lead of Offsite Solutions Scotland (a local industry timber offsite community of practice employing over 1000 people). His most recent book publication is “Mass Timber: an introduction to solid laminate timber systems”.

Mohamed Al-Hussein is a professor at the University of Alberta, Hole School of Construction Engineering, and a highly sought researcher and consultant in the areas of construction automation, lean construction, process improvement, equipment selection and utilization, CO_2 emission quantification and reduction, urban development, and building information modelling. Dr. Al-Hussein’s current research initiatives include prototyping of automated and semi-automated machinery for fabrication of steel and wood-framed construction components, application of lean and ergonomic principles to improve the safety and productivity of industrialization construction operations, and development of plant layout and process improvement measures for panelized and modular construction. Dr. Al-Hussein holds the NSERC Industrial Research Chair in the Industrialization of Building Construction, and is Director of the Nasseri School of Building Science and Engineering. Dr. Al-Hussein’s research has been published in well over 200 peer-reviewed journals and conference proceedings.

MODELING AND SIMULATION OF CABINET MANUFACTURING PROCESSES: EVALUATION AND RECOMMENDED CONTROLS

Osama Mohsen[(a)], Sina Abdollahnejad [(b)], Narges Sajadfar [(c)], Yasser Mohamed [(d)], Simaan AbouRizk [(e)]

[(a),(c)] PhD candidate, Hole School of Construction Engineering, University of Alberta, Canada
[(b)] M.Sc. student Hole School of Construction Engineering, University of Alberta, Canada
[(d),(e)] Professor, Hole School of Construction Engineering, University of Alberta, Canada

[(a)]omohsen@ualberta.ca, [(b)]sabdolla@ualberta.ca, [(c)]sajadfar@ualberta.ca
[(d)]yasser.mohamed@ualberta.ca [(d)]abourizk@ualberta.ca

ABSTRACT
Project uniqueness and high degree of customization have always been a challenging aspect of construction projects and many related operations. This paper describes the simulation of a production line in a kitchen cabinet manufacturing facility that was carried out with the aim of better understanding and improving the production process. Discrete event simulation (DES) using Simphony.NET, a simulation modeling tool developed at the University of Alberta, is used to investigate and analyze an existing facility process to increase manufacturing productivity, reduce work-in-progress, and decrease idle time. The cabinet manufacturing company in the presented study operates different production lines, produces different product types, and uses varying materials. In the case study, specifically, the simulation tool is used to explore the challenges associated with increasing production to satisfy the rising demand of customized products. The result of the simulation study provides valuable information to achieve this goal.

Keywords: production process simulation, kitchen cabinet manufacturing, construction manufacturing, workflow improvement

1. INTRODUCTION

Customization is a challenging aspect inherent to construction projects and related products. In recent decades, the construction industry has been seeking methods and techniques to reduce process waste, increase productivity, improve quality, reduce cost, and ensure the timely delivery of projects. One popular and effective approach is the application of manufacturing principles and techniques to the construction industry—a paradigm that is associated with offsite construction methods such as modularization.

In offsite construction, components for residential, commercial, or industrial buildings are produced in a controlled factory environment. Components are then delivered to work sites, in sequence, for installation. This approach reduces construction waste, improves product quality, and minimizes onsite safety incidents (Koskela, 1992). By adopting modular construction methods, structures can be erected quickly and efficiently without requiring resources to build them to be co-located within the construction site. This, typically, results in the construction of a high-quality buildings within a short amount of time (Mohsen et. al 2008). Despite these advantages, offsite construction is associated with several challenges, especially with respect to the efficient adaptation to product uniqueness as well as the coordination between factory operations and onsite activities. Other challenges involve the need for seamless flow of information beginning at the design stage, proceeding to the factory production lines, and finishing at site delivery and installation, together with a systematic and robust feedback mechanism for future improvements. The successful completion and profitability of projects completed in this manner are highly dependent on the design and implementation of efficient and agile business and manufacturing processes that adapt effectively to a wide range of product designs.

In this context, the processes and operations of a kitchen cabinet manufacturing facility (hereafter referred to as "the facility" or "the manufacturer") are analyzed and simulated. The facility provides complete cabinetry solutions, including kitchen cabinets, bathroom vanities, and hutches, to individual customers, home builders, and trade contractors. Contending with rapidly broadening product demand, the manufacturer is facing various operational and customer satisfaction challenges related to quality, order fulfilment, supply chain, material handling, and labor productivity and stability. The goal of the research presented in this paper is to analyze and simulate the current-state operations and to recommend improvements based on the DES study.

2. STUDY BACKGROUND

Market conditions are rapidly changing, and the demand for increasingly customized and unique products is expanding. More and more, customers are demanding customized dwellings to reflect their cultural tastes and personal preferences. Cabinets in bathrooms, kitchens, or otherwise constitute a large portion of the visible customization in which customers are interested.

Mass customization is the process of delivering products that are modified to satisfy specific customer needs (Phuluwa et al. 2013). Mass customization is a marketing and manufacturing technique that combines the flexibility and personalization of custom-made products with the low unit costs associated with mass production (Pine, 1993). Kotha (1996) concluded that, in the implementation of mass customization, several external and internal conditions are necessary for successful achievement of mass customization, which include, but are not limited to, (i) close proximity of suppliers; (ii) industry-wide increase in number of products; (iii) development of an information network with a selected group of retailers; (iv) investment in manufacturing and information technologies and human resources; (v) knowledge creation to develop manufacturing capabilities; and (vi) marketing efforts to promote individualized products. Benros & Duarte (2009) developed a framework for the mass customization of the housing industry that seeks to achieve customization by combining flexible design, data communication, and industrialization of building processes.

The ability to experiment with different scenarios and analyze the consequences of implementing various tools and techniques intended to improve the production of building components is very valuable. Over the past few decades, discrete-event simulation (DES) has been widely and increasingly used in the construction and manufacturing industries to emulate the various activities and processes of real-world systems. Tremendous amount of research work has been performed in the areas of simulation language development, simulation model design, model optimization, and to combine statistical design-of-experiment techniques with simulation modeling (Callahan et. al 2006).

Altaf et al. (2014) applied DES using Simphony.NET for the purpose of investigating various resource allocation strategies at a panelized construction factory. In addition, Altaf et al. (2015) and Altaf et al. (2017) developed an online simulation-based production control system in a wall panel prefabrication factory using RFID technology. This system evaluates production performance based on real-time data acquired by the RFID system. Kamaruddin et al. (2011) used the WITNESS simulation package to examine the performance of different layout designs relating to two experimental factors: (i) model variability and (ii) headcount variability. They concluded that these two factors do impact the performance measures (throughput time, lateness, and labor productivity) of the flow line, job shop, and cellular layout.

Mohsen et. al (2008) used Simphony.NET to analyze and examine the on-site assembly aspect of the modular construction process used to construct 5 new dormitory buildings for college students. In the study, building modules were considered as the model entities that were processed by different resources. Construction sites often have particular site restrictions that effect production. The results of their simulation study agreed with actual onsite performance, with only one day variation in completion time.

Sharma et al. (2007), meanwhile, used simulation to quantify the impact of proposed changes on existing workflow and resource allocation as part of a broader initiative to study and optimize the service management processes in selected hospitals in Germany. They concluded that, by applying lean principles, targeted work orders could be completed efficiently using fewer resources. In another field, Bruzzone and Longo (2013) used simulation-based approach to support decision making in the food industry. They developed an advanced java-based simulation model using the AnyLogic software to assist production managers with the investigation of various production scenarios in a hazelnut processing facility. The simulation model was an updated version of a simpler model proposed by Longo et. al (2012), which was based on (i) standard library objects provided by AnyLogic software, without any ad-hoc java coding, and (ii) the assumption of infinite availability of raw materials.

Simulation models are traditionally built based on assumptions and approximations of input data. Chung et al. (2006) applied Bayesian updating techniques to improve the quality of inputs to a simulation model for an actual tunneling project, the North Edmonton Sanitary Trunk, in Canada. They demonstrated how statistical distributions of input parameters can be updated using Bayesian techniques, based on actual data as the project progresses, and that these updates can mitigate simulation uncertainty and enhance prediction accuracy, thereby improving overall project control over schedule and cost. Hu and Mohamed (2011) used simulation experiments with embedded Artificial Intelligence (AI) planning tools to identify the optimal fabrication sequence of pipe spools. Experimenting with different AI planners, their results demonstrated that these tools require extensive preparation of input data, which greatly affected results. They concluded that shop productivity could be improved by changing the spool fabrication sequence. Jahangirian et. al (2010) conducted an extensive literature review of simulation application within the manufacturing and business sectors between 1997 and 2006. Their study was characterized by wide coverage, broad scope of simulation techniques, and a focus on real-world applications.

Applying mass customization principles, together with underlying lean techniques, can benefit offsite construction companies. In this regard, offsite construction enterprises producing wall panels, complete modules, or specific portions of buildings (e.g., cabinets, doors, windows, etc.) can greatly improve their operations by developing simulation models that can be used for investigating the application of mass customization, lean principles, and knowledge discovery in data. Significant contributions to both academia and industry can result from the application of data mining techniques and mass customization philosophy to production systems using simulation and various data acquisition tools. It is anticipated that the present research will thus offer numerous benefits to the cabinet manufacturing industry at large, as the current research contributions in this area are limited.

3. MODEL DEVELOPMENT

In this section, we describe how the simulation experiment was carried out by outlining the study methodology, the underlying assumptions used to build the simulation model, and our interpretation of the results.

3.1. Methodology

Development of the model began with the research team observing the daily operations at the facility to identify relevant boundaries of the system and abstract its processes. The problem domain was restricted to the production lines of regular and custom kitchen cabinets. A workflow diagram and flow chart were developed as shown in Figure 1. The pertaining assumptions and requirements, upon which the conceptual model was built, were recorded and are documented in Section 3.2. A discrete-event simulation (DES) model was created using Simphony.NET, a simulation modelling tool developed at the University of Alberta. Experimentations with the model through what-if analysis, as well as verification and validation, were performed throughout development until satisfactory results were obtained. Results and a summary of the validation approaches used are presented in Sections 5 and 6, respectively.

3.2. Assumptions

In the model, the basic entity simulated is a kitchen cabinet purchase order. Each order is associated with one project, and the time unit employed is days. The database used in the simulation model is created using information provided by the manufacturer. The main characteristics of the orders used in the simulation are the number of cabinets and the type of material used in cabinets. In this context, we refer to an unfinished product as a “box”; a finished product, which has the box assembled together with its drawers and doors, is referred to as a “cabinet.” Each project consists of several cabinets that make up the kitchen, and the simulation is based on an average order size of 22 cabinets. Each cabinet consists of three main components:

- The cabinet box, which is the skeleton of the cabinet;
- The cabinet doors, which may be supplied or manufactured in-house; and
- The cabinet and kitchen components that are referred to, here, as accessories.

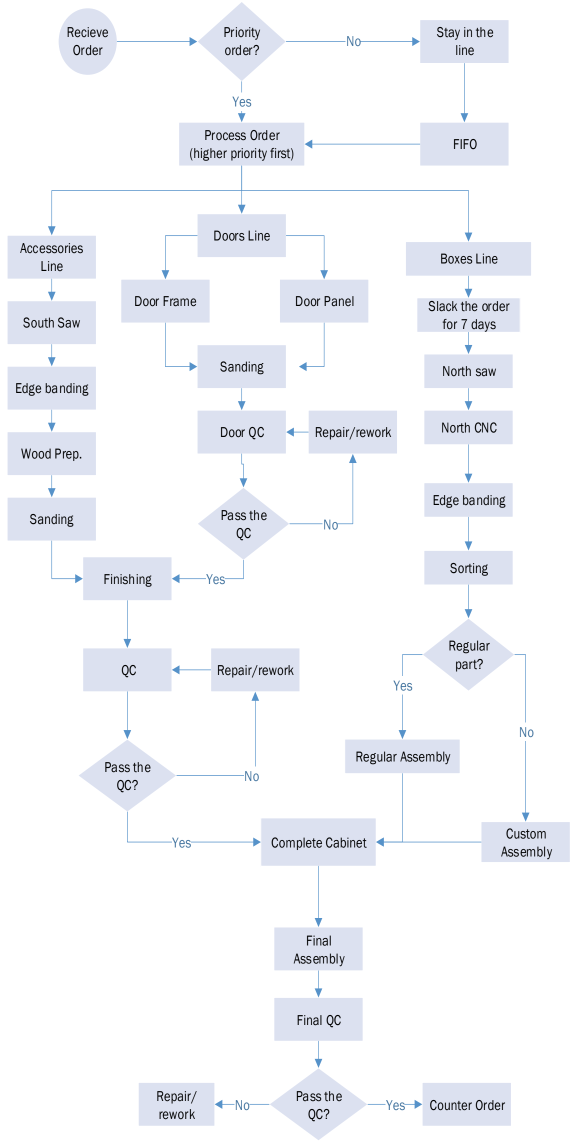


Figure 1: Flowchart of processes at the cabinet manufacturing facility

Each of these components is routed to its corresponding department once the raw material has been cut using one of two available saws. The following is a summary of the assumptions used to build the simulation model:

- Approximately 80% of the cabinets are regular, non-custom types, and 20% are custom cabinets. Custom cabinets require more time to assemble.
- Availability of raw materials is not a constraint, meaning that orders are infinitely flowing into the model as they occur with no shortage of supply. In a simulation model, this entails creating a large number of entities at the beginning of the simulation.
- Task durations are based on the historical data available. When such data is not available, the

durations are fit into triangular distributions based on observations.

- Approximately 10% of products passing through the quality control station will require rework.

4. MODEL LAYOUT AND DESCRIPTION

4.1. Facility Layout and Process Description

The manufacturer operates a cabinet production line that produces cabinets and transports them for in-home assembly. The cabinet manufacturing process consists of three lines: accessories, doors, and boxes.

The accessories line process begins with the cutting task, which is performed by the south saw. The saw can simultaneously cut an average of four panels at a time. Accordingly, panels are batched into average groups of four and cut into pieces at once by the south saw. After cutting, pieces are routed to the edge banding station, where pieces are edge-banded one-at-a-time. Since each element requires a different type of edge banding, setup time is allocated for this task and is incorporated into the duration of the task. Melamine material type, it should be noted, does not require sanding and painting. If the material type is veneer or hardwood, however, pieces are routed for wood preparation and sanding.

After sanding, the prepared hardwood accessories wait for their corresponding door parts, which are being processed in the door manufacturing line. Because there is no need for preparing or sanding melamine type, such pieces skip wood preparation and sanding and proceed directly to final assembly. Melamine account for approximately 20% of the company's material inventory.

In the doors manufacturing line, however, veneer and melamine account for almost 40% of the material, while hardwood accounts for 60%. If the material type is veneer or melamine, the panels proceed to the accessories production line to be cut and edge-banded, as per the aforementioned reason, before finally proceeding to the final assembly line. For hardwood types, the panels go through both doorframe and door panel processes before being routed to the finishing station and, then, to the final assembly line.

The doorframe process includes outer and inner edge profiling, while the door panel process includes cutting, gluing, clamping, raised panel cutting, and thickness sanding for raised panels. Then, the panels from the doorframe and door panel lines converge at the sanding stations, which include automatic and manual sanding stations. A quality control station immediately downstream from the sanding stations flags approximately 10% of the pieces as failing to pass the quality check, with the sanding operation prone to account for the majority of the defects identified. The hardwood doors (~60%) merge with their corresponding pieces from the accessories line and proceed to the finishing area, where they are painted. The second quality control station is located after the finishing station. Approximately 10% of the pieces are returned to the finishing station after failing to pass the quality control check, while the remaining pieces wait for their corresponding pieces from the box production line to arrive.

Table 1: Process Activities and Durations

Operation Station	Activity	Duration (minutes)
North Saw	Cutting	Triangular (1.25,1.75,1.5)
South Saw	Cutting	Triangular (0.75,1.25,1)
Accessories Edge banding	Edge banding	Triangular (0.75,1.25,1)
Edge-banding of boxes	CNC router/drilling	Triangular (1.75,2.25,2)
	Edge banding	
Sorting	Staging and sorting	Triangular (0, 0.75, 0.5)
Final Assembly	Drawer storage	Triangular (0.75,1.25,1)
	Door staging	
	Press door hinges	
Regular Assembly	Sort	Triangular (2.75,3.25,3)
	Case clamp	
Custom Assembly	Assembly station	Triangular (9,11,10)
	Case clamp	
QC	Quality check	Triangular (0.75,1.25,1)
Door Sanding	Sanding machine	Triangular (0.75,1.25,1)
	Manual sanding	
Accessories Sanding	Sanding machine	Triangular (1.25,1.75,1.5)
	Manual sanding	
Finishing	Finishing kitchen parts and cabinet parts	Triangular (9,11,10)
Door Frame	Sized lumber for door frames	Triangular (0.75,1.25,1)
	Cutting and profiling for door frame	
Door Panel	Flat panel saw	Triangular (5.25,6.75,6)
	Cutting lumber to width	
	Cutting lumber to thickness	
	Sized lumber for raised panel	
	Gluing station	
	Clamping machine for raised panels	
	Raised panel saw	
WP	Wood preparation	Triangular (0.75,1.25,1)

Seven days after an order arrives at the factory, cabinet box production begins at the box production line. The first station in the box manufacturing line is the north saw, by which the panels are cut in maximum groups of six. At the next station, pieces are drilled with a CNC machine and are edge banded. Similar to the edge banding machine in the accessories line, pieces are edge-banded one-at-a-time.

The pieces are then sorted as either regular or custom and are routed to their respective partial assembly lines. The partial assembly line includes assembly of boxes and clamping. The last step is the final assembly line, in which all parts of an order converge and are assembled together. Figure 1 shows the conceptual workflow at the cabinet manufacturing facility. Table 1 lists the activities and durations of each activity, with the distribution of the activities approximated using triangular distribution.

4.2. Simulation Model

A detailed simulation model of the cabinet manufacturing production line (shown in Figure 2) is developed in Simphony.NET. In the simulation model, cabinets are the model entities. All cabinet attributes, such as priority, number of cabinets, number of custom cabinets, and type of wood, are extracted from the database at the beginning of the simulation run. Each entity goes through a different task element based on its attributes. A task element represents a work package or a group of work packages in a workstation such as door panel, door frame, or finishing. Each piece of machinery (e.g., north saw, CNC machine, sanding machine, edge-banding machine) is defined as a resource in the simulation model.

Once the entities have been created using the database, they go through the "Generate" element so that cabinets can be created based on the attributes of each order. The conditional branch separates the different types of materials, routing non-hardwood materials to the accessories line for further processing. Given that the south saw in the accessories line can cut several panels simultaneously, an average of four panels are batched together at cut at the same time. Since some orders have a higher priority than others, the captured resource elements with higher priority are used for each task. This results in the work on the current order being stopped and switched to those with higher priority.

5. SIMULATION RESULTS

Based on real data, the company produces an average of 220 cabinets per day when the system proceeds without any interruptions or unexpected breakdowns. To maintain the company's confidentiality, data are masked by being multiplied by an arbitrary number.

The results derived from the simulation indicate that an average of 44,935 minutes or 94 days (assuming a work shift of 8 hours/day) is required to produce 17,314 cabinets. The rationale for choosing a large number of entities is to generate valid results by ensuring that all stations are fully loaded. As there will not be any cabinets produced until Day 14 (due to the lead time of placing the first order until it is promised to be ready), fourteen days are subtracted from the final number of days shown in the results. The simulation results estimate a production rate of 216 boxes per day, which is close to the rate that was obtained using actual performance. The simulation is run for 30 times to analyze the stochastic nature of activity durations. The average daily production rate has not changed and remains at 216 with minimal variance. It is noted that the north CNC and edge-banding stations as well as the regular assembly line are the bottlenecks of the whole process due to over utilization at 71% and 93%, respectively. Several scenarios have been examined to mitigate this issue. One solution we propose is to double the number of resources at these stations.

Implementing the proposed changes will increase the average production rate to 333 cabinets per day and resolve the bottlenecks identified in the model; however, the custom assembly line become the bottleneck at 89%. The proposed scenario demonstrates that, by increasing some resources, the production rate increases, and the processing time decreases.

On another note, cabinets associated with each order are either of regular or custom type. Assembly of custom cabinets takes three times longer than regular cabinets to complete. To examine the effect of cabinet-type mix on the production rate, different percentages of regular versus custom cabinets were assigned for each order. Figure 3 demonstrates that, to reach the maximum simulated production rate of 216 cabinets per day, total orders should consist of 75% regular cabinets and 25% custom cabinets.

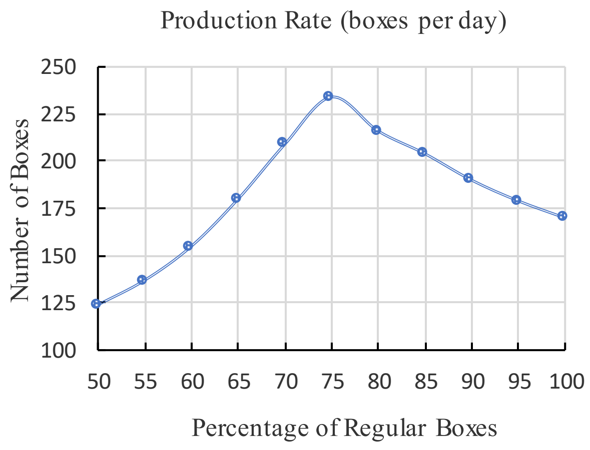


Figure 2: Daily Production of Cabinet Mix

6. MODEL VALIDATION

Several model verification and validation approaches are used to ascertain the accuracy and reliability of the proposed simulation model.

Conceptual model validations are used to evaluate the system's logic. The input, process, and output of each production line, as well as the assumptions described above, were analyzed and confirmed by subject-matter experts.

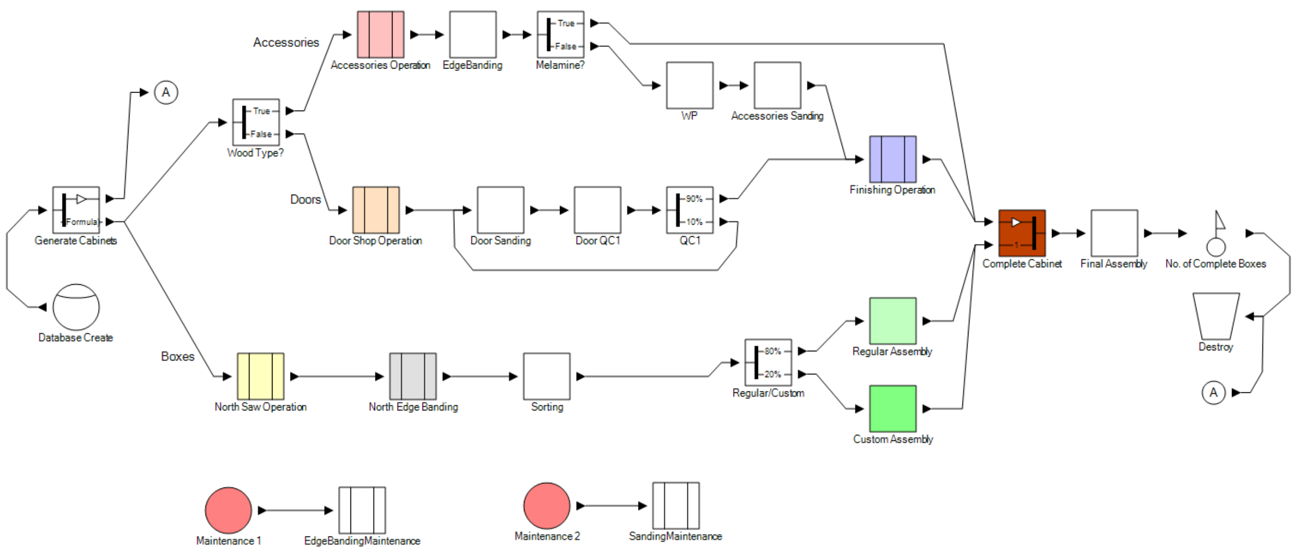


Figure 3: Simulation model of the cabinet manufacturing facility in Simphony.NET

In addition, the results of the computerized model were compared to the historical data of the company, and sensitivity analysis was used to determine the responsiveness of the model to various scenarios. As evidenced in Figure 3, the maximum production rate of 234 cabinets per day occurs when regular boxes represent 75% of total cabinets produced.
In this study, the simulation model is validated at two levels:

(i) Duration of each order, and
(ii) Number of boxes.

On the first level, the duration of each line—doors, accessories, and boxes—is analyzed. The average simulation cycle time for activities in the door line, the accessories line, and the box line were 11.5, 11, and 5.5 days, respectively. These are very close to the actual durations of each line, which are determined, based on historical data, to be 11, 10, and 6 days, respectively, as shown in Table 2. These difference between actual and simulated times are considered acceptable. For the next level of validation (i.e., the number of boxes) historical data are compared to simulated results; results are found to be very close as well, thereby validating the simulation model. The average daily production rate derived from the simulation is 216 cabinets, which deviated from actual measurements by 1.8%. A few possible explanations could account for the deviation: for example, breakdown activities or different types of rework (and rework duration) are not fully incorporated in the proposed simulation model.

7. CONCLUSION

In this paper, a simulation study of the production line in a kitchen cabinet manufacturing company is conducted. A simulation model illustrating the current production process of the company is developed using the discrete event simulation approach.

Table 2. Comparison of actual and simulated results

	Actual	Simulated	Diff.
Cycle time for doors (days)	11	11.5	4.5 %
Cycle time for accessories (days)	10	11	10%
Cycle time for boxes (days)	6	5.5	8.3%
Daily production rate (cabinets/day)	220	216	1.8%

The proposed simulation model is found to be capable of producing reliable results that are representative of the actual system. Moreover, the simulation model uncovered potential scenarios that are anticipated to improve the production process, reduce wait times, and increase productivity. The model is validated by comparing the total order completion time, as well as the average number of boxes produced per day, that are obtained by the simulation model with historical data. Generated results are close to historical data. One of the most important finding of the simulation result is the recognition of bottlenecks. Based on the utilization report, the regular assembly line is determined to be the bottleneck of the process. Several what-if scenarios are examined to determine what improvements are applicable. Adding additional resources to the regular and assembly lines, as well as additional resources to the north CNC and edge-banding stations, are anticipated to increase the average production rate from 216 to 333 cabinets per day, and improvement of more than 50%. This study demonstrates the advantage of utilizing simulation models to experiment with real-world systems and of using simulation to investigate the impact of various what-if scenarios before deploying conceptual solutions into practice.

ACKNOWLEDGMENTS

This research was funded through an Engage Grant by the Natural Science and Engineering Research Council of Canada (NSERC). Special thanks to the industry collaborator for their support throughout this study.

REFERENCES

AbouRizk, S. (2010). "Role of Simulation in Construction Engineering and Management." *Journal of Construction Engineering and Management*, 136(10), pp. 1140-1153. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000220

Altaf M., Bouferguene A., Liu H., Al-Hussein M., Yu H., (2017). "Integrated production planning and control system for a panelized home prefabrication facility using simulation and RFID." *Automation in Construction*, 85 (2018), pp. 369-383. https://doi.org/10.1016/j.autcon.2017.09.009

Altaf M., Liu H., Al-Hussein M., (2015). "Online Simulation Modeling of Prefabricated Wall Panel Production Using RFID System." *Proceedings of the 2015 Winter Simulation Conference*, pp. 3379-3390. December 6-9, Huntington Beach, CA, USA. https://doi.org/10.1109/WSC.2015.7408499

Altaf M., Liu H., Al-Hussein M., AbouRizk S., (2014). "Resource Planning in Panelized Construction based on Discrete-Event Simulation." *Proceedings of the 14th International Conference on Construction Applications of Virtual Reality*, pp. 26-33. November 16-18, Sharjah, UAE.

Benros D., Duarte J.P., (2009). "An integrated system for providing mass customized housing." *Automation in Construction*, 18(2009), pp. 310-320. https://doi.org/10.1016/j.autcon.2008.09.006

Bruzzone, A. & Longo, F. (2013). "An Advanced Modeling & Simulation Tool for Investigating the Behavior of a Manufacturing System in the Hazelnuts Industry Sector." *International Journal of Food Engineering*, 9(3), pp. 241-257. https://doi.org/10.1515/ijfe-2013-0039

Callahan, R., Hubbard, K. & Bacoski, N. (2006). "The use of simulation modeling and factorial analysis as a method for process flow improvement" *International Journal of Advanced Manufacturing Technology* 29(1-2), pp. 202-208. https://doi.org/10.1007/s00170-004-2497-5

Chung T., Mohamed Y., AbouRizk S., (2006). "Bayesian Updating Application into Simulation in the North Edmonton Sanitary Trunk Tunnel Project." *Journal of Construction Engineering and Management*, 132 (8), pp. 882-894. https://doi.org/10.1061/(ASCE)0733-9364(2006)132:8(882)

Hu D., Mohamed Y., (2012). "Pipe spool fabrication sequencing by automated planning." *Proceedings of Construction Research Congress 2012 Conference*, pp. 495-504. May 21-23, West Lafayette, Indiana, USA. https://doi.org/10.1061/9780784412329.050

Jahangirian, M., Eldabi, T., Naseer, A., Stergioulas, L. and Young, T. (2010). "Simulation in manufacturing and business: A review." *European Journal of Operational Research*, 203(1), pp. 1-13. https://doi.org/10.1016/j.ejor.2009.06.004

Kamaruddin S., Khoo S.Y., Khan Z.A., Siddiquee A.N., (2011). "The effect of layout design on productivity: an empirical study." *International Journal of Productivity and Quality Management*, 7 (4), pp.484–500. https://doi.org/10.1504/IJPQM.2011.040545

Koskela, L. (1992). "Application of the New Production Philosophy to Construction." Technical Rep. No. 72, Center for Integrated Facility Eng, Dept. of Civil Engineering, Stanford Univ., Stanford, Calif.

Kotha S., (1996). "From Mass Production to Mass Customization: The Case of the National Industrial Bicycle Company of Japan." *European Management Journal*, 14 (5), pp. 442-450. https://doi.org/10.1016/0263-2373(96)00037-0

Longo, F., Massei, M., Nicoletti, L. (2012). "An application of modeling and simulation to support industrial plants design" International Journal of *Modeling, Simulation, and Scientific Computing*, 3 (1), art. no. 1240001. https://doi.org/10.1142/S1793962312400016

Mohsen O.M., Knytl P.J., Abdulaal B., Olearczyk J., and Al-Hussein, M. (2008). "Simulation of modular building construction" *Proceedings, 2008 Winter Simulation Conference*, Austin, TX, pp. 2471-2478. https://doi.org/10.1109/WSC.2008.4736356

Phuluwa S., Mpofu K., Ramatsetse B., Tshawe C., (2013). "Role of Reconfigurable Manufacturing Systems in Minimizing Import Dependency in the South African Press Tool Enterprise." *Proceedings of the 5th International Conference on Changeable, Agile, Reconfigurable and Virtual Production*, pp. 97-103. October 6-9, Munich, Germany. https://doi.org/10.1007/978-3-319-02054-9_17

Pine, B. Joseph, (1993). Mass Customization: The New Frontier in Business Competition. Harvard Business School Press. ISBN 0-87584-372-7.

Sharma V., Abel J., Al-Hussein M., Lennerts K., Pfrunder U., (2007). "Simulation application for resource allocation in facility management processes in hospitals." *Facilities*, 25 (13/14), pp. 493- 506. https://doi.org/10.1108/02632770710822599

SIMULATION-BASED PLANNING AND OPTIMIZATION OF AN AUTOMATED LAUNDRY WAREHOUSE USING A GENETIC ALGORITHM

Marcel Müller[(a)], Tobias Reggelin[(b)], Stephan Schmidt[(c)]

[(a),(b),(c)] Otto von Guericke University Magdeburg, Germany

[(a)]marcel1.mueller@ovgu.de, [(b)]tobias.reggelin@ovgu.de, [(c)]stephan.schmidt@ovgu.de

ABSTRACT
The planning of logistics systems is a complex task with important decisions to make. Simulation models can help already in the early planning process of these systems. Usually they only provide a visualization of the different planned concepts but with modern genetic algorithm it is possible to provide even more support. Numerous parameters are not yet fixed at this point of a planning process. A dynamic model with flexible parameters and a genetic algorithm (GA) can already deliver good approximated solutions. An exemplary procedure for using this GA in the context of the research and development project "LOCSys" (Laundry Order Consolidation System) is presented in this paper. We use the genetic algorithm to find good approximated parameters for an optimised throughtput by changing the dimensions and control parameters of a warehouse with an automated picking and retrieval unit.

Keywords: genetic algorithms, simulation modelling, optimization, automation

1. INTRODUCTION

The complexity of planning logistics systems has been steadily increasing for years. One of the reasons for this is the increasing automation, digitization and networking of logistics objects. A secure planning process as the basis for making the right decisions is of immense importance to the logistics planner. The creation of simulation models of logistical systems is meanwhile part of the classical planning. The software for creating simulation models has evolved in recent years and now offers comprehensive integrated optimization tools. This not only makes it possible to visualize and simulate the logistics system, but also to optimize it. Thanks to modern algorithms, well-developed interfaces for data import and flexible model creation, the model creation and optimization is partly already automated. Initial application examples in the automotive industry already show the effectiveness of such combinations in the planning (Kirchhof and Stoehr 2016).

Our approach is used as an example in the picking area of an industrial laundry. Brandau, Weigert and Tojulew (2015) showed various application of simulation in industrial laundries already and pointed out the difficulties such as "insufficient basic data, increasingly complex customer and product structures, and a lack of transparency", which are typical for an industrial laundry. Generally, compared to the automotive business, industrial laundries mean smaller dimensions and supposed lower complexity because the good is simpler. The microscopic modelling effort, however, can be higher in the picking area, since there are no fixed storage locations/dimensions and the goods are restricted to certain stacking sequences. The planned logistical system for solving the picking process includes the designing of a special storage and retrieval unit (SRU). Figure 1 shows an early concept of the planned picking system "Laundry Order Consolidation System (LOCSys)" which aims to improve the picking & storing processes in the clean area of an industrial laundry through automation.

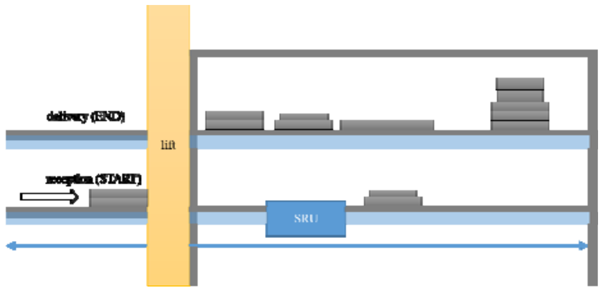


Figure 1: Early Concept of LOCSys

Typical sorter and batcher systems do not work for laundry items because the items are pliable (form unstable) and additional outer packaging is usually not allowed. The approach of LOCSys is therefore a new picking system with a storage and retrieval unit (SRU) which reliably handles laundry items with a special gripping mechanism. The intended warehouse system with dynamic storage allocation and dynamic storage dimensions will address the issue of low available space in industrial laundries. The stacking-order-problem of laundry items is similar to the mathematical puzzle "Tower of Hanoi" and will be addressed with an algorithm which is similar to the ones that solve the Tower of Hanoi problem (Hinz, et al. 2013).

LOCSys in its current state consists only of one rack row and the SRU. The parameters of the rack row are flexible and have to be determined depending of the input material flow. To model a flexible rack row, there is the need to model numerous individual models for each parameter combination or an automatic model creation. We opted for the latter.

2. AUTOMATED GENERATED SIMULATION MODEL

2.1. Challenges and Current State of the Model

In the early planning process many parameters and characteristics of the planned logistics system are not yet determined. A simulation model must therefore be highly flexible at this point in the planning process in order to be able to react to major changes in the model structure. The automated creation of the model structure based on predefined parameters, rules or templates is one way to solve this challenge.

The approach is not new and a lot of research has already been done on this topic. Researchers often point to the great potential of automatic modelling, despite the challenges of applying this approach:

Automated generated simulation models have the "[…] potential to radically reduce the time, cost, and skills of creating complex models […]" (Lucko, et al. 2010).

According to Bergmann and Strassburger (2010), there are four main challenges for automatic model generation:

- Incomplete data in external systems
- Generation of dynamic / complex behaviour
- Support of cyclic approaches involving multiple model generations cycles
- Support of multiple life cycle phases of the production system

Bergmann and Strassburger (2010) try to counter these challenges with mechanisms that "[…] support automatic testing and adaptability".

In our case of application, we decided only to partially automate the model creation because some model building blocks do not change in different scenarios. We use a discrete event and microscopic simulation approach to model LOCSys because discrete rate and mesoscopic simulation models (Tolujew and Reggelin 2008) would aggregate the objects properties too much.

Figure 2 shows the current state of the simulation model of LOCSys in the Siemens PLM Software "Tecnomatix Plant Simulation". The yellow marked area consists of the manually created objects and the automatically generated are in the green area. Adjustment parameters are currently in the blue area but in perspective they will be transferred into a more appealing user interface.

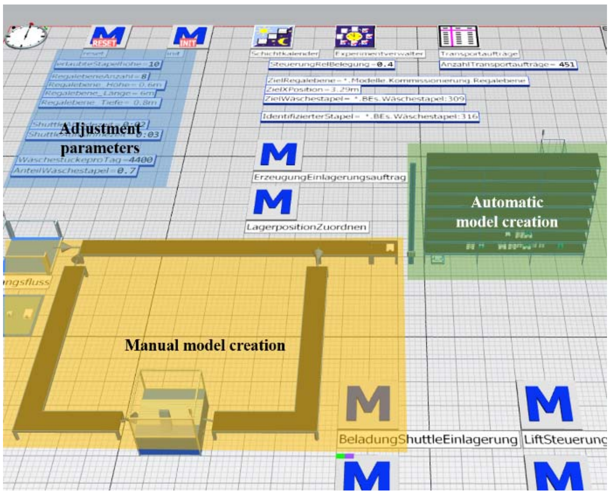


Figure 2: Current Simulation Model of LOCSys

Adjustment parameters include the shelf length and the number of shelf levels. It is assumed that the SRU can only accept individual or entire laundry stacks. This assumption is due to the difficult handling of laundry and the restriction that no further outer packaging may be used. A new developed extern picking algorithm will take over the control of the SRU in the future. Until then, there is a small test control module in the simulation model, which initially tries to store the laundry stack on the lowest shelf level. But as soon as a certain relative occupancy of the shelf level is reached, the next higher shelf level is used. This relative occupancy is also an adjustment parameter related to the dimensioning of the rack row. The adjustment and other parameters are stored as global variables in the simulation model.

The automatic model creation is realized by the method "INIT", which is triggered every time the simulation model is initialized. The following pseudo-code shows a part of the logic inside the method:

```
RackRow.Length := ParRackRoW.Length
RackRow.Depth := ParRackRow.Depth
ShelfLevel.BaseHeigth :=
ParShelfLevel.BaseHeigth
LiftRail.Length := ShelfLevel.BaseHeigth *
ParRackRow.NumberLayers + Lift.Length / 2
ShuttleRail.Length := RackRow.Length
ShuttleRailReception.BaseHeigth :=
RackRow.BaseHeigth
ShuttleRailDelivery.BaseHeigth :=
RackRow.BaseHeigth * 2

-- Create Shelves
for i := 0 to ParRackRow.NumberLevels - 1
    .ShelfLevel.createObject (.Models.LOCSys,
    400, 320)
    if  i = 0
        ObjectName := "ShelfLevel"
    else
        ObjectName := to_str("ShelfLevel", i)
    end

    Object := str_to_obj(ObjectName)
    Object.coordinate3D[3] :=
    Object.coordinate3D[3] +
    .ShelfLevel.BaseHeigth * i
next
```

In the first part, the input parameters, which variable names begin with "Par", can easily set the object attributes. Simple calculations, such as for calculating the length of the elevator, are calculated based on the entered parameters. For the generation in the 3D area, however, a for-loop is necessary because the object creation can only be addressed by the x and y coordinates of the area, i.e. without the z ones. The problem arises that we have to refer to objects that do not exist at the beginning of the initialization of the simulation model to change their positions in the model. This is achieved by converting the generated string of the object name to the corresponding object. This is possible, because it is known how the software names the automatically generated objects: "ShelfLevel", "ShelfLevel1", "ShelfLevel2" and so on. This method to address objects that do not exist in the beginning is also heavily used by other control modules in the simulation model.

2.2. Verification and Validation

Since the planned picking solution is still in an early planning phase, the simulation model is also highly flexible. For this reason, first of all, verification processes take place, which internally ensure that the simulation model has been correctly transformed from the planning sketches into the simulation software.
One form of verification was the 3D animation of the SRU and the laundry stacks. This allowed the modeller to verify that the route finding / procedure was correct and whether the stacking order and stacking rules (smaller dimensions only on same/larger dimensions) were complied with. Another verification process took place in regular project meetings. Through a review process with the experts who plan the warehouse system, it was ensured that the modelled picking solution also corresponded to the planned, real picking solution.
Further verifications will continue in the future with the project partners in order to reduce iteratively and preventively planning costs and to avoid wrong planning decisions. The validation will take place later in the project with appropriate representatives from industrial laundries.

3. GENETIC ALGORITHMS

Genetic algorithms (GA) have been discussed intensively in science since the 1980s. Goldberg and Holland gave genetic algorithms a greater reputation, when they defined genetic algorithms as "probabilistic search procedures designed to work on large spaces involving states that can be represented by strings. " (Goldberg and Holland 1988). Genetic algorithms are part of evolutionary algorithms and work with natural mutations and recombination to create new individuals. The best individuals in relation to the fitness function are taken over into the next generation. After that they are adjusted and mutated again to find even better solutions. Figure 3 shows the fitness of the offspring of an example in a diagram. In this case, we made five observations per individual and created five generation with 30 individuals (1st generation only 15), which are divided by the red line in the diagram. The improvements from generation to generation are not big but visible. The second generation has for example three individuals with a fitness value below 400, while the 5th generation no longer consists of fitness values below 400. The penalties show individuals where the combination of the parameters created an error in the simulation model. This can also be used to validate the simulation model if the errors are not supposed to happen owing to impossible parameter combinations.

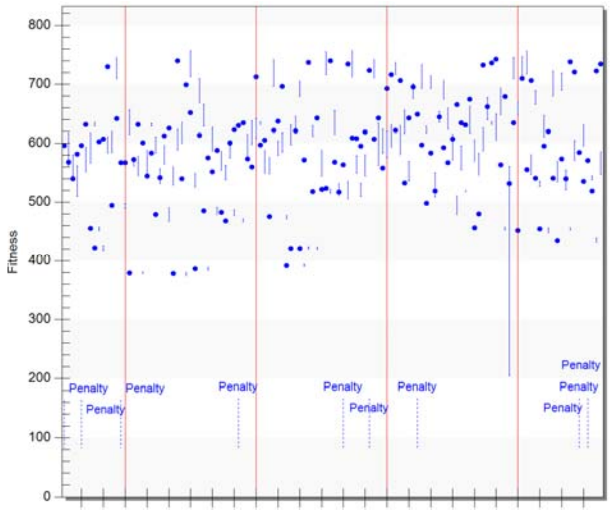


Figure 3: Fitness of the offspring

Genetic algorithms are nowadays widely used and common in the optimization of logistics systems. Not only as the already given example in the automotive industry but they are also applicable to the steel industry (Paul and Chanev 1997), manufacturing systems (Can, Beham and Heavey 2008), closed looped supply chain battery recycling models (Ramkumar, et al. 2010) to give just a few examples of countless more. Researchers are also trying to optimize genetic algorithms furthermore. For example, Deep, Shashi and Katiyar (2011) developed a new mutation operator that "prevents premature convergence of an algorithm". Many modern simulation software offer nowadays integrated GA assistants which are easy to adjust. We also used such an assistant to optimize the parameters of the rack row in our simulation model.

4. EXPERIMENTS AND RESULTS

4.1. Optimization Goals

LOCSys is an automated order picking system, which should generate the highest possible throughput of goods at the lowest possible costs. As not all cost figures are known at the moment, we only expect a maximization of the throughput. The throughput of laundry items is measured at the sink of the system with the number of moving elements leaving the system.

With just one target figure, weighting the goals was unnecessary.

4.2. Optimization Parameters

The chosen optimization parameters can be divided into three categories:

- Warehouse dimensions
- Controls
- Material flow quantities

The parameters were determined based on physical constraints that exist in our use case or realistic estimates based on experience and understanding of the system. We discussed already the flexible parameters of the rack row. The genetic algorithm changes in this category only the rack row length and the number of shelf levels. For the experiments we used the following values:

Table 1: Parameters for the Rack Row Length [m]

Parameter	Value
Lower limit	3
Upper limit	20
Increment	0.25

Table 2: Parameters for the Number of Shelf Levels

Parameter	Value
Lower limit	1
Upper limit	9
Increment	1

Regarding the control, there are three types of processes: placing goods in storage, retrieving/picking them and relocating them. The SRU processes the orders one after the other according to the FIFO principle. The process “placing in storage” occurs when laundry items arrive and allocate a random free storage space in the shelf for the item. This provisional control prevents the algorithm from simply choosing the largest rack row dimensions, as there are no costs implemented yet. As a result, longer rows of shelves now also create longer paths and in consequence less throughput.

An artificial clocking triggers the methods for creating picking and relocating orders. The timing between the two orders are also adjustable parameters for the genetic algorithm:

Table 3: Parameters for the Clocking of the creation of Picking/Relocating orders [s]

Parameter	Value
Lower limit	5
Upper limit	90
Increment	5

Another control parameter is the relative occupancy of a shelf level, which determines also how much the lift or the horizontal movement will be used:

Table 4: Parameters for the Target Relative Occupancy of the shelf levels

Parameter	Value
Lower limit	0.1
Upper limit	0.9
Increment	0.1

We also decided to test a flexible input material flow to find the maximum throughput of the system, taking into account the previous mentioned parameters that can be changed:

Table 5: Parameters for the Interval of the Laundry Item Creation of the Source [s]

Parameter	Value
Lower limit	1
Upper limit	15
Increment	0.5

For the general settings, we have an initial population of 30 and a population 60 for each additional generation. Ten generations of individuals should be generated. In order to minimize stochastic fluctuations, ten simulation runs were performed for each individual with a simulation period of 18 hours (two shifts).

4.3. Results

As expected and according to the planning of the experiments 670 individuals were generated. Figure 4 shows the development of the fitness values with every new generation. The genetic algorithm already finds individuals with high fitness values in the first generations. The average fitness value increases strongly until the 5th generation. In the last generations there is nearly no visible improvement anymore, although two of the five best individuals only were found in the 10th generation.

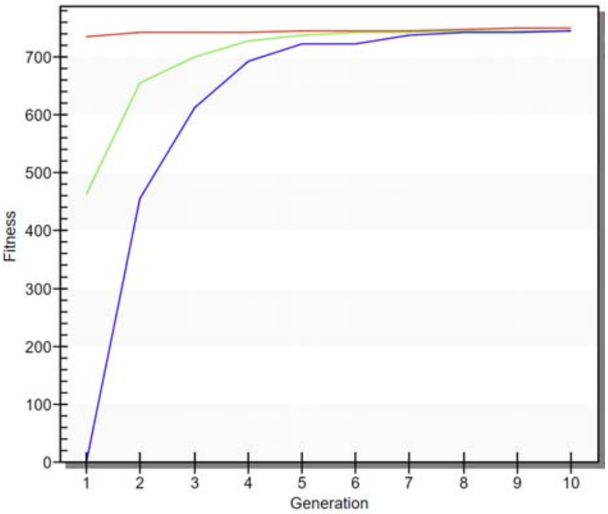


Figure 4: Development of the fitness values of the generations (Red = best solution, green = average of solutions, blue = worst solution)

The sharp increase in the first half of the generations' leads to the conclusion that a smaller number of generations would also have led to a very good solution for our specific application. In other cases, with additional optimization parameters, smaller increments and larger limits, more generations would probably be necessary.
The best individual with a fitness value of 750.3 resulted from the listed parameters in Table 6.

Table 6: Parameters for the Best Found Solution

Parameter	Value
Number of shelf levels	9
Rack row length [m]	7
Target relative occupancy	0.7
Clocking relocating [s]	15
Clocking picking [s]	15
Interval of the source [s]	6

This means that high rack rows of short length are good solutions on condition that the other parameters do not change in the future. The system could also handle more laundry items than expected

5. CONCLUSION AND OUTLOOK

This paper showed that simulation can not only visualize in an early planning process but also optimize first concepts. One way to do that is to use genetic algorithms.
The first approximate solutions can already be a good guidance for the logistics planners and decision makers to parameterize their system. There is no need for a high number of generations because there is no search for the exact optimal solution yet in this state of the planning.
In our particular scenario, the optimal solution tends to high rack rows of short length. This can of course change very quickly if the planned acceleration or maximum speed of the planned lift and SRU change during the conceptual phase.
A disadvantage of this approach is that the results of both the simulation and the genetic algorithm can't always be fully understood. In our case for example, we do not know yet which adjustment parameter has the biggest impact on the fitness value. A sensitivity analysis could help in this regard.
In the further course of our project we will adapt the parameters to fit our steadily changing picking solution. We are curious to see which solutions other parameters and the external picking algorithm will generate in the future and excited to share the results as well.

ACKNOWLEDGMENTS

As part of the "Zentrales Innovationsprogramm Mittelstand" (ZIM) of the Federal Ministry for Economic Affairs and Energy of Germany, a research and development project was launched to find an innovative solution for automating the picking area in industrial laundries. The project is called "LOCSys - Laundry Order Consolidation System". We are very thankful for the financial support of the ministry which allows us to research for new picking solutions in the laundry industry.

REFERENCES

Bergmann, Sören, and Steffen Strassburger. "Challenges for the Automatic Generation of Simulation Models for Production." *Proceedings of the 2010 Summer Computer Simulation Conference.* Ottawa, Ontario, Canada: Society for Computer Simulation International, 2010. 545-549.

Brandau, Annegret, David Weigert, and Juri Tojulew. "Applications of Simulation to Improve Operations." In *Simulation in Production and Logistics*, by Markus Rabe and Uwe Clausen, 289-298. Stuttgart: Fraunhofer Verlag, 2015.

Can, Birkan, Andreas Beham, and Cathal Heavey. "A comparative study of genetic algorithm components in simulation-based optimisation." *Proceedings of the 2008 Winter Simulation Conference.* Miami: IEEE, 2008. 1829-1837.

Deep, Kusum, Shashi, and V.K. Katiyar. "A New Real Coded Genetic Algorithm Operator: Log Logistic Mutation." *Proceedings of the International Conference on Soft Computing for Problem Solving (SocProS 2011) December 20-22, 2011.* India: Springer, India, 2011. 193-200.

Goldberg, David E., and John H. Holland. "Genetic Algorithms and Machine Learning." *Machine Learning 3.* USA: Kluwer Academic Publishers-Plenum Publishers, 1988. 95-99.

Hinz, Andreas M., Sandi Klavžar, Milutinović Uroš, and Petr Ciril. "The Classical Tower of Hanoi." In *The Tower of Hanoi – Myths and Maths*, by Andreas M. Hinz, Sandi Klavžar, Milutinović Uroš and Petr Ciril, 71-130. Basel: Birkhäuser, 2013.

Kirchhof, Patrick, and Tobias Stoehr. "Optimization of Storage Allocation Using an Automatically Generated Warehouse Simulation Modell." *Proceedings of the 2016 Winter Simulation Conference*. Duesseldorf: Winter Simulation Conference, 2016. 3554-3555.

Lucko, Gunnar, Perakath C. Benjamin, Kannan Swaminathan, and Michael G. Madden. "Comparison of Manual and Automated Simulation Generation Approaches and their use for Construction Applications." *Proceedings of the 2010 Winter Simulation Conference*. Baltimore: Winter Simulation Conference, 2010. 3132-4144.

Paul, R.J., and T.S. Chanev. "Optimising a Complex Discrete Event Simulation Model Using a Genetic Algorithm." In *Neural Computing & Applications, Volume 6, Issue 4*, 229-237. London: Springer-Verlag London Limited, 1997.

Ramkumar, N., P. Subramanian, T.T. Narendran, and K. Ganesh. "A genetic algorithm approach for solving a closed loop supply chain model: A case of battery recycling." In *Applied Mathematical Modelling, Volume 35, Issue 12*, by M. Cross, C.J. Malmborg, N.C. Markatos and J.I. Ramos, 5921-5932. Elsevier Inc., 2010.

Tolujew, Juri, and Tobias Reggelin. "Mesoscopic Simulation: In Between of Continuous and Discrete-event Simulation." In *Advances in Simulation for Production and Logistics Applications*, by Markus Rabe, 585-594. Stuttgart: Fraunhofer IRB Verlag, 2008.

AUTHORS BIOGRAPHY

MARCEL MÜLLER is a research fellow at Otto von Guericke University Magdeburg. He received a master degree in industrial engineering for logistics from Otto von Guericke University Magdeburg. His research interests include modelling and simulation of logistics systems and forecasts of logistics figures.

TOBIAS REGGELIN is a research and project manager at Otto von Guericke University Magdeburg and the Fraunhofer Institute for Factory Operation and Automation IFF. He received a doctoral degree in engineering from Otto von Guericke University Magdeburg. His research interests include modelling and simulation of production and logistics systems and the development and application of new modelling and simulation methodologies.

STEPHAN SCHMIDT is an assistant professor for "autonomous vehicles" at Otto von Guericke University Magdeburg. He received his doctoral degree in Mechatronics in 2013. His research interests include modelling and simulation of mechatronic systems as well as planning and control strategies for automated vehicles.

APPROACH TO EVALUATING AND PLANNING INDUSTRIAL LAUNDRIES BY USING DISCRETE EVENT SIMULATION AND PERFORMANCE MEASUREMENT SYSTEM

David Weigert[(a)], Frank Ryll[(b)], Marcel Müller[(c)]

[(a), (b)] Fraunhofer Institute for Factory Operation and Automation IFF, Magdeburg (Germany)
[(c)] Otto von Guericke University Magdeburg, Magdeburg (Germany)

[(a)]david.weigert@iff.fraunhofer.de, [(b)]frank.ryll@iff.fraunhofer.de,
[(c)] marcel1.mueller@ovgu.de

ABSTRACT

Insufficient basic data, increasingly complex customer and product structures and a lack of transparency in the process structure increase the amount of laundry items in circulation, reduce machine utilization, delay deliveries and increase the error frequency. Competition is making it essential for laundries to operate quickly, reliably and cost effectively. The often non-transparent internal processes further worsen the economic situation of the companies. The evaluation of laundries based on logistic key figures has been solved individually and cumbersome until today. The reason is the partial lack of ability to determine the logistic performance indicator. The article presents on the one hand the implementation of a simulation study for the determination of the order of orders as well as the increase of the transparency by the development of a logistic performance measurement system using logistic indicators and characteristics.

Keywords: KPI Logistics, Industrial Laundry, Logistic Operating Curves Theory

1. INTRODUCTION AND MOTIVATION

Industrial laundries in Germany provide comprehensive services for the deployment, cleaning and treatment of washable textiles. The laundry requirements increase with increasing competition and pronounced customer orientation. Individual customer in the health care and hospitality sector require a timely and flexible cleaning and delivery of laundry at a low prices. Due to the historical development of many laundries, their planning, control and monitoring are characterized by the experience of the employees. The closed process of a laundry differs from a manufacturing company in that the product, the laundry, is not made from existing resources, but runs through a laundry cycle. The core processes such as washing and processing (ironing, drying, folding, stacking) are highly automated processes, their interfaces often still manually run. Due to the increased complexity in the customer and product structure, laundries handle a diverse number of garments. To anticipate and optimally plan a production, it is necessary to know the type and number of existing resources in the goods receipt. This requirement is only partially fulfilled in the case of the laundries. The type and number of incoming dirty linen

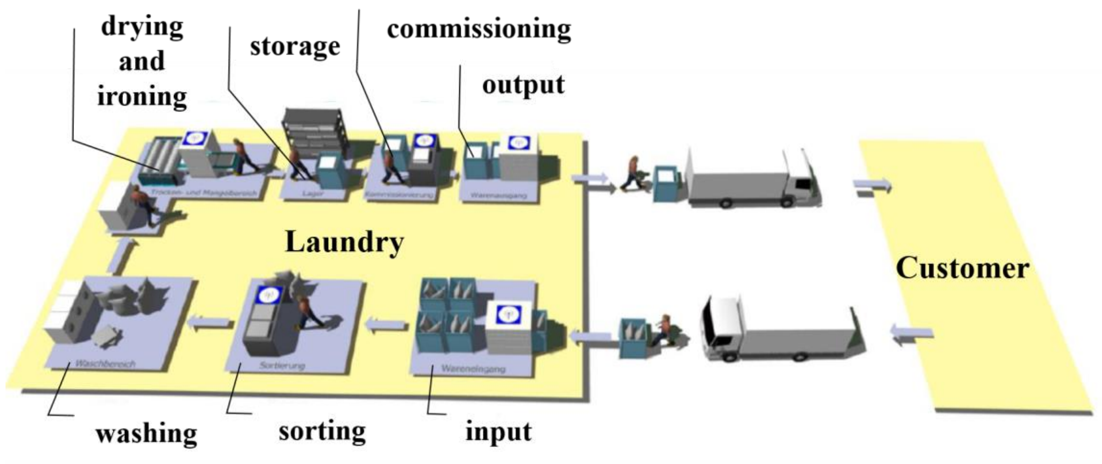


Figure 1: Schematic representation of the general laundry process

of customers is not known in advance. The content of the container is visible through the closed transportation difficult. Planning the washing sequence order is thus based on uncertain data, resulting in costly rescheduling in the operations. The reasons for this development are manifold and can be attributed to the growing demand for laundries, the increasing outsourcing of laundry services and the increasing demands on the purity of the laundry due to higher standards. However, in spite of these developments it has not been possible to evaluate or compare laundries by means of generalized key figures, so that the detection of improvement potentials is difficult and cumbersome. Still exists in small laundries little storage space, a new building is possible in rare cases. These reasons lead to high circulation levels of laundry, poor machine utilization, delayed deliveries and increased susceptibility to faults in the process sequence. Due to the current market and competition situation, it will be of great importance for wholesalers to work quickly, reliably and at low cost. This article is reproduced here. On the one hand, the aim is to carry out simulations-based measures for the laundries and, on the other hand, to identify general key figures from the areas of business management and logistics for industrial laundries. Thus making it possible to compare different industrial laundries. On the basis of the key figures thus obtained, a tool is to be developed which enables the current state of a laundry to be assessed using these and to develop alternatives.

2. DEVELOPMENT OF A PERFORMANCE MEASUREMENT SYSTEM

Key figures from production and logistics are gaining in importance from an economic and production logistical point of view. The key figures are increasingly used in process design and investment decision-making at customers and operators. Over the last decades, various approaches have been adopted for the assessment of machines and systems that go beyond traditional cost accounting and are to identify the true cost of investment and maintenance. The field of industrial laundry is a special feature in the consideration and calculation of life cycle costs. Many companies in the industry do not have information on the cost of machinery and equipment over the scheduled period of use. By increasing customer and product diversity and increased competition a missing, holistic view of the cost factor in business can lead to economic difficulties. The life cycle costs and process costs of washing centrifugal machines, cycle washers or drying systems cannot be determined unequivocally. Supported by a lack of technological equipment for the machinery and plant as well as small investment volumes of companies, new methods have to be used to determine the life cycle costs of machines and systems in industrial washing machines.

2.1. Key Figures and Key Performance Indicators

Key figures are therefore used to compare and assess alternatives or facts (Koether 2011; Reichmann et al. 2017). In doing so, they support the decision-makers in the decision-making process by reducing the multitude of information from a real situation to the relevant information and presenting it in an easily comprehensible form (Weber and Großklaus 1995; Meyer 2007)

This high level of information has led to the fact that key figures have established themselves as an indispensable management tool in business processes (Grochla 1983). They are used for internal and inter-company comparisons and thus enable efficient management, planning and control classification of the key figures can be carried out on the basis of various aspects. A widely distributed division in the literature is based on the statistical methodological approach in absolute and relative key figures. Absolute key figures describe key figures, which are calculated from sums, differences or average values or are present as a single number (Morris and Carter 2005; Meyer 2007). Such as, for example, the profit or the turnover. Relative ratios, which are also referred to as ratios, are, however, quotients of two absolute numbers and are subdivided into:

- structural numbers,
- relation numbers,
- index numbers.

This includes the relationships to the most important key figures (Probst 2014; Siegwart et al. 2010). This property is important because it places unequal data in a meaningful relationship to one another and thus generates a wealth of key figures. As already mentioned, there is the problem with individual characteristic numbers or separately considered key figures that these have a limited meaningfulness. This is due in part to a possibly ambiguous interpretation of the individual key figures. On the other hand, there may be a lack of interdependency.

2.2. Overall objective of the industrial laundry

The objectives of a company can vary greatly according to the requirements of the market and the political environment. In order to better subdivide the multitude of possible goals, a subdivision into three categories is divided (Wöhe and Döring 2010). Which also divides the objectives into economic, ecological and social goals.

- The economic objectives include the classic objectives of a business enterprise, such as the long-term profit maximization or the profitability protection of the company.
- The ecological objectives, for example, focus on the protection of the resources, the limitation of pollutants and the prevention and recycling of waste.

- The social objectives are geared to the needs of employees and thus, for example, to a fair pay as well as good working conditions.

This subdivision can also be applied to the objectives of a laundry, although the category of social goals, despite their indisputable relevance, should not be the focus of this study, but should be concentrated on the economic and ecological objectives.
The economic objectives of a laundry can be derived from the pre-defined arrangement of laundry service with the service companies with a service production. **Fehler! Verweisquelle konnte nicht gefunden werden.** shows the relationship between different aspects of goals and their key figures of a laundry.

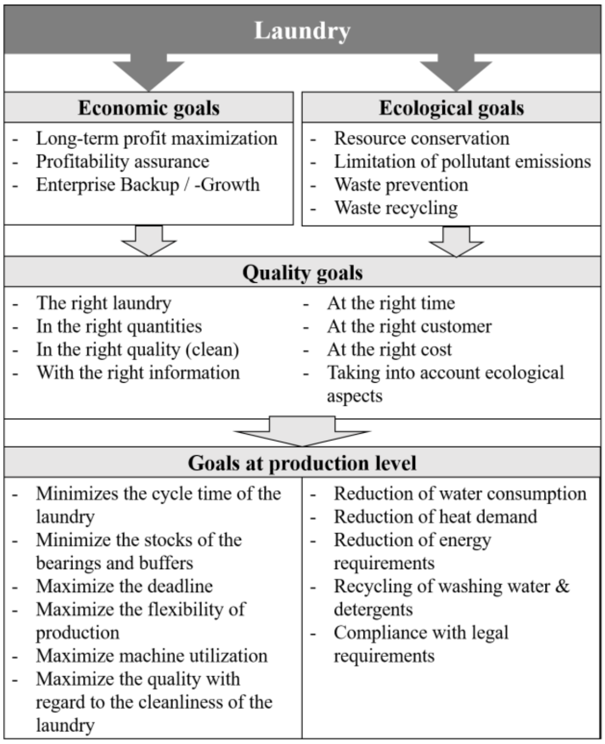


Figure 2: Evaluation of the goals of a laundry with key figures

The topic of laundry is considered in research from various perspectives. In addition to the area of optimization of the washing and drying processes, a great deal of attention is given to the area of energy efficiency or resource conservation (Mechel 2017; Devoldere et al. 2009). Measures are being initiated by various institutes and associations to reduce water consumption and the need for detergents, to recycle existing water to save fresh water or to heat recovery during drying processes.

2.3. Logistics characteristic as a basis

Logistics characteristic curves are generally used for the qualitative and quantitative description of the interrelationships of different key figures as a function of changing boundary conditions (Nyhuis and Wiendahl 2012; Arnold et al. 2008; Cooper and Ellram 1993; Swink 1995). The characteristic curves can be determined either by means of practical test series or by means of different modeling approaches. In the following, the approach of Nyhuis and Wiendahl (2012) is to be concentrated. This has developed a deductive experimental modeling approach to address the conflicting objectives of logistics, which are present, for example in the form of simultaneous demand for low inventories and a high volumetric efficiency, which allows the logistics characteristics of one or a group of parallel arranged working systems using approximate equations (Schönherr 2015; Nyhuis and Wiendahl 2012; Gimenez and Ventura 2005; Schneider 2004). The characteristic theory according to Nyhuis and Wiendahl (2012) represents a robust and error-tolerant model. Therefore, an application is also possible if some of the required parameters are not available or can only be determined in a specific way. Even violations of the required process stability can be compensated. This is, however, to be verified within the framework of a valid validation of the respectively produced model, whereby a practice-oriented approach, for example in the form of a comparison with historical data, is recommended (Nyhuis and Wiendahl 2012). Further work has emerged from the approach developed, which also allows the derivation of time characteristics (Yu 2001) and process characteristics in addition to the production characteristics of individual systems (Guide 2000). The application possibilities of the characteristic theory are correspondingly manifold. This extends from the modeling of the current state of the system in question to the modeling of alternative possibilities for action, as well as the logistical positioning of the system, to the analysis and prognosis of the impact of planned measures. This makes it a way to support decision-making processes within the company. This is characterized by simple comprehensible calculations and a low modeling effort.

3. RESTRUCTURING OF THE RECEIVING AREA IN AN INDUSTRIAL LAUNDRY

The introduction of RFID technology, as well as an automatic collection and identification system in the receiving area provide the necessary transparency in order to plan the wash sequence order not only with experience but to use different known strategies from other production systems. Three strategies for the planning of the wash sequence were selected and compared with a simulation model with regard to different disturbance scenarios according to the target criterion of delivery reliability

3.1. Combined storage strategy

The first step in the development of a combined stock strategy is the selection of an appropriate storage technology. For this purpose, a detailed analysis of the events in the goods receipt as well as the available storage aids was carried out. It has been found that the bottom line storage is best suited in the goods entrance since the laundry is bundled and transported in roll containers.

3.2. Combined storage strategy

The first step in the development of a combined stock strategy is the selection of an appropriate storage technology. For this purpose, a detailed analysis of the events in the goods receipt as well as the available storage aids was carried out. It has been found that the bottom line storage is best suited in the goods entrance since the laundry is bundled and transported in roll containers. In order to determine an appropriate allocation strategy correlations between the complex customer and product structure as well as the treatment processes using expert interviews were analyzed. The result of the analysis is shown in Figure 2. The various structures show that a zoning as a placement strategy is very suitable. A chaotic as well as fixed-location storage means that each parking space is directly accessible at any time. Due to the product and customer structure, however, this is not absolutely necessary in laundries. In addition, the existing storage area is generally not sufficient. Therefore zoning according to the different structures is a very good solution. From the various contexts (economic type, type of customer, type of handling and customer group), 48 combinations for determining the zones can be derived. Priority rules were set up to determine the order of allocation. First, the laundry is assigned according to the type of customer, then by customer category, then by business type and then by the type of handling.

3.3. Validation of the combined storage strategy and determination of the size of the zones

To the practical application of storage strategy and zoning validate a simulation study was conducted. The logistical system is the laundry entrance. The input data are different roll containers, which arrive with different contents as truck deliveries. Consequently, the event-discrete simulation approach is selected for imaging the container flows. Stochastic distributions show the arrival frequencies and contents of the truck deliveries as well as the driving and loading times of the driverless transport systems. The storage strategy with the appropriate zoning, the number of actuators and buffer areas as well as the frequency and number of roll containers are parameterized via the influencing variables. The output data are described by constituents in the incoming goods, utilization, still and wait times of the automatic collection system and the driverless transport system. The discrete event simulation model was created using Tecnomatix Plant Simulation (Figure 2). The input variables such as the number, content and frequencies of the incoming roll containers are generated randomly.

The storage consists of 48 zones, which are approached by the driverless transport system. To this end, the storage area was divided into different sections. The storage strategy is carried out based on the previous priority of the roll container and corresponds to the combined storage strategy. The outsourcing strategy corresponds to the 1: 3 rule (to a wash load of dry washing followed by three wash loads flatwork) which is often used in laundries. Through the monitoring and control of individual processing steps of the incoming roll container in the simulation model (trace file analysis), the verification has been successfully performed. Because of an insufficient data basis from the real laundry for the validation of the simulation model, the partial results were validated by expert discussions. These were conducted by field inspection and recording of process times and process descriptions in direct cooperation with the laundries. The experiment plan includes the selection of scenarios that are tested to determine the utilization of the zones in the incoming goods. The various test scenarios consist of a total of 3750 randomly generated roll containers, which pass through the simulation model 50 times. The observation period is 24 hours with a transient phase of 16 hours. Consider the average throughputs of the respective occupied zones over the 50 cycles. The utilization of the respective zone results in the zoning of the storage spaces.

3.4. Results

The experiments show that is no disturbances in the materialflow through the combined storage strategy. In the process in receiving are no delays in dispatching the goods. The average throughput per zoning can be evaluated with the scenarios. As shown in Figure 5, the containers 30 produced occupy up to 48 possible zones. The Top 3 zones with the highest average throughputs

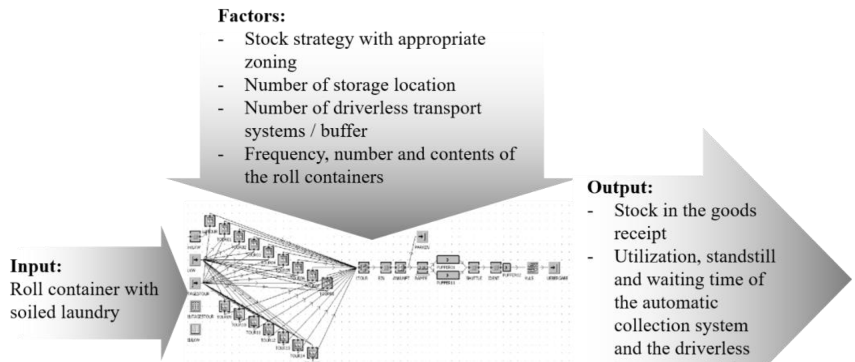


Figure 3: Input and output data as well as influences of the simulation model for the restructuring of the incoming goods

are NPRIO13 with 49 containers, NPRIO12 with 32 containers and NPRIO10 with 25 containers. The zones receive containers from hotel customers with rental and paywear, the majority of which consists of drying laundry. The zones EPRIO14 and EPRIO32 EPRIO34 represent the smallest zones, each with a container.

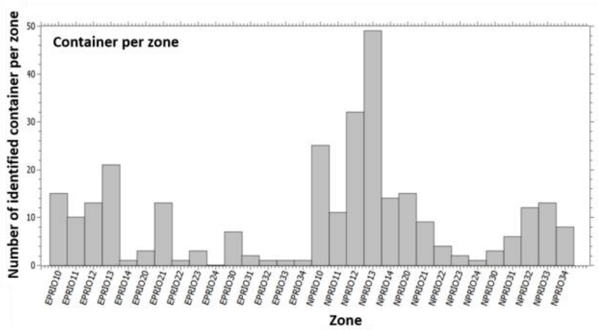


Figure 4: Number of containers per zone

These zones include 24-hour customers with an increased percentage of inadequate laundry. The overall results give a recommendation for the distribution of the zones and the number of pitches in the zone. The combination of zoning and priority rule corresponds to the combined acquisition and storage strategy.

3.5. Conclusion

The low storage space, increasing customer and product structure and the large number of manual operations in the laundries prevent a transparent storage process. Only when transparency is given adequate zoning and storage space allocation can take place. Through the simulation study and the development of a combined allocation and storage strategy, a recommendation for the zoning and the number of parking spaces in the respective zone could be given. With up to 48 potential zones, laundries are able to store different roll containers with different product compositions.

4. PLANNING OF WASH ORDER SEQUENCES

The introduction of RFID technology as well as an automatic identification system in the goods receipt provides the necessary transparency in order to plan the washing order order not only with experience but to use different known strategies from other production systems. There were chosen three strategies for planning the washing order sequence and compared with a simulation model with respect to different disturbance scenarios for the outcome measure delivery reliability. At the moment, the washing orders are planned with the employees' experience. Due to the transparency gained in the goods receipt, further strategies for order planning can now be implemented. Three typical planning strategies were selected:

- First-Come First-Served (FCFS)
- Shortest Processing Time (SPT)
- Longest Processing Time (LPT)

Depending on the tour schedule, it is specified which customer is washed when. The wash orders of a customer are then adjusted according to the experience of the employees in a ratio of 1: 3 (see chapter 3.3).

4.1. Simulation model for the evaluation of wash order sequences

To evaluate the three selected strategies for the wash order sequences, the laundry production was modeled as a simulation model. For this, the material flow from the receiving area to the goods issue is considered. Individual pieces of laundry are combined to wash orders and the wash supplied. After the washing line, the orders are partly separated and reassembled into processing orders. In the warehouse and at the picking stations, the laundry items are separated again and picked to customer orders. The simulation model is thus intended to map the flow of individual items of laundry. However, this can lead to long simulation times because of the diversity and large number of items of laundry. The different types of laundry are represented as product types and the different orders than product portions. The process times of manual treatment processes have been mapped as stochastic distributions. The verification of the model was performed with a trace-file analysis. Individual model building blocks were checked for their correctness and the flow of the individual wash orders was followed by production.

In order to validate the model, the results from the preexisting surplus calculation were compared with the results from the test runs of the simulation model. For the evaluation of the wash order, a basic scenario was created, which includes the delivery of laundry by different customers as well as the delivery of customer orders with delivery dates. For this scenario, the three selected strategies for order planning were used and validated with the simulation model. In addition, the delivery reliability among the three successor strategies as well as among three disturbance scenarios:

- Disorder of a washing machine,
- disorder ironer,
- occur of an rush job.

In total, 3 simulation runs were carried out with 24 hours per scenario and strategy.

4.2. Results

For all three strategies and the four scenarios, the delivery dates of the runs were recorded and their average values were considered (Table 1).

Table 1: Mean values the completion time of the different scenarios and order strategies

	Scenario without disorder	Scenario disorder washing machine	Scenario disorder ironer	Scenario rush job
FCFS	14:02	14:20	16:34	14:20
SPT	14:13	13:46	16:06	14:15
LPT	13:48	13:47	16:45	14:15

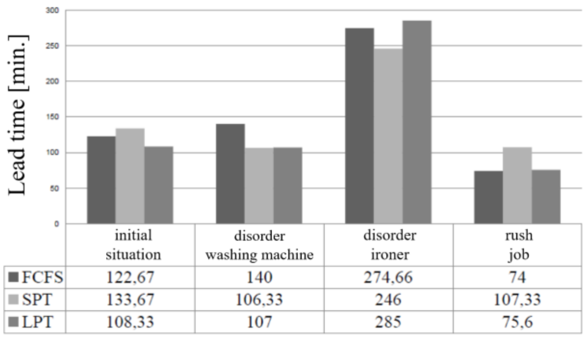


Figure 5: Mean values of the throughput times of the various scenarios and sequence strategies

The required target delivery date was at 4:30 pm. It can be seen that the strategies FCFS and SPT meet the delivery time under normal conditions as well as under interfering influences. With the LPT strategy, the requested delivery date is adhered to, except in the scenario of a lack of a fault. This suggests that the failure of a deficiency has a major impact on the process times and the end of the laundry cycle. The completion dates of all three strategies are delayed by two hours in case of a fault. This relationship is also shown when considering the average throughput times (Figure 5).

4.3. Conclusion

In summary, it can be said that the production of a laundry can be mapped with a discrete-event simulation model. The comparison and evaluation of the different strategies for order planning have shown that all three strategies are suitable, but that delays can occur under interfering influences. Particularly the failure of a defect is a big problem in the laundry in question and leads to delays. Therefore, a failure should be eliminated as soon as possible or prevented.

5. SELECTION AND DEVELOPMENT OF KEY FIGURES AND PERFORMANCE MEASUREMENT SYSTEMS FOR INDUSTRIAL LAUNDRY

The resulting key figures are used to develop a tool which, on the one hand, allows an assessment of the current state of the laundry in question, and can also be used in the course of planning and control. The pre-defined requirements profile of the tool to be developed reflects the characteristics of a model-oriented decision support system with a descriptive and prognostic character (Gaul and Both 1990). These are to be understood as assistants of the users involved in the decision-making process as well as with the help of methods and models of computer science, data analysis, statistics and operations research, such as those involved in control and planning tasks, (Müller and Lenz 2013; Watson and Wixom 2007; Lassmann 2006; Avison and Fitzgerald 2003; Gaul and Both 1990). The approach developed by Nyhuis and Wiendahl (2012) and further developed by Schneider (2004) represents a good alternative to the support of decision-making processes in laundries, which is characterized by simple comprehensible calculations and a low application effort. The calculated results are displayed visually in the characteristic diagrams and can be modified by varying the input parameters. The key figures obtained during the calculation reflect those key figures for the logistics performance targets of a laundry. The complexity of the varying laundry types and structures within the laundry can also be represented well and flexibly on the basis of the approximations and the use of mean values. In addition, an implementation in Excel with the support of Visual Basic for Applications (VBA) can be implemented well. The selected approach for the calculation of the logistical characteristics requires that the respective work systems of the production area be examined in advance for the evaluation of a production area.

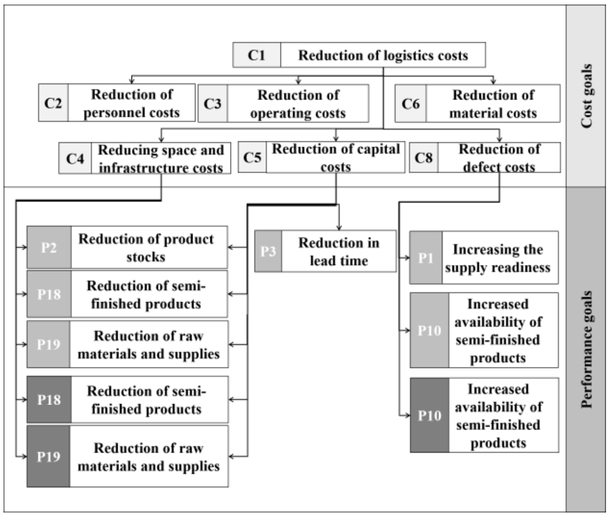


Figure 6: Cost target system adapted to industrial laundry

The calculation of the logistic characteristics via the approximation functions is the basis for the order times of the orders which are passed through the work system. The respective order time of an order is calculated over the individual time per unit of measure, which is to be understood as the default time for the work system, the set-up time, as well as the lot size and the laundry items contained in the wash order.

Proc. of the Int. Conference on Modeling and Applied Simulation 2018,
ISBN 978-88-85741-07-2; Bruzzone, De Felice, Frydman, Longo, Massei and Solis, Eds.

5.1. Prototypical implementation

The tool is divided into two different modules which contain different functions. The first module makes it possible to examine a single work system with a fixed number of identical, parallel workstations without considering a higher-level production area. An example of such a working system would be a single washing centrifugal machine or a region in which there are several identical workstations for sorting the incoming jobs. In this module, it is possible to evaluate the current status of the work system by means of key figures and characteristics, and to identify potentials with regard to the power, the transit time and the stock, as a target value, by varying the mean stock. In addition, a modification of the input data, the creation of variants of the working system, a variation of the input data and the storage of the results obtained is possible. The second module of the tool is used to investigate a production area, which corresponds to the production area of a laundry in this tool. It is necessary first to record and investigate the work systems that are present in the laundry and must be taken into account within the framework of the investigation.

As a result, the user is not only provided with the data of the production area, but also all the data in the working systems under consideration.

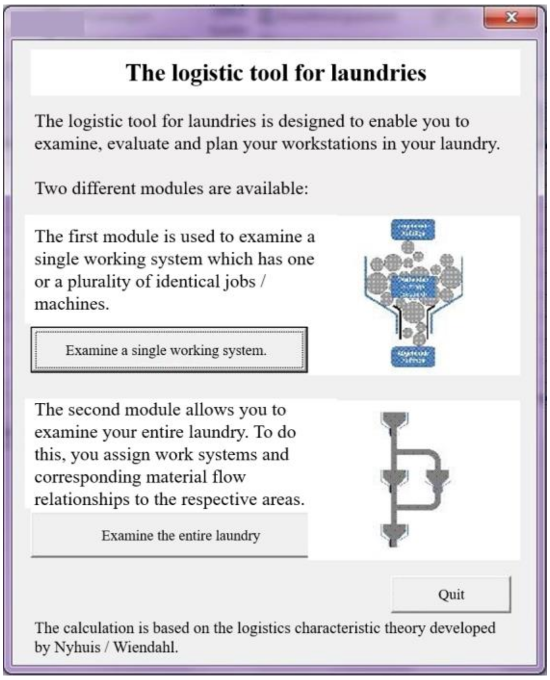


Figure 7: User form of the key figure tool

The data thus obtained, together with the information on the proportion of the material flow, is used to calculate the characteristics of the production area and to provide the user with information regarding the stock, the performance, the processing time and the bottleneck system of the production area under consideration. The input of the data of the operating systems is also carried out with the help of a user form, whereby a user interface identical to the description is used for the first input of the data, the change of the input data and the variation of the data.

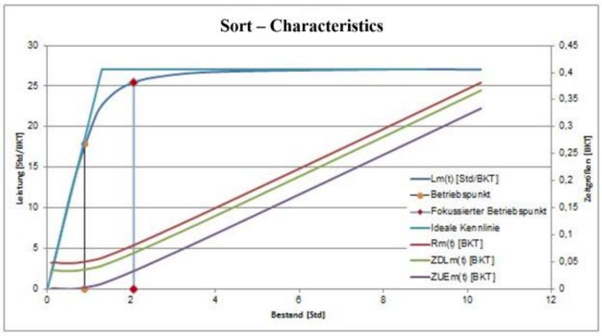


Figure 8: Logistics characteristics of the "Sorting"

The user form allows the creation of customized user interfaces, which serve as an interface between the tool and the user for input and variation of data. When the tool is started, the user is presented with a startup interface that informs about the purpose and the alternative work modules of the tool.

Furthermore, it allows a selection of the respective module via buttons and allows the completion of the program via a third button (Figure 7).

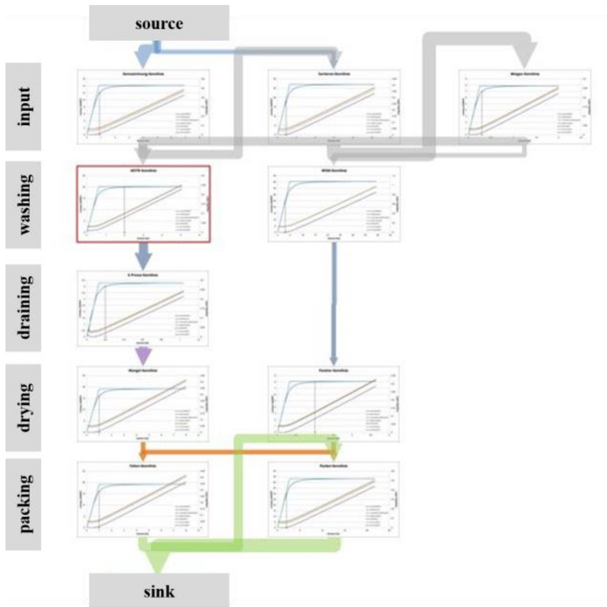


Figure 9: Illustration of the material flow

The calculations listed are already stored in this. When the input data is transferred, the key figures and the characteristics are calculated automatically. The key figures are displayed in a table and contain information on the stock, the processing time, or the system utilization. In addition to the table, there is a diagram area in which the characteristics and the current operating state of the system under consideration are displayed (Figure 8). Using buttons stored in the Excel sheet, the user is able to change the input data, create a variant, vary the medium stock, simulate and focus a different operating state, and save the Excel sheet. A further function, which is available in the course of viewing a production area, is the possibility of displaying the material flow of the current production area in a Sankey diagram (Figure 9).

6. COMBINATION OF SIMULATION AND PERFORMANCE MEASUREMENT SYSTEM

After the simulation study carried out in the laundry, it became clear that without a large use of key figures the significance of the simulation results is low. The difficulty of determining logistical data within industrial laundry was extensively discussed in Chapter 3. Through the development of a uniform measurement system, it is now possible to model and evaluate simulation studies for laundries significantly better. The extensive investigations into key figures and measured variables are no longer necessary. The tool can already be used to make a pre-selection. The creation of the conceptual model for any problem in the laundry can thus be carried out quickly and inexpensively. The presentation of the individual work systems provides a first qualitative image of the laundry. By using Sankey diagrams the flow of information can be viewed in advance and changes are made. The developed measurement system complements the modeling and simulation of the problems of industrial laundry.

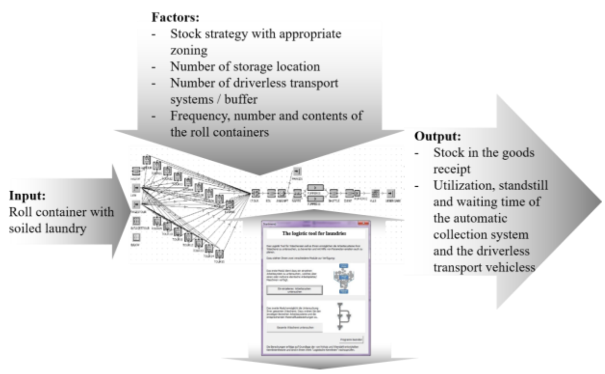


Figure 10: Schematic representation of the input and output data as well as influences of the simulation model supplemented by the function of the tool.

The verification and validation of the tool was carried out using the simulation study described in chapter 2. The results were interpreted much better and stronger and the results will be better used. There was, of course, a considerable time-saving in the production of the model. The verification and validation of the model could be carried out more quickly by the representation of the conceptual model from the determination of the individual work systems. The tool affects neither positive nor negative, the results of the simulation runs. An expansion of the tool and greater integration into the simulation is planned. For this purpose, the determination of key figures and the representation of the concept model direct could be transferred directly into the simulation model. These tasks must be solved in the future.

7. SUMMARY AND OUTLOOK

In this article two simulation studies for laundries were presented. In the future, the results of the restructuring of the goods warehouse will be implemented in a real laundry. The simulation of the laundry processes, in addition to discrete event simulation approach still performed with a mesoscopic approach. Further strategies for sequence planning are to be tested as well as other interfering scenarios are developed and their effects assessed. The paper shows that the laundry cycle offers a great potential for improving process flows, which can be exhausted with simulation studies.

In addition, it was possible to determine key figures from the laundry. In the present study, a tool was developed to evaluate real and planned laundry using pre-established general logistical data. The actual development of the tool has been preceded by a detailed analysis of the system of industrial laundries. For this purpose, the processes and the classification of the concept of the laundry were thoroughly examined and the results of previous work were considered. It was possible to derive general logistical performance and cost target systems for weighing plants This has been used to derive general logistical indicators for laundries and have thus resulted in 21 general metrics, 14 of which are the description of the performance targets and seven the description of the cost estimates. With regard to the transferability of the results obtained, it can be said that the developed tool is generally suitable for the imaging of work systems as well as production areas and for minor adjustments in data entry as well as a certain verification and validation, also in areas away from laundries can come. The derived key figure and target systems can generally be applied to laundries and can be adapted as required to the needs. The combination between the key figure system and the simulation model will be focused even more strongly in the future. In this context, important functions in the process transparency for industrial laundry and the machines are explored with future research topics of automated modeling and real-time connection as well as data visualization.

REFERENCES

Arnold, D.; Isermann, H.; Kuhn, A.; Tempelmeier, H.; Furmans, K. (ed.): Handbuch Logistik. Berlin: Springer 2008.

Avison, D.; Fitzgerald, G.: Information system development: Methodologies, techniques, and tools. London [etc.]: McGraw-Hill 2003.

Cooper, M.C.; Ellram, L.M.: Characteristics of Supply Chain Management and the Implications for Purchasing and Logistics Strategy. The International Journal of Logistics Management 4 (1993) 2, p. 13–24.

Devoldere, T.; Willems, B.; Duflou, J.R.; Dewulf, W.: The eco-efficiency of reuse centres critically explored – the washing machine case. International Journal of Sustainable Manufacturing 1 (2009) 3, p. 265.

Gaul, W.; Both, M.: Computergestütztes Marketing. Berlin: Springer 1990.

Gimenez, C.; Ventura, E.: Logistics-production, logistics-marketing and external integration. International Journal of Operations & Production Management 25 (2005) 1, p. 20–38.

Grochla, E.: Erfolgsorientierte Materialwirtschaft durch Kennzahlen: Leitfaden zur Steuerung und Analyse der Materialwirtschaft. Baden-Baden: FBO-Verl. 1983.

Guide, V.R.: Production planning and control for remanufacturing. Journal of Operations Management 18 (2000) 4, p. 467–483.

Koether, R. (ed.): Taschenbuch der Logistik. München: Fachbuchverl. Leipzig im Carl-Hanser-Verl. 2011.

Lassmann, W. (ed.): Wirtschaftsinformatik: Nachschlagewerk für Studium und Praxis. Wiesbaden: Gabler 2006.

Mechel, C.: Ökoeffizienzanalyse zum Vergleich heterogener Unternehmen: Springer Fachmedien Wiesbaden GmbH 2017.

Meyer, C.: Betriebswirtschaftliche Kennzahlen und Kennzahlen-Systeme. Sternenfels: Verl. Wiss. & Praxis 2007.

Morris, M.; Carter, C.R.: Relationship Marketing and Supplier Logistics Performance. The Journal of Supply Chain Management 41 (2005) 4, p. 32–43.

Müller, R.M.; Lenz, H.-J.: Business Intelligence. Berlin, Heidelberg: Springer Vieweg 2013.

Nyhuis, P.; Wiendahl, H.-P.: Logistische Kennlinien: Grundlagen, Werkzeuge und Anwendungen. Berlin, Heidelberg: Springer 2012.

Probst, H.-J.: Kennzahlen: Richtig anwenden und interpretieren. München: Redline Verlag 2014.

Reichmann, T.; Kißler, M.; Baumöl, U.; Hoffjan, A.; Palloks-Kahlen, M.; Richter, H.J.; Schön, D.: Controlling mit Kennzahlen: Die systemgestützte Controlling-Konzeption. München: Verlag Franz Vahlen 2017.

Schneider, M.: Logistische Fertigungsbereichskennlinien. Düsseldorf: VDI-Verl. 2004.

Schönherr, M.: Wertorientiertes Logistikmanagement, Dissertation, 2015.

Siegwart, H.; Reinecke, S.; Sander, S.: Kennzahlen für die Unternehmensführung. Bern: Haupt 2010.

Swink, M.: The Influences of User Characteristics on Performance in a Logistics DSS Application. Decision Sciences 26 (1995) 4, p. 503–530.

Watson, H.J.; Wixom, B.H.: The Current State of Business Intelligence. Computer 40 (2007) 9, p. 96–99.

Weber, J.; Großklaus, A. (ed.): Kennzahlen für die Logistik. Stuttgart: Schäffer-Poeschel 1995.

Wöhe, G.; Döring, U.: Einführung in die allgemeine Betriebswirtschaftslehre. München: Vahlen 2010.

Yu, K.-W.: Terminkennlinie: Eine Beschreibungsmethodik für die Terminabweichung im Produktionsbereich. Düsseldorf: VDI-Verl. 2001.

AUTHORS BIOGRAPHY

DAVID WEIGERT studied Industrial Engineering with specialization in Logistics at the Otto-von-Guericke-University Magdeburg. He became a research associate at the Chair Logistical Systems at the Otto-von-Guericke-University Magdeburg and scientific project assistant at the Fraunhofer Institute for Factory Operation and Automation IFF Magdeburg. His areas of competence are the analysis and optimization of logistics processes, as well as modelling, simulation and optimization of logistics systems.

FRANK RYLL is a research and project manager as well as Deputy Business Manager Logistics and Factory Systems at the Fraunhofer Institute for Factory Operation and Automation IFF. He received a doctoral degree in engineering from Otto von Guericke University Magdeburg. His research interests include Factory Planning, Energy and Resource Efficiency (Total Resource Management) as well as maintenance and service management (strategies, TPM, RCM).

MARCEL MÜLLER studied Industrial Engineering with specialization in Logistics at the Otto-von-Guericke-University Magdeburg. He became a research associate at the Otto-von-Guericke-University. His areas of competence are the analysis and optimization of logistics processes.

ACO TOPOLOGY OPTIMIZATION: THE PHEROMONE CONTROL FOR CONSIDERING THE MECHANICAL KANSEI

Nanami Hoshi[(a)], Hiroshi Hasegawa[(b)]

[(a)(b)] Graduate School of Engineering and Science, Shibaura Institute of Technology, Japan

[(a)]mf17061@shibaura-it.ac.jp [(b)]h-hase@shibaura-it.ac.jp

ABSTRACT
In a structural concept design, to create topology as solution principle from functional requirement is greatly dependent on the accumulated engineering knowledge, experience and know-how. These engineering senses are called as The Mechanical Kansei. Ant Colony Topology Optimization (ACTO) was applied major principal stress in topology optimization. ACTO considered major principal stress and Von Mises stress using by pheromone control and can apply the accumulated engineering knowledge. However, A part of experience and know-how which are had designer can consider, because experience and know-how is different each designer. Therefore, ACTO is not perfect for the Mechanical Kansei. In this study, we confirm that ACTO is possible to consider individual differences in the Mechanical Kansei each of the designers using by pheromone control. In addition, we improve and change the method to consider individual differences in the Mechanical Kansei.

Keywords: Mechanical Kansei, Topology Optimization, Ant Colony Optimization, Principal stress

1. INTRODUCTION

The Mechanical Kansei means a decision capability to evaluate an impression unconsciously, intuitively with informational synthesis in order to mechanics (Architectural Institute of Japan 2014). In a structural concept design, to create topology as solution principle from functional requirement is greatly dependent on the accumulated engineering knowledge, experience and know-how(Fischer J., Mikhael J.G., Tenenbaum J.B. and Kanwisher N. 2016.). These engineering senses are called as The Mechanical Kansei. However, the mechanical Kansei is difficult to apply for a product and system. For example, topology optimization has not considered the mechanical kansei.

Topology optimization has been used for structural optimization, and the various techniques and approaches of topology optimization have been developed and researched since about 1985 (Nishiwaki, Izui and Kikuchi 2012). Topology optimization can change shape, size and number of holes therefore topology optimization is the most flexible methodology in structural optimization. For the maximum stiffness of structures, the mean compliance is often applied in most of Topology Optimization methods as the objective function, and the constraint is imposed on a somewhat arbitrarily specified volume. Moreover, for stress-based of structures, either the number of iterations or the volume is used as the termination condition for Topology Optimization (Q.Q. Liang and G.P. kteuen 2002). However mechanical product designers often consider “a weight minimization under a stress constraint” and the force flow i.e. principal stress vector when design optimal structure. These mechanical thinking is based on their accumulated engineering knowledge, experience and know-how. This is the thinking which can search by engineering sense, named the Mechanical Kansei. If Mechanical Kansei cannot be well used to design topology which satisfied various functional requirements, its design process has a difficulty to draw attractive topology for engineer.

We applied major principal stress in topology optimization using Ant Colony Optimization (ACO), named Ant Colony Topology Optimization (ACTO) (Hoshi, N. and Hasegawa, H. 2017).

ACO is a stochastic local search method that has been inspired by the tracks of pheromone and following a behavior of ant swarm intelligence (Dorigo, M. and Stützle, T 2009). Structure Optimization using ACO has been suggested from 2004(Camp and Bichon 2004, Kaveh, Hassani, Shojaee and Tavakkoli 2008). ACTO considered major principal stress and Von Mises stress using by pheromone control. ACTO is one of the few topology optimization approach considering principal stress. ACTO can apply the accumulated engineering knowledge, it is based on principal stress and Von Mises stress. However, A part of experience and know-how which are had designer can consider, because experience and know-how is different each designer. Therefore, ACTO is not perfect for the Mechanical Kansei.

In this study, we confirm that ACTO is possible to consider individual differences in the Mechanical Kansei each of the designers using by pheromone control. In addition, we improve and change the method to consider individual differences in the Mechanical Kansei.

2. ANT COLONY OPTIMIZATION APPROACH

Ant Colony Optimization (ACO) is a nature-inspired metaheuristic for the solution of hard combinatorial optimization problems (Dorigo, M. and Stützle, T 2009). Ants deposit a pheromone on the path while walking. After iterate deposits pheromone, ants can create a trail to back from the nest to food sources and can sense the pheromone trails and follow the path to food discovered. Also, pheromone has a character that is evaporation. In not optimal path, pheromone evaporation is the mechanism that decreases over time the pheromone deposited by previous ants. Pheromone evaporation and deposits is important for ACO, its called pheromone update. The pheromone update is commonly implemented as:

$$\tau(t+1) = \mu\tau(t) + \sum_{i=1}^{m} \Delta\tau_i \qquad (1)$$

μ: Reduction factor(pheromone evaporation)
m: individual Number i:selected route $\Delta\tau_i$:pheromone deposit t: Generation Number

3. ANT COLONY TOPOLOGY OPTIMIZATION APPROACH

Ant Colony Topology Optimization (ACTO) is the Topology Optimization approach that is used ACO (Hoshi, N. and Hasegawa, H. 2017). Also, Genetic Algorithm (GA) is often used Topology Optimization (Ohsaki, M. 1995).

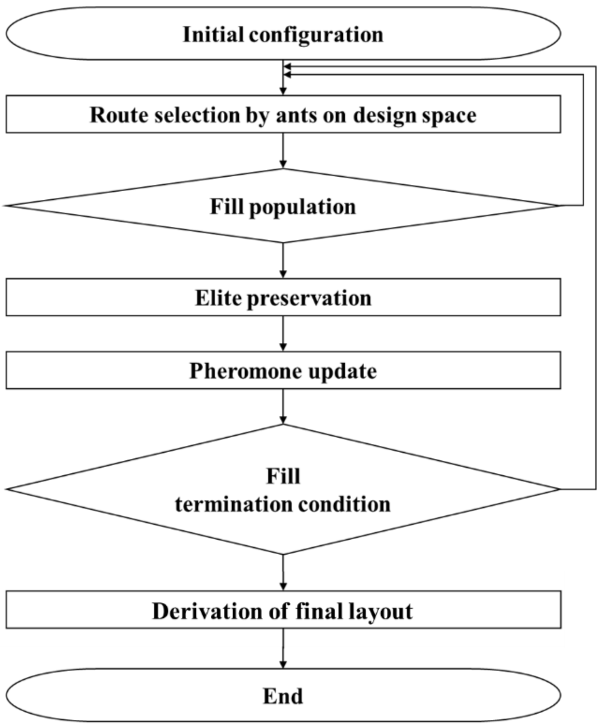


Figure 1: Flow chart of Ant Colony Topology Optimization

The mesh selected by the ant is judged as a structure on ACTO. Ultimately, one structure layout is completed and called the individual. Several individuals are created in one generation. When the route search of the ant in the generation is completed, the pheromone is updated based on the distribution of Von Mises stress and the trajectory of the ant walking route, for each individual. Ants chooses routes in the next generation using by the pheromone updated. After completion of the set number of generations, banalization is performed using the threshold value, and a final structure layout is created. The flow of this method is divided into sections and explained in detail:

1) Initial Configuration

In initial configuration, object model is applied meth of Finite Element Method (FEM). This section should be setting the initial pheromone for each mesh. Result of FEM in object model is become the initial pheromone. We choose Von Mises stress as the initial pheromone, because Von Mises stress which is scalar value, is easy to apply in the proposal method.

2) Ant of route selection

The set first pheromone is use by ants to generate routes. Also, several ants create routes This section is used some method, as follow:

- Ranking selection

It takes time to derive the optimum solution, when is no difference in the pheromone value for each mesh. Therefore, ACTO is used ranking method which is adjusted the convergence of the solution. Figure 2 shows the procedure of the ranking method. we will reset the value according the pheromone. The ranking method redistributes the value according the pheromone. (Table is shown the redistribution value in the proposal method.) It is possible to control the probability in route search regardless of the difference in the pheromone.

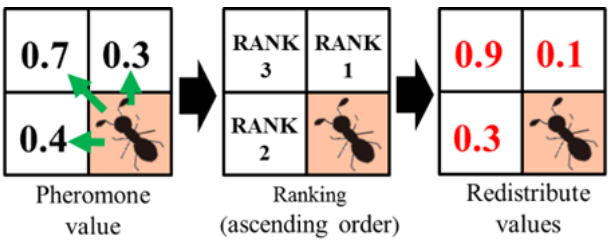


Figure 2: Ranking method (ranked in order of the pheromone)

- Roulette wheel selection

Roulette wheel section (Lipowski and Lipowska 2012) is the probability of selection is proportional to redistribute values of the liner lank method. The better fitted redistribute values of the liner lank method, the larger the probability of selection (Figure 3). This method considers N individuals, each characterized by redistribute values of the liner lank method. Selection of an individual choose randomly. The selection probability of i-th individual P_i is follow as:

$$P_i = \frac{\text{Exp Val}(r,t)}{\sum_{i=1}^{N} \text{Exp Val}(r,t)} \quad (i = 1,2,3 \dots, N) \qquad (2)$$

Exp Val(i,t): redistribution value of liner lank method
i: individual r: rank of liner lank method

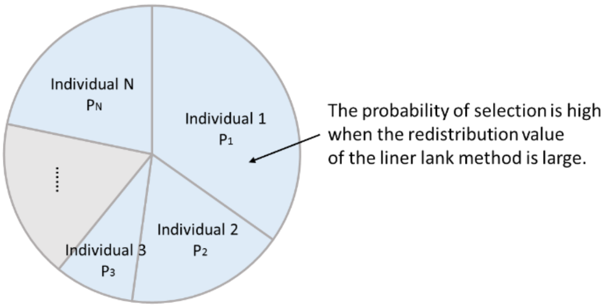


Figure 3: Roulette wheel selection

-Consider with the principal stress vector

The green arrow in Figure 4 is the arrow of the principal stress. The probability of selected the element increase when the element has the principal stress vector (Ito, Hoshi and Hasegawa 2016).

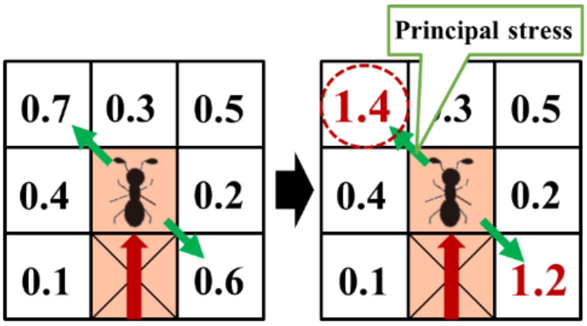


Figure 4: Consider with the principal stress vector

3) **Fills population**

Ants repeat route selection in design space until the set number of individuals is satisfied.

4) **Pheromone update**

This section executes pheromone update where paths' ants have walked and Von Mises stress. Also, existing pheromones decrease because of evaporation of pheromone. Reduction rate of evaporation of pheromone is often use 0.8 to 0.98.

5) **Fill individual**

ACTO repeats generating optimal structures until the set number of generations is satisfied.

6) **Derivation of final layout**

A binary image is generated by binarizing the pheromone. The binary image become the final layout of ACTO.

4. PHEROMONE CONTROL AND ANALYSIS SETTING

4.1. To alleviate stress concentration using by pheromone control

A complex structure often has stress concentration. stress concentration is structure part where stress is concentrated. stress will become strong where the area of structure part is small. Figure 5 shows stress concentration proceed by the small area of structure part. A final layout obtained by topology optimization has the small area of structure part, so it become stress concentration on FEM. ACTO should consider stress concentration on pheromone update, because it is a big effect and problem in ant of route selection. area of structure part. In order to alleviate stress concentration, stress concentration is alleviated by changing the weight nonlinearly as shown in the Figure 6. The Von Mises stress by changed nonlinearly is applied deposits a pheromone on pheromone update.

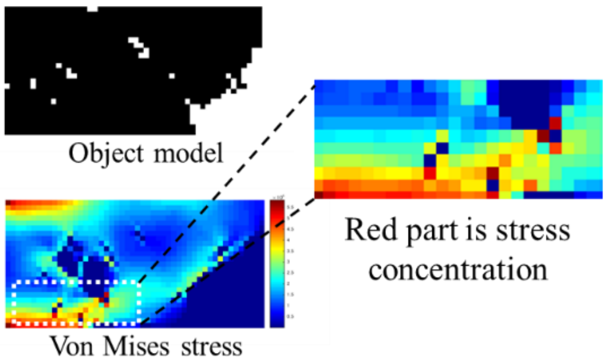


Figure 5: stress concentration in Von Mises stress

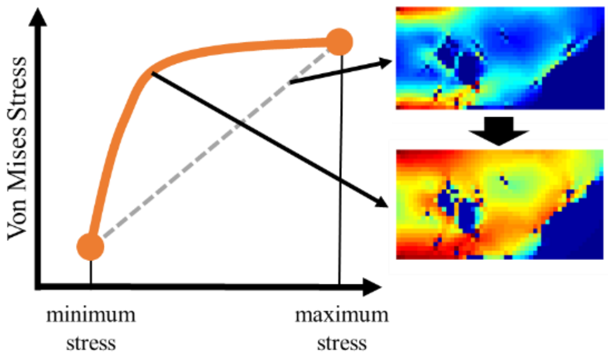


Figure 6: How to alleviate stress concentration

4.2. Analysis Setting

Figure 7 shows object model in this study. In addition, the analysis settings of this paper are shown in Table 1. In this study, we analyze 3 trial considering the Mechanical Kansei, because we want to know the characteristics of the final layout with pheromone control that takes into consideration the principal stress and

Mises stress. Table 2 shows the Mechanical Kansei what is considered in Trials of ACTO.

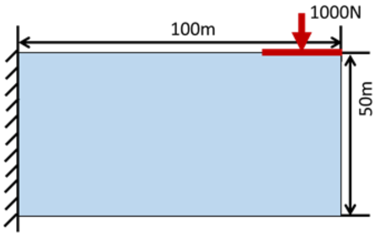


Figure 7: object model

Table 1: Parameters of ACTO

Name	Value
Mesh division Vertical	25
Mesh division Horizontal	50
The number of generation's individuals	5
The number of ant on design space	10
Generation number	100
Reduction factor(pheromone evaporation)	0.8
Factor of deposits pheromone	0.05 - 0.09

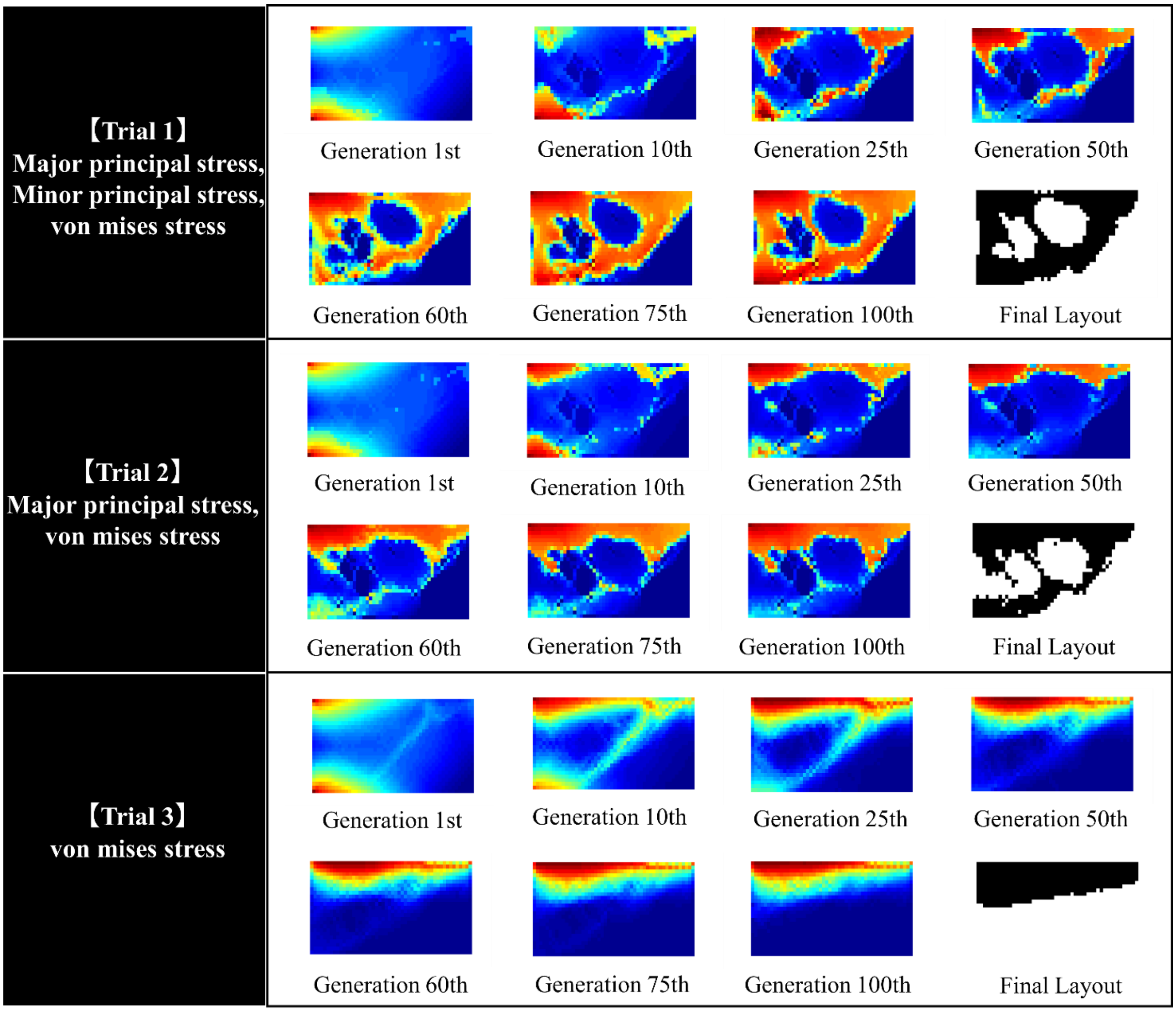
Table 2: Analysis setting of ACTO

	the Mechanical Kansei was considered on Trial
Trial 1	Major principal stress, Minor principal stress, von mises stress
Trial 2	Major principal stress, von mises stress
Trial 3	Von mises stress

5. RESULT AND DISCUSSION

Table 3 shows the final layout obtained by ACTO. We were able to get the final layout with different features in each trial. Trial could get a structure similar to the final layout of the other method shown in Figure 8.
However, the other methods did not consider the principal stress, different internal structures could be obtained in Trial 1. Trial 2 is considered the maximum principal stress and Mises stress. Trial 2 has characteristics similar to Trial 1, the upper beam is led out thickly. Trial 3 has the same characteristic that this is thickly upper beam. Therefore, considering Mises stress leads to deriving the upper beam thickly. The reason is

Table 3: Result of ACTO



that the shortest route is selected as the original characteristic of ACO. Therefore, it is considered that the characteristic that the upper beam near the goal is derived from the start position appears. the upper beam of Trial 1 and Trial 2 is weakened better than trial 3, because it is considered the principal stress.

The pheromone control based on the principal stress and Mises stress is possible to change the final layout depend on the designer's Mechanical Kansei.

However, considering only the Mises stress of other methods has the characteristic that is similar Trial 1, so we need reconsider the setting of ACTO. For example, in the current method, the start point of the path selection of ants is set as the load point, but changing the setting such as random determination is necessary as improvement of ACTO.

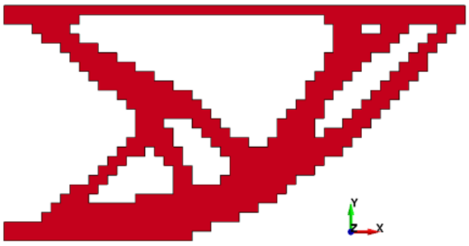


Figure 8: Result of the other method (LS-Tasc)

6. CONCLUTION

We confirmed ACTO could create the final layout that has many characteristics based on the Mechanical Kansei using by the pheromone control. The pheromone control depended on the major principal stress, the minor principal stress and Von Mises stress. Also, we confirmed that the final layout changes by changing the consideration of the principal stress and Von Mises stress using the pheromone control.

ACTO become a supportive tool for embodying each designer's Mechanical Kansei with a structure layout, because ACTO can change consideration of the principal stress and Von Mises stress using the pheromone control. We should improve some section to suggest good final layout for the mechanical Kansei.

REFERENCES

Architectural Institute of Japan., 2014. Architectural Form and The Mechanical Kansei. Maruzen., (in Japanese).

Fischer J., Mikhael J.G., Tenenbaum J.B., Kanwisher N. ,2016. Functional Neuroanatomy of Intuitive Physical Inference. Proceedings of the National Academy of Sciences of the United States of America, E5072–E5081.

Nishiwaki, S., Izui, K., and Kikuchi, N., 2012. Topology Optimization, The Japan Society for Computational Engineering and Science, MARUZEN-YUSHODO Company, (in Japanese).

Liang, Q. Q., Steven, G. P., 2002. A performance-based optimization method for topology design of continuum structures with mean compliance constraints. Computer methods in applied mechanics and engineering, 191(13-14), 1471-1489.

Hoshi, N., Hasegawa, H., 2017. ACO topology optimization: The geometrical constraint by learning overlaid optimal ants route. In 16th International Conference on Modeling and Applied Simulation, MAS 2017. CAL-TEK Srl.

Dorigo, M., & Stützle, T., 2009., Ant colony optimization: overview and recent advances. Techreport, IRIDIA, Universite Libre de Bruxelles, 8.

Ohsaki, M., 1995. Genetic algorithm for topology optimization of trusses. Computers & Structures, 57(2), 219-225.

Lipowski A. and Lipowska D. 2012. Roulette-wheel selection via stochastic acceptance. Physica A: Statistical Mechanics and its Applications, 391(6), 2193-2196.

AUTHORS BIOGRAPHY

Nanami Hoshi received her B. E.(2017) from Shibaura,Institute of Technology(SIT), Japan. She is a master student at Division of Systems Engineering and Science-Graduate School of Engineering and Science, SIT. Her research interests are optimization methods.

Hiroshi Hasegawa received his B.E. (1992) and M.E.(1994) from Shibaura Institute of Technology(SIT), Japan. He received PhD (1998) in Mechanical Engineering from Tokyo Institute of Technology, Japan. He has been working at SIT, and currently is a Professor. He is member of JSME, ASME, JSST, JSCES and JSDE. His research interests include computer-aided exploration, especially multi-peak optimization, robust design and multi-disciplinary optimization.

SCENARIO-BASED ANALYSIS IN HIGH-MIX LOW-VOLUME PRODUCTION ENVIRONMENT

István Gödri[a], Csaba Kardos[b], András Pfeiffer[c], József Váncza[d]

[a]Aventics Hungary
[b],[c],[d]MTA SZTAKI, EPIC Centre of Excellence

[a]istvan.godri@aventics.com, [b]csaba.kardos@sztaki.mta.hu, [c]andras.pfeiffer@sztaki.mta.hu,
[d]jozsef.vancza@sztaki.mta.hu

ABSTRACT

The challenge of high-mix low-volume production has reshaped manufacturing systems causing increased complexity in processes and growing sensitivity to the mix and temporal distribution of demand. Efficient evaluation and experimenting for decision support in such an environment is of key importance, however it is also extremely difficult as the complex interrelation between the affecting factors and the size of the input domain would require a large number of experiments to get reliable results. The paper introduces a method based on advanced data analysis for defining typical input scenarios, aiming to reduce the computational complexity of Discrete Event Simulation (DES) analysis. The presented approach was tested in a real-life combined (manufacturing and assembly) production line and the results showed that using scenarios for representing the typical input allowed reducing significantly the number of experiments required to execute sensitivity analysis of the structural (e.g. buffer size or workforce) and the operational (i.e. sequencing) parameters.

Keywords: Simulation, data analysis, robust planning, sensitivity analysis

1. INTRODUCTION

Today's manufacturing companies have to face increasing product variability and decreasing lot sizes at the same time. This results in a growing complexity both for all planning as well as execution levels of the production.

The challenge of high-mix low-volume production has reshaped manufacturing systems and having efficient evaluation techniques for decision support in such an environment is of key importance. However, it is also extremely difficult as the complex interrelation between the affecting factors and the size of the input domain would require a large number of experiments to get reliable results.

The DES approach, available for modelling productions systems' behaviour on a detailed level, has been applied for decades mainly for the evaluation and support of decisions in planning and control (Banks 1998; Law and Kelton 2015; O'Rielly and Lilegdon 1999). The simulation models that are used for making or evaluating these decisions (e.g., by projecting the values of different key performance indicators, KPIs in time) generally represent the flow of materials to and from processing machines and the operations of machines themselves (Rabelo et al. 2003).

In the paper research results of an in-depth investigation and improvement of the delivery performance of manufacturing plants with a special focus on high mix–low volume production are presented.

The suggested novel top-down process modelling methods are validated and verified by simulation experiments.

2. PROPOSED NOVEL DELIVERY PROCESS STABILIZATION METHOD

2.1. Factors affecting delivery performance related KPI-s

Competitiveness of manufacturing companies is defined in general by three Key Performance Indicators (KPIs):

- cost efficiency;
- quality of products and production processes;
- delivery performance.

If these KPIs are in line with the worldwide benchmark figures produced by manufacturing companies of similar type, then the profitability of the company should be on a level expected by the investors as well.

When assessing delivery performance, one should distinguish between two related KPIs: the fulfilment of the *requested delivery date* (RDD) and the *confirmed delivery date* (CDD). Our focus will be set on CDD which is depending on a number of factors such as raw material inventory level, total supply chain lead-time, manufacturing /machining capability, suppliers' capability, as well as quality of planning (this list is not exhaustive).

1. *Manufacturing/machining capability:* The availability of manufacturing and machining capacities clearly affect CDD. Here, a number of various factors have to be considered, such as the machines' age and condition, the efficiency of implemented total productive maintenance (TPM) policy, the flexibility, changeability and compatibility of resources, the skill and availability of maintenance personnel, the change-over time applied in high mix – low volume environment, the capability of moving a product from one machine center to another one in case of machine break-

down, and the utilization of the machines in general. Investments in new capacities, personnel or maintenance, or flexibility and changeability will definitely incur extra costs on the one hand, but have a positive impact on the delivery performance on the other hand.
2. *Raw material inventory level:* Material availability determines when the product order can be launched in production. The probability to have raw material always available in production can be increased with higher stock levels. However, inventories incur costs as well: typically, the cost of capital of raw material inventory is calculated with a given percentage level defined by each company internally (obviously, this figure is always higher than the actual banking interest rate).
3. *Suppliers' capability:* Performance of suppliers is a key influencing factor in the CDD of a plant. Supply channels are controlled by contracts referring to minimum order quantity (MOQ), required quality, item prices and transportation cost, tooling, as well as supplier's flexibility. While having a reduced supplier network may result in more efficient and frequent deliveries, lower transportation costs and overall prices, it makes, at the same time, the plant more dependable (and vulnerable) to supplier contingencies.
4. *Total supply chain lead-time:* The shorter the total supply chain lead-time, the better are the chances to deliver products at CDD. Production and supply planning have a key role in lead-time reduction. In an ideal case instead of the manufacturing processes, the real bottleneck is the lead-time of raw materials from suppliers. Note that lead-time reduction on the supplier side will not only improve CDD performance of the plant but, at the same time, reduce also the required (safety) stock level.
5. *Quality of planning:* While the requested delivery date (RDD) is an exogenous factor in managing production, the CDD is the result of planning. If the promise confirmation is given to the customer on the basis of careful and principled planning that takes into consideration future load, resource and material availability, and does it in a robust way, then the chances of keeping this promise are clearly better. On the cost side, however, advance planning requires precise and up-to-date status information, appropriate information and communication technologies (ICT), disciplined and orchestrated management of a number of planning functions, and sophisticated decision-making mechanism.

2.2. Delivery process stabilization method

The above (far from exhaustive) list of issues show that the CDD performance is determined by a number of *internal* and *external* factors. Some of the factors are cross-correlated, and efforts in improving delivery performance in any way may easily deteriorate other KPIs, most importantly, *cost*. With other words, CDD improvement is never for free, and a trade-off has to be fund when setting target levels of attainable KPIs. Since in our actual problem domain *product quality is not negotiable*, the key problem boils down to finding ways to improve CDD performance at an acceptable cost that warrants both customer satisfaction and profitability of production.
Hence, the method is articulated around the following – closely related – stages, formulating a top-down approach:

1. *Scope setting and characterization:* Classify situations in high mix–low volume production, delimit those cases where CDD performance can be warranted by traditional techniques of production managements (e.g., by inventory control, or capacity planning). Make an in-depth investigation of cases which are critical, determine the main factors – both internal and external – that affect CDD performance.
2. *Sensitivity analysis and selection of factors:* Make a sensitivity analysis for assessing the impact of the above factors on CDD performance. Select those factors for further investigation whose influence – both positive and negative – on CDD are the most significant.
3. *CDD improvement techniques:* By relying on the selected factors, define those techniques that are implementable in a given production environment and contribute to the improvement of CDD without deteriorating product quality (which is taken as a non-violable requirement).
4. *Delivery performance* – cost trade-off: Assess the cost impacts of the selected (most promising) delivery performance improvement policies and find a balanced trade-off between meeting these two main KPIs. Results expected in form of an implementable method for measuring cost implications of CDD improvement techniques, waging CDD performance against cost and finding an acceptable trade-off.

In the paper methods related to the second item (sensitivity analysis and selection of most influencing factors) are introduced via the analysis of large-scale real-life datasets generated from the archives of a high mix–low volume production facility by applying simulation and data analysis techniques.

3. SCENARIO-BASED EXPERIMENT DESIGN APPROACH

Handling all the influencing factors within one experimental scenario, as described in the previous section, it can be considered as a problem of intractable complexity in the simulation domain. Therefore, the factors have to be separated, moreover, the number of factors and the number of resulting scenarios have to be reduced. In the following section a novel method is proposed for supporting scenario definition for simulation studies having significantly less number of experiments with the same expected output quality.

Classical *Design of Experiment* (DoE) techniques enables reducing computational efforts by selecting and focusing on the factors affecting most significantly the output variable (KPI). A well-known method is *factorial experiment design*, when assigning a possible/usual low and high value to all factors defined. After identifying the most significant factors, a more comprehensive study by the combination of the remaining factors are required. Note that for categorical (qualitative) factors (e.g., dispatching rule for jobs entering the system) the number of possible values is limited. However, non-categorical (quantitative) factors may have "infinite" values to be assigned with. The categorization of these values may reduce the number of scenarios when designing the experiments and not to lose diversity of input degrading modelling accuracy.

The proposed approach introduced here focuses on identifying the similarity between several values of *Demand* (Figure 1), taken as the main categorized input factor.

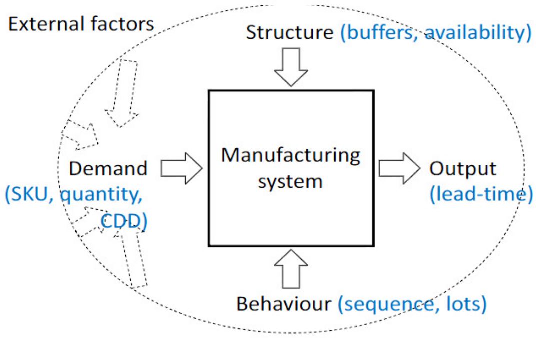


Fig. 1. Influencing factors of a manufacturing system on the production control level.

External factors are covered/modelled by the input factor *Demand* (Figure 1). *Demand* represents SKU (product code) of products, quantity to be produced and CDD, defined at the production planning level in form of a daily input-mix of products to be produced in the system. Product-related features are only considered in the simulation model.

5. *Structure*, internal: structural parameters, e.g., buffer sizes, availability of operators and machines.
6. *Behaviour*, internal: control parameters, e.g., selected routings, sequence of jobs, lot-sizing.
7. *Output*: lead-time and lead-time variance.

The purpose of the experiments is twofold:

a. Validate if the methods are applicable for the reduction of the number of scenarios needed for a comprehensive simulation analysis on system-sensitivity (reducing lead-time variance and so, improving CDD).
b. Categorization of the input factor *Demand* could result in assigning situation-related sequencing and lot-sizing rules to a finite number of demand patterns. This would support the planner in creating production schedules/sequences by applying a formalized method, considering the reduction of lead-time variance, as a primary target.

4. COMPUTATIONAL EXPERIMENTS

4.1. Case study

The production line under analysis is for making complex products, which includes both machining of raw materials and assembly of the final product. In general machinery and assembly areas and operations are separated in the production environment, as in machining usually the equipment is in the centre of decisions, while in assembly the focus of the process analysis is on the human workforce. This *combined* production line, therefore can be viewed as a factory within the factory, which also creates complex interrelated connections between the KPIs and the structural and operational parameters of the line.

Model building and abstraction is essential for any kind of analysis and decision support process. For sensitivity analysis in such a complex environment, where analytical solution is probably out of the scope, the application of declarative tools such as Discrete Event Simulation (DES) of material flow is widely accepted.

4.2. Description of the material flow

The machining area contains four CNC machine centres and two CNC turning machines and a conveyor line. The washing operation for machined parts is positioned between the machining and assembly area. There is a machine operator for each machine.

The process starts with the aluminium tube cutting, followed by the machining of the tubes – according to the production schedule. The piston rod manufacturing is done in parallel with the machining of the tubes. After these operations, the tubes and piston rods belonging to the same order are sent to be washed. After the washing operation these parts are put in a box. These boxes are in line, waiting for the final assembly operation.

Conveyor belts transport the cut material to four milling and two turning machines where the machined surfaces are finished. Each machine requires human operators for the change-overs and setups before the processes, served by a pool of machinery workers in the area, which also means that no dedicated workforce is assigned to any of the machines. Accordingly, the list of assigned tasks is defined by rules for each worker. The final step in the machinery is a washing station with manual material transport.

The machinery area provides supply to the preassembly area, where semi-finished products required in multiple product families are assembled. The final assembly and test of the finished products are performed in two assembly work cells, each one operated by one or two dedicated workers in a one-piece-flow production with manual material transport. Most equipment used in the assembly are designed especially for the products, however –in order to handle multiple product families and variants– there is a wide variety of applied fixtures and tools, which require manual change-over.

The conveyor belt has a length-dependent buffer-capacity and each station has a specified buffer area as a fixed-size number buffer.

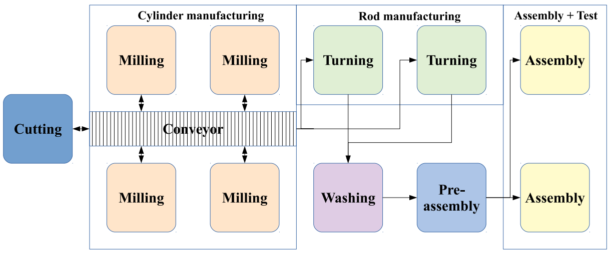


Fig. 2. Overview of the material flow in the production line.

4.3. Simulation modelling

The simulation model of the production line was created in the Plant Simulation DES software. The processes durations are modelled as product-dependent, stochastic cycle times, while the change-over times are given by a 4-dimensions matrix, characterized by the following product parameters: product family, diameter and stroke length. The sequence of the daily product mix is given as an input for the model.

The goal of the simulation experiments is *to analyse the sensitivity of the output* to the key structural parameters and to the input. These can be divided into two categories, namely structural and operational parameters. The operational parameters are the essential input product sequence of the model. The structural parameters define aspects of how the elements of the simulation model behave. The following were identified as structural parameters: size of buffers, machine availability and workers availability.

In order to execute the sensitivity analysis, the ultimate *outputs* of the simulation model are the completion date of the orders and the total throughput time for the daily production executed in 3 shifts.

4.4. Preliminary simulation experiments

Analysing the structural parameters is a typical application of DES tools and despite the possibly large domain of the variables such experiments can be handled efficiently. The domain of the structural parameters are summarized in Table 1.

	Min	Max	Values
Buffer size	0	∞	6
Human availability	60%	100%	4
Machine availability	85%	100%	11

Table 1. The domain of the structural simulation parameters.

The average runtime of an experiment is below 5 seconds, which means that –based on the defined values of the domain– the structural parameters can be evaluated in ~4000 seconds (assuming 3 experiments with each setting). This is well within the usual time requirement of complex simulation experiments; however, this assumes only one product sequence along every experiment. Unfortunately, the operational parameters have a significant impact on the analysed structural parameters and therefore they cannot be studied separately. Another consequence is that, due to the larger domain of the input sequence (a daily product mix can contain over 40 items) evaluating every combination is no longer feasible in the available time.

Fig. 3. shows the results of a preliminary simulation study, executed on experimentally defined scenarios, using baseline data as input. It can be stated that the same structural parameters resulted in completely different outcome showing it is important to include the effect of the product sequence (compared to *exp1* as baseline) into the sensitivity analysis (denoted as *exp19-exp21* in Fig. 3.).

A possible solution for handling this complexity is to aggregate the input domain into a set of input scenarios, which contain defined settings of both structural and operational parameters.

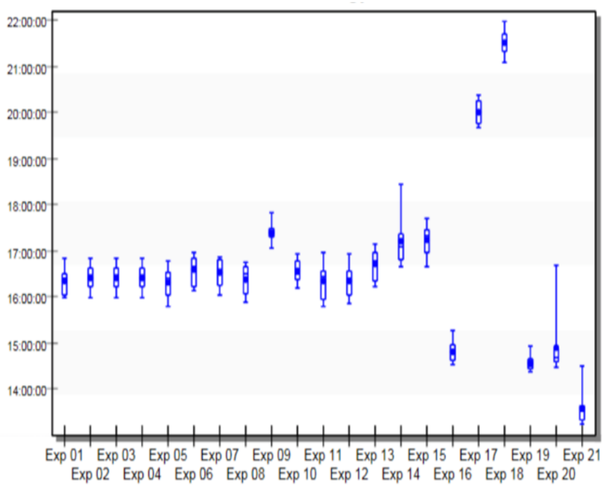


Fig. 3. Experiment settings (horizontal axis) exp16-21 deal with buffer sizes (structural parameters) and sequences of jobs, showing significant relevance on output (makespan, vertical axis, in hours).

4.5. Reducing the domain size of the input mix

The first step in handling the complexity is to reduce the domain size defined by the daily input mixes. A daily input mix is defined by the number of each product type planned to be in production on a given day. Thus, it can be formalized as a feature vector where each product type is described by its planned daily amount. Collecting data from a one-year time frame resulted in data for 243 workdays, where 39 different types of products were in production. It is an important assumption that this period is considered as representative for analyzing the behavior of the system and, therefore, it is set as the baseline of the analysis. The p=39 different products define the length of the feature vector, with n=243 observations.

The aim of the reduction here is to define a set of representative input mixes with a smaller cardinality than n, which can provide a similar behavior of the system. The system's behavior is evaluated by using the simulation model and comparing the following KPIs at each experiment:

- The average net lead-time (LT) of products.
- The total makespan (MS) required to finish the production of the input-mix.

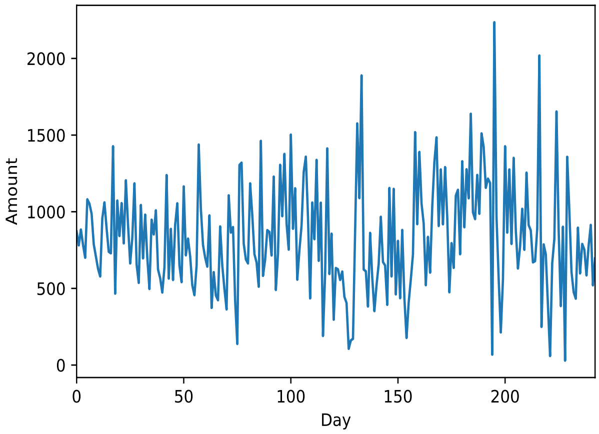


Fig. 4. Fluctuating daily workload

In order to create the reduced cardinality sets of input mixes the baseline data was analyzed using unsupervised learning, specifically clustering techniques. In this approach the only features or descriptors applied during the analysis were the above-mentioned daily workload of each product type. Therefore, the same feature vectors were used for measuring the dissimilarity between the daily input mixes. In order to eliminate the distortion caused by the fluctuation of the daily workload the baseline data is normalized at each day (Fig. 4 shows the fluctuation of the total daily workload).

On the normalized baseline data, the cluster analysis is performed by *hierarchical clustering* as –by using *dendrogram*– it provides an adequate visualization of the dissimilarity even with a considerable number of features without specifying the number of clusters beforehand. The *dendrogram* is created by applying the average or UPGMA algorithm. Fig. 5 shows the results of the hierarchical clustering. Note that–because of the normalization–the min (0) and max (1) values denote the largest and smallest workload calculated for each day, respectively.

Using the *maxclust* criteria the observations (i.e. the daily input mixes) are grouped into a set of clusters, with the following (decreasing) cardinality: 100, 50, 20, 10, 5.

This means that the original *n* observations were represented by 100, 50, 20, 10, 5 cluster centroids each of which is obtained by the geometric mean of the observations assigned to each cluster. An example result of the clustering, where the applied number of clusters is 10, is shown in Fig. 5.

Fig. 6. illustrates how the observations are assigned to clusters in each case. Note, that for lower number of clusters the majority of observations lie in one cluster. Compared this with Fig 5, it can be concluded that these major clusters represent days where a single product rules the majority of the daily workload, while days with more distributed workload form smaller clusters. The higher the number of clusters the more distributed they become over the daily observations.

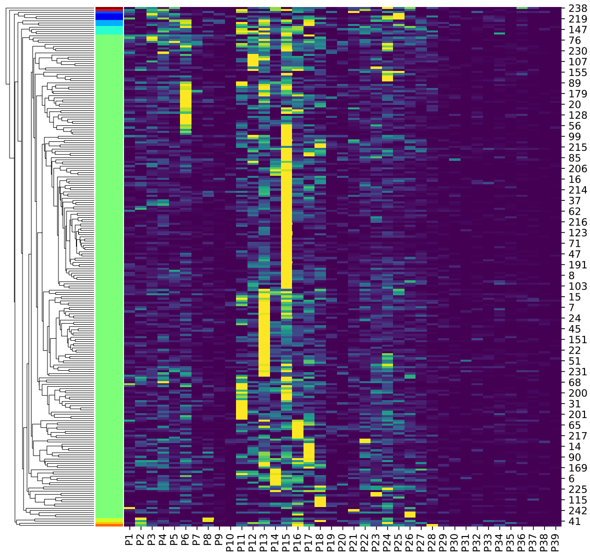


Fig. 5. Clustering the daily demands by quantity for SKU-s

4.6. Discussion of the simulation results

In order to evaluate the results of the clustering the centroids, which are using the normalized data, were multiplied by a constant representing the planned average daily workload, which is given by the company. The newly calculated daily input mixes then were evaluated by using the simulation model and compared to the baseline data. The simulation experiments are carried out with two settings. In the first run the daily workload of each product is handled as a single batch, while in the second run a simple lot splitting rule was applied, which forms lots with a maximal size of 49 (a value chosen as a best practice by the company).

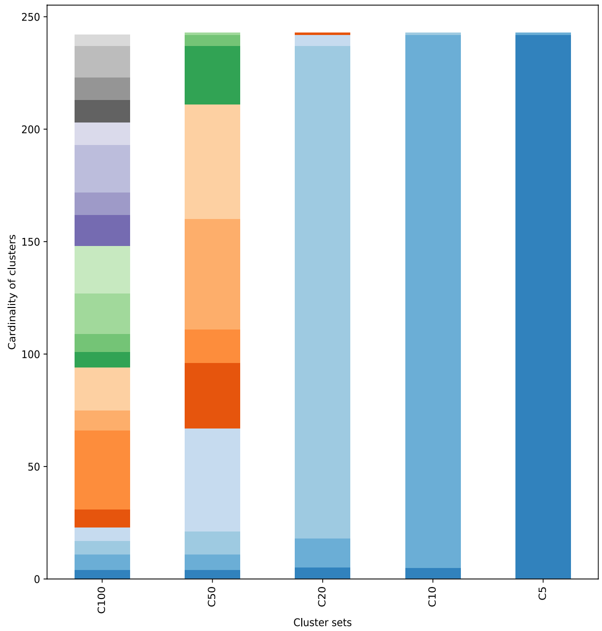


Fig. 6. The distribution of cluster labels for the 5 cluster sets

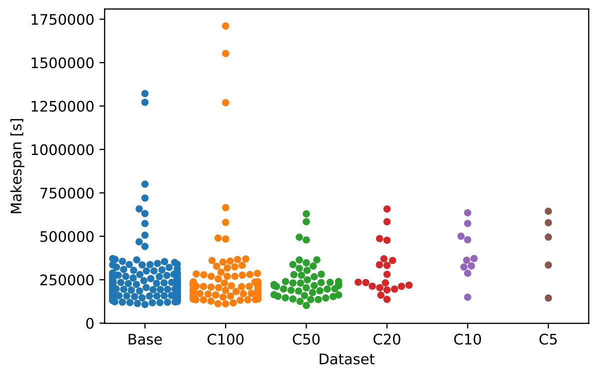


Fig. 7. Simulated makespan values for the baseline data and the five datasets created by clustering without lot-size control.

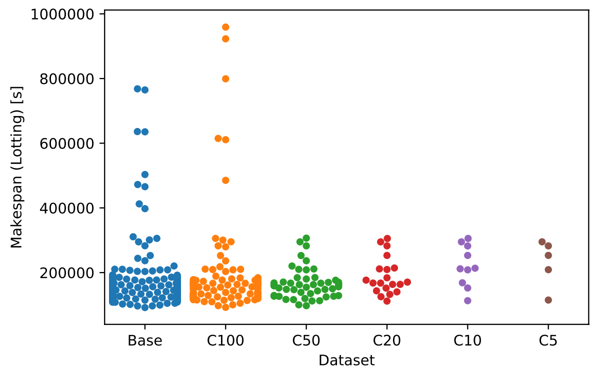


Fig. 8. Simulated makespan values for the baseline data and the five datasets created by clustering with lot-size control.

Fig. 7. shows the total makespan of daily input mixes without lot-size control, and Fig. 8. shows the total makespan with lot-size control. The figures display the results by using *swarmplots*, which are able to show cardinality of the clusters and the distribution of the values as well. Note, that the total makespan decreases when lot-size control is applied. This is probably due to the fact that large-sized lots can easily block resources for a long period of time, therefore causing low resource utilization. It can be concluded that even the low cardinality clusters represent the spread of the baseline makespan data well for the majority of the observations. However, it is also visible that the most extreme values do not appear in clusters where the domain size reduction is in the order of magnitude (C50, C20, C10, C5). This phenomenon is even stronger when lot size control is applied.

Fig. 9. shows the average lead-time for products in daily input mixes without lot-size control, and Fig. 10. shows the average lead-time with lot-size control. In these cases, the approach appears to perform better, as–without lot-size control–the spread hardly shrinks until dataset C50 and only dataset C5 shows significant shrink. When lot-size control is applied the results are even more consistent until dataset C5.

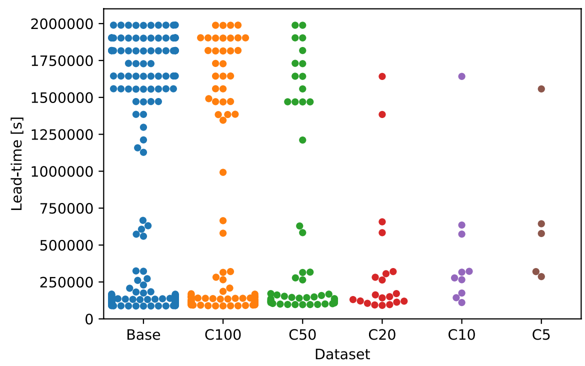


Fig. 9. Simulated average lead-time values for the baseline data and the five datasets created by clustering without lot-size control.

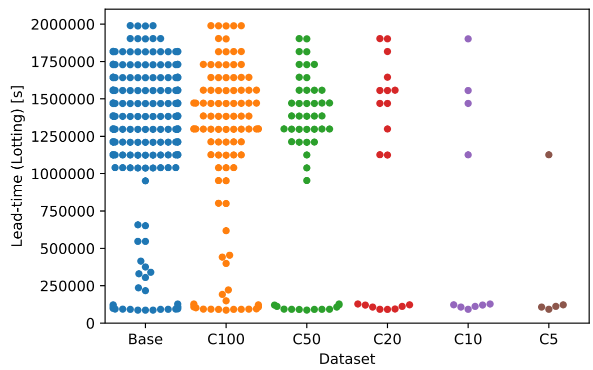


Fig. 10. Simulated average lead-time values for the baseline data and the five datasets created by clustering with lot-size control.

5. SUMMARY AND OUTLOOK

The paper introduced a method based on advanced data analysis for defining typical input scenarios, aiming to reduce the computational complexity of Discrete Event Simulation (DES) analysis.

The presented approach was tested in a real-life combined (manufacturing and assembly) production line of a high-mix low volume environment. The results showed that using scenarios for representing the typical input sets allowed significantly reducing the number of experiments required to execute sensitivity analysis of both the structural (e.g. buffer size or workforce) and the operational (i.e. sequencing) parameters of the line.

A method was introduced, as a possible solution for handling this complexity, in order to aggregate the input domain into a set of input scenarios, which contain defined settings of both structural and operational parameters. Thus, reduced cardinality sets were provided by clustering techniques on the input mixes of the baseline data, formulating a set of representative input mixes with a smaller cardinality.

In order to evaluate the results of the clustering, the newly calculated daily input mixes were evaluated by using the simulation model and compared to the baseline data.

It can be stated that by applying the proposed methods even the low cardinality clusters represent the spread of the baseline makespan and lead-time data well for the majority of the observations. However, it is also visible that the most extreme values do not appear in clusters where the domain size reduction is in the order of magnitude

As an outlook, the research work presented in the paper had two distinct goals. On the one hand, to validate, if the methods are applicable for the reduction of the number of scenarios needed for a comprehensive simulation analysis on system-sensitivity. On the other hand, categorization of the main input factor (*Demand*) could result in assigning situation-related sequencing and lot-sizing rules to a finite number of demand patterns. This would support the planner in creating production schedules/sequences by applying a formalized method, considering the reduction of lead-time variance, as a primary target.

The proposed solution is intended to be extended by more comprehensive analysis on applying and comparing different clustering methods, as well as introducing new dissimilarity measures for the clustering algorithms. Moreover, a set of new experiments on selecting other production related KPIs would be necessary, by applying the new data available from the clustering.

REFERENCES

Bagchi, S.; Chen-Ritzo, C.; Shikalgar, S.T.; Toner, M.: A full-factory simulator as a daily decision-support tool for 300mm wafer fabrication productivity. In: Proc. of the 2008 Winter Simulation Conference. 2009, pp. 2021-2029.

Banks, J., 1998, Handbook of Simulation, Principles, Methodology, Advances, Application and Practice. John Wiley & Sons Inc 1998.

Jahangirian, M., Eldabi, T., Naseer, A., Stergioulas, L.K. & Young, T., 2010, Simulation in manufacturing and business: A review, European Journal of Operational Research, vol. 203, no. 1, pp. 1-13.

Law, A.; Kelton, D., 2015, Simulation modeling and analysis. New York, McGraw-Hill, 5th Ed. 2015.

Longo, F., 2013, "On the short period production planning in industrial plants: A real case study", International Journal of Simulation and Process Modelling, vol. 8, no. 1, pp. 17-28.

Negahban, A. & Smith, J.S., 2014, "Simulation for manufacturing system design and operation: Literature review and analysis", Journal of Manufacturing Systems, vol. 33, no. 2, pp. 241-261.

O'Reilly, J.J.; Lilegdon, W.R., 1999, Introduction to FACTOR/AIM. In: Proceedings of the 1999 Winter Simulation Conference, 1999, pp. 201-207.

Pfeiffer, A.; Gyulai, D.; Kadar, B.; Monostori, L., 2016, Manufacturing lead-time estimation with the combination of simulation and statistical learning methods. PROCEDIA CIRP vol. 41., 2016, pp. 75-80.

Rabelo, L.; Helal, M.; Jones, A.; Min, J.; Son, Y.J.; Deshmukh, A.: A hybrid approach to manufacturing enterprise simulation. In: Proc. of the 2003 Winter Simulation Conference, 2003, pp. 1125-1133.

THE ROLE OF DIGITAL TWIN IN A CYBER-PHYSICAL PRODUCTION ENVIRONMENT WITH PRESCRIPTIVE LEARNING

János Jósvai[(a)], András Pfeiffer[(b)], Norbert Szántó[(c)], Gergő Monek[(d)]

[(a),(c),(d)]Széchenyi István University, Department of Automobile Production Technology, Győr (Hungary)
[(c)] Institute of Computer Science and Control, Hungarian Academy of Sciences, Budapest (Hungary)

[(a)]josvai@sze.hu, [(c)]szanto@sze.hu, [(d)]monek.gergo@sze.hu
[(d)]pfeiffer.andras@sztaki.mta.hu

ABSTRACT
In this paper, we introduce the establishing of digital twins for cyber-physical production systems. Using industrial PLC network and a discrete event simulation tool a live and interactive connection is made between the physical and cyber world. Based on this achievement a modified proactive framework with prescriptive simulation will be presented.

Keywords: cyber-physical production system, digital twin, proactive and prospective planning

1. INTRODUCTION
In the last years Industry 4.0 has got an increasing role within production environments. It is based on cyber-physical systems (CPS), where the physical and cyber (virtual) elements are in close interaction with each other. In such environments, the data collection and analysis is made in a networked structure. In production environments, cyber-physical production systems (CPPS) are evolving. This progress is based on three main science fields: computer science (CS), info communication technologies (ICT) and manufacturing science technologies (Monostori et al. 2016)(Schleich et al. 2017)(Enke et al. 2018).

Cyber-physical production systems are horizontally and vertically integrated environments, where the planning processes of products, technology, production and logistical system are digitally interconnected and permeable (Ochs and Riemann 2018)(Vogel-Heuser, Bauernhansl, and Ten Hompel 2017)(Ustundag and Cevikcan 2018, 4).

One of the main features of cyber-physical production systems is the Digital Twin model. These virtual models are mirroring the real world into the cyber space with the appropriate level of detail (Deloitte 2017)(Negri, Fumagalli, and Macchi 2017)(Post, Groen, and Klaseboer 2017).

Through the digitalisation of production, the Digital Twins can have a live connection to the real world, bringing its physical state changes directly into the virtual world. This enables a reactive or proactive planning using simulation as well (Dr. Pfeiffer 2017)(Stark, Kind, and Neumeyer 2017)(Uhlemann et al. 2017).

Figure 1. presents a simulation-based prediction framework. In Figure 1, Plant represents the underlying production system, which is generally controlled through the MES. Thus, green arrows represent production related data provided by the plant (e.g., resource status, job completion, or, the performance measure of interest KPI in the current case), either gathered by the MES and stored as log data, or, monitored on-line by, i.e. the simulation framework. Contrary, grey arrows represent an interaction or information exchange, e.g., the Decision-maker might control the process of the production (highlighted as Reaction) of the plant by the MES system.

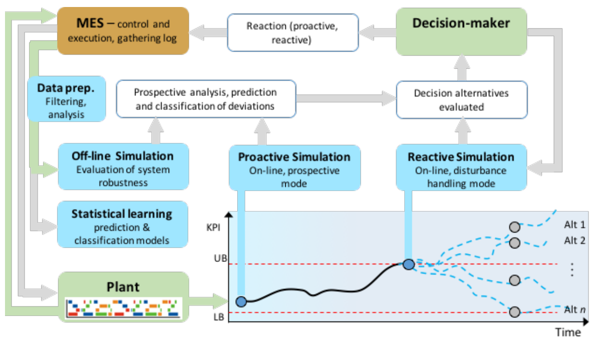


Figure 1: Simulation-based prediction framework (Dr. Pfeiffer 2017)

In a real-world application, three main distinct operation modes follow each other during operation.

- Off-line validation, sensitivity analysis and statistical modelling of the system. Evaluation of the robustness of the system against uncertainties (e.g., different control settings, thresholds and system load levels). Consequently, this scenario analysis might point out the resources or settings which can endanger the normal operation conditions. In Figure 1 off-line simulation represents the comprehensive model of the plant.

- On-line, anticipatory recognition of deviations from the planned operation conditions by running the simulation parallel to the plant activities; and by using a look ahead function, support of situation recognition (proactive operation mode, Fig. 1).
- On-line analysis of the possible actions and minimization of the losses after a disturbance already occurred (reactive operation mode, Fig. 1), e.g., what-if scenario analysis.

In this paper, we would like to present how a real-world application and a simulation-based predictive and prescriptive framework can be connected.

2. DECRIPTION OF THE CYBER PHYSICAL SYSTEM

The laboratory of the Department of Automobile Production Technologies at the Széchenyi István University has a FESTO Didactic System with modular elements. Figure 2. represents the 3D digital model examples of those elements.

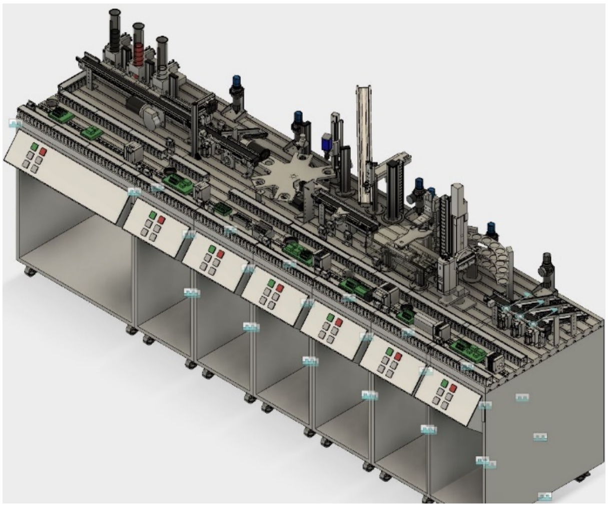


Figure 2: FESTO Didactic System 3D digital model

The modular FESTO Didactic line components of this system are controlled by Siemens S7-300 PLCs. These are linked through a switch and then a main central control is delivered by an S7-1500.

The physical connection between the computer and the system is provided by the S7-1500 PLC. The data communication between machine-to-machine is realized by the currently available most modern communication protocol OPC UA. As the OPC UA server S7-1500 PLC provides the central control.

Two example system components are shown by Figure 3. The first one is the starting station; this unit delivers the basic parts for the small production process. The delivery sequence can be predefined using the systems own control mechanism, or through the later described network architecture. The second example on the same figure is a machining station with a rotation table for different operations. It is planned to change the operation sequence depending on the parameters of the arriving part. The goal is to achieve a lot size 1 automated production environment without interruptions.

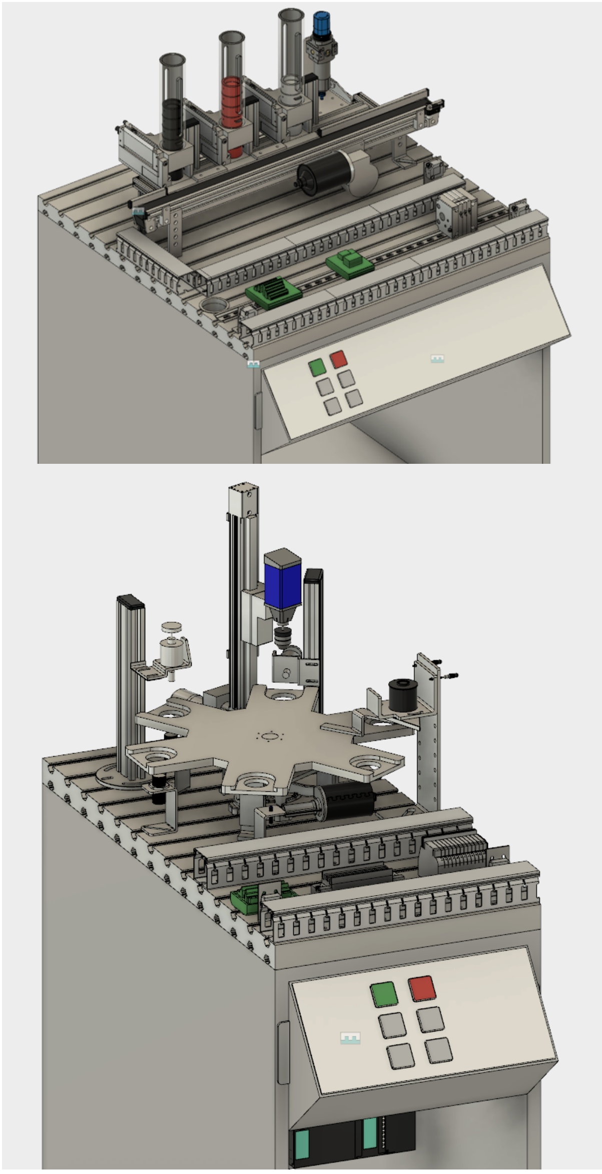


Figure 3: FESTO Didactic System example elements

The applied discrete event simulation (DES) tool is Tecnomatix Plant Simuation, which supports OPC UA communication, this software can collect and process data as a client.

Figure 4. Illustrates the architecture of the network and the connection with Plant Simulation.

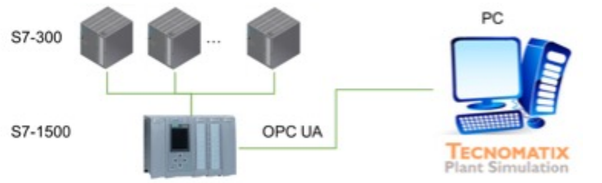


Figure 4: Network architecture using Plant Simulation

Originally the control and functionality of this mini factory is driven by a specialized software tool. In our work, we built up a live connection and communication with the simulation environment. Using the 3D digital models and the live communication a living digital twin has been established in DES environment.

There are several ways to establish communication among the machine-to-machine elements of the desired network architecture. The objects in our scope are Siemens products, therefore, the manufacturer's recommended solutions are considered, which are also used in industrial environments. These possible solutions are the following:

- Simit: Simit is a required interface at hardware and software level in a networked architecture. (PLC-Simit-PS)
- SimaticNET: it connects physically the tools with Ethernet functionality, and it also provides OPC support. The PLC-PC (PS, TIAPortal, SimaticNET) connection is created by the OPC server which is installed on the computer. It can be created using the support of softwares: TIAPortal and SimaticNET.
- OPC UA: the S7-1500 PLC 2.0 or newer firmware update provides the UPC-UA use on itself, furthermore version 14 of Plant Simulation supports OPC UA with a dedicated object. This way a direct contact between the PLC and PS using UPC UA communication protocol is provided.

The OPC Unified Architecture was born out of the desire to create a true replacement for all existing COM (Component Object Model) -based specifications without losing any features or performance. Additionally, it must cover all requirements for platform-independent system interfaces with rich and extensible modelling capabilities being able to describe also complex systems. The wide range of applications where OPC is used requires scalability from embedded systems across SCADA and DCS up to MES and ERP systems (Mahnke, Leitner, and Damm 2009).

OPC is used as system interface today; therefore, the reliability for the communication between distributed systems is very important. Since network communication is not reliable by definition, robustness and fault-tolerance are the important requirements, including redundancy for high availability. Platform-independence and scalability is necessary to be able to integrate OPC interfaces directly into the systems running on many different platforms (Mahnke, Leitner, and Damm 2009).

In the case of an active connection, the operating actual state of the PLC and its changes are continuously visible at the signal level in the simulation environment (Figure 5). By using sensor signals, the actual process can be mapped and modelled in real time, which is one of the pillars of creating a living digital twin. The simulation environment is then able to collect detailed real-time data to analyse system and process behaviour.

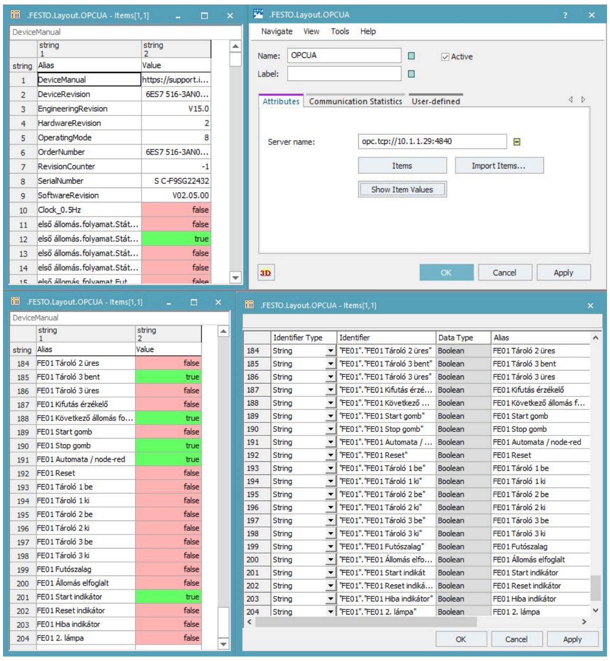


Figure 5: Mapped and modelled sensor signals for the Digital Twin

The OPC UA communication capability not only allows the reading of PLC signals but has also the capability of writing them. This architecture allows higher level of configuration and management, more complex logic can be applied than on the original PLC level.

The digital model of the physical system in the simulation environment is completed, the Digital Twin is able to monitor the individual workflows in real-time and collect the performance data (Figure 6).

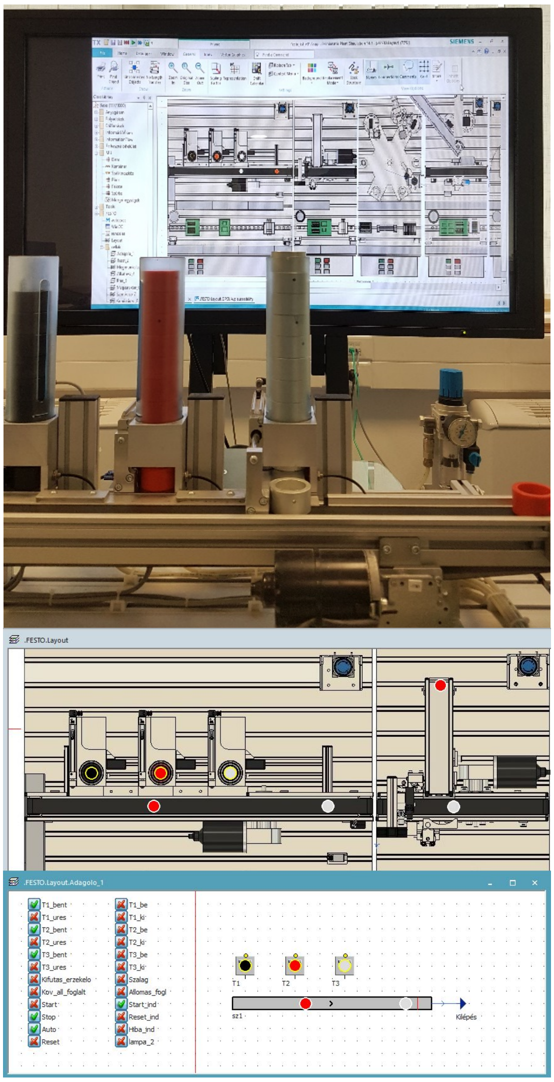


Figure 6: Physical system and its real-time Digital Twin

3. PROACTIVE PLANNING AND PRESCRIPTIVE SIMULATION WITH THE HELP OF DIRECT LINKED DIGITAL TWIN

The simulation based predictive framework introduced on Figure 1 can be extended with the functionality of prescriptive simulation (Figure 7.). The live and interactive connection between the physical and cyber counterparts enables data gathering, analysing and system configuration. Based on collected data, on statistical learning and on the ability to simulate different scenarios prescriptive alternatives can be evaluated. The results serve as basics for the decision making by classifying reactions on possible scenarios and deviations of the system under investigation.

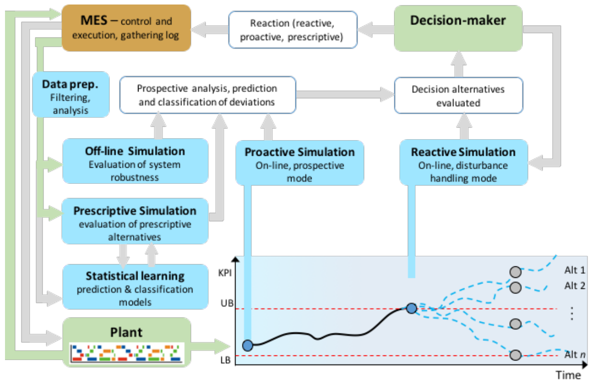


Figure 7: Proactive planning and prescriptive simulation

The simulation using the Digital Twin model of the system under investigation can provide prescriptive alternatives for different possible deviations or parametrizations. This can be achieved through statistical and machine learning functionalities. For this purpose, a separate but identical Digital Twin is used, which has no real-time connection with its physical counterpart, but is able to use control and log data of the MES system. Through experiment planning and several simulation-runs prescriptive alternatives are evaluated and stored. The cognitive decision making – which can also be a specialized Digital Twin model – can evaluate the current status of the plant or production process. This analysing step is able to use the prescriptive alternatives as the base of similar cases or historical data. After decision making the configuration is able to react back on the system. With the earlier introduced network architecture, a direct PLC field level configuration is possible as well.

4. SUMMARY AND OUTLOOK

In this paper, we introduced the establishing of digital twins for cyber-physical production systems. Using industrial PLC network and a discrete event simulation tool a live and interactive connection was established between the physical and cyber world. Based on this achievement a modified proactive framework with prescriptive simulation was presented.

In our future work, we are working on standardisation for automatic model building, and on machine learning capabilities using the above introduced framework. Furthermore, we are investigating the signal transmission and network communication speed, and its effects on the real-time digital twin.

REFERENCES

Deloitte. 2017. ‘Industry 4.0 and the Digital Twin’. Deloitte University Press.

Dr. Pfeiffer, András. 2017. ‘Improving the Accuracy of Cycle Time Estimation for Simulation in Volatile Manufacturing Execution Environments’. ASIM 2017.

Enke, Judith, Rupert Glass, Antonio Kreß, Jens Hambach, Michael Tisch, and Joachim Metternich. 2018. ‘Industrie 4.0 – Competencies for a Modern Production System’. Procedia Manufacturing 23:

267–72. https://doi.org/10.1016/j.promfg.2018.04.028.

Mahnke, Wolfgang, Stefan-Helmut Leitner, and Matthias Damm. 2009. OPC Unified Architecture. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-68899-0.

Monostori, L., B. Kádár, T. Bauernhansl, S. Kondoh, S. Kumara, G. Reinhart, O. Sauer, G. Schuh, W. Sihn, and K. Ueda. 2016. 'Cyber-Physical Systems in Manufacturing'. CIRP Annals 65 (2): 621–41. https://doi.org/10.1016/j.cirp.2016.06.005.

Negri, Elisa, Luca Fumagalli, and Marco Macchi. 2017. 'A Review of the Roles of Digital Twin in CPS-Based Production Systems'. Procedia Manufacturing 11: 939–48. https://doi.org/10.1016/j.promfg.2017.07.198.

Ochs, Th., and U. Riemann. 2018. 'Smart Manufacturing in the Internet of Things Era'. In Internet of Things and Big Data Analytics Toward Next-Generation Intelligence, edited by Nilanjan Dey, Aboul Ella Hassanien, Chintan Bhatt, Amira S. Ashour, and Suresh Chandra Satapathy, 30:199–217. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-60435-0_8.

Post, Jan, Manso Groen, and Gerrit Klaseboer. 2017. 'PHYSICAL MODEL BASED DIGITAL TWINS IN MANUFACTURING PROCESSES'. Forming Technology Forum 2017. https://www.researchgate.net/publication/320445237.

Schleich, Benjamin, Nabil Anwer, Luc Mathieu, and Sandro Wartzack. 2017. 'Shaping the Digital Twin for Design and Production Engineering'. CIRP Annals 66 (1): 141–44. https://doi.org/10.1016/j.cirp.2017.04.040.

Stark, Rainer, Simon Kind, and Sebastian Neumeyer. 2017. 'Innovations in Digital Modelling for next Generation Manufacturing System Design'. CIRP Annals 66 (1): 169–72. https://doi.org/10.1016/j.cirp.2017.04.045.

Uhlemann, Thomas H.-J., Christoph Schock, Christian Lehmann, Stefan Freiberger, and Rolf Steinhilper. 2017. 'The Digital Twin: Demonstrating the Potential of Real Time Data Acquisition in Production Systems'. Procedia Manufacturing 9: 113–20. https://doi.org/10.1016/j.promfg.2017.04.043.

Ustundag, Alp, and Emre Cevikcan. 2018. Industry 4.0: Managing The Digital Transformation. Springer Series in Advanced Manufacturing. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-57870-5.

Vogel-Heuser, Birgit, Thomas Bauernhansl, and Michael Ten Hompel, eds. 2017. Handbuch Industrie 4.0. Bd. 4: Allgemeine Grundlagen. 2. Auflage. Springer Reference Technik. Berlin: Springer Vieweg.

AUTHORS BIOGRAPHY

JÁNOS JÓSVAI is associate professor and head of the Department of Automobile Production Technology at the Széchenyi István University. He studied at the Budapest University of Technology and Economics and at the Karlsruhe Institute of Technology. He holds a PhD from the Széchenyi István University. He is a member of the ASIM group Simulation in Production and Logistics (SPL) and of the ASIM SPL workgroup on the Investigation of Energy-related Influences in SPL since its formation. His research interests include resource efficient production, Digital Factory and Cyber-Physical Production Systems. His email address is josvai@ga.sze.hu.

NORBERT SZÁNTÓ is assistant lecturer of the Department of Automobile Production Technology at the Széchenyi István University, Győr. Education: Phd Student – currently, Mechatronics engineering MSc, Mechanical engineering BSc. Participating in the education and also industrial research-development projects. Research area is the production planning, process and logistics simulation.

GERGŐ MONEK is assistant lecturer of the Department of Automobile Production Technology at the Széchenyi István University, Győr. Education: Phd Student – currently, Logistics engineering MSc, Mechanical engineering BSc. Participating in the education and also industrial research-development projects. Research area is the production planning, process and logistics simulation, quality management.

ANDRÁS PFEIFFER, earned his PhD in 2008 at the Budapest University of Technology and Economics. Currently he is a senior research fellow and deputy director at the Computer and Automation Research Institute, Hungarian Academy of Sciences (SZTAKI). His current interest includes decision support in production planning and control, as well as the simulation and emulation modeling of complex production systems, self-building simulation systems and digital twins.

ASSEMBLY LINE BALANCING TECHNIQUES: LITERATURE REVIEW OF DETERMINISTIC AND STOCHASTIC METHODOLOGIES

Giuseppina Belfiore [(a)], Domenico Falcone [(b)], Luca Silvestri [(c)]

(a),(b) Department of Civil and Mechanical Engineering
University of Cassino and Southern Lazio
(c) Niccolò Cusano University
00166 Roma (RM) – Italy

[(a)]g.belfiore@unicas.it, [(b)]falcone@unicas.it, [(c)]luca.silvestri@unicusano.it

ABSTRACT

A production line consists of a sequence of work stations where the operations, necessary to create a product unit, are carried out. Therefore, the design of a production line is linked to the search of the smallest number of stations to be activated in order to obtain advantages in terms of costs, time and resources. The literature review carried out in the following research was organized according to the nature of task times (deterministic and stochastic), concerning to heuristic and metaheuristic techniques.

Keywords: work station, cycle time, stochastic task times, deterministic task times

1.INTRODUCTION

The layout of a production system represents the disposition of a set of vehicles, operators and equipment within a production area, depending on products or services to be produced. The optimal layout allows to simplify the production process, minimizing the cost of transport of materials, production stocks and stored materials. So, a plant, from the point of view of the layout, can be divided into:

- line layout or for product;
- layout for units or for process;
- fixed position.

When few products are realized in large quantities, we refer to mass production systems that adopt a line layout. In this case, the production is to "stock" and the machines are arranged respecting the sequence of processing that the raw material must undergo. These systems are characterized by low flexibility, and by high fixed costs for investments in specialized plants. Instead, production systems characterized by a fixed layout and process layout are used when many products are produced in small quantities. The production is made "on order," the mix of products is extremely variable, and the different processing sequences are defined according to the material to be processed and the product to be obtained. The design of a fixed layout favors the production of large products, where the product remains stationary at one point while the resources must find an appropriate position.

In this paper, line production was taken into consideration. The inputs for the design of the assembly line system are as listed below:

- Precedence network of tasks;
- Task times;
- Cycle time or number of workstations.

Priority relationships between tasks are assigned by theprecedence network, and each activity requires a task time toexecute. The cycle time is defined as the greater between the times of crossing of a product within a station of the line and is given by the following relation:

$$TC = \sum productiontimes + \sum transfertimes$$

To evaluate the efficiency of the balance, the following formula is provided:

$$B.E. = \frac{Sumofalltasktimes}{Numberofworkstation * Cycletime} \times 100$$

So, the efficiency is given by the relationship between the sum of the task times and the total time provided to perform all the activities.

In this work, a literature review on the production line balancing problem was carried out, focusing on the nature of the task times, namely the time needed to complete a given operation.

This interval, in the production lines can be classified in deterministic (known and constant over time) and stochastic (variable due to random factors).

Seven methodologies have been identified to balance the production lines, four with deterministic task times and three with stochastic task times. In the paper, the various methodologies were presented. At the end a brief discussion of the research is present.

1. LINE PRODUCTION

Line production is characterized by a sequence of work stations in which the operations, necessary to realize a unit of product, are carried out. Moreover, the time in which a station is being processed during a cycle is called workload. The line production characterizes companies that produce few products in large

quantities, requiring high investments in specialized plants. These companies, usually, carry out all the production process based on demand forecasting, so the production is to "stock". The machines, generally dedicated, are arranged respecting the processing that the raw material must undergo. The workpiece is transferred from one station to another, through a rigid movement system.
Therefore, the line is called at cadency if the material handling takes place at fixed intervals (Fig.1); if instead, the machines are decoupled from buffers, the line is called an asynchronous transfer (Fig.2).

Figure 1: Line at cadence

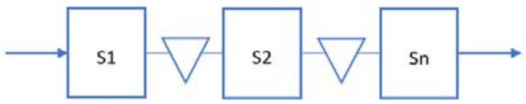


Figure 2: Asynchronous transfer line

Production planning and control problems are reduced to the sizing and sequencing of lots, so that the assignment of the operations must generate a balanced workload between the stations along the line, in order to avoid bottlenecks. The layout of line production is characterized by advantages and disadvantages. The strengths are: reduced labor demand, low WIP, high saturation of the machines/operators, uniformity of the quality characteristics of the products. Among the disadvantages we find: low flexibility, high investment, risk of rapid obsolescence, high vulnerability to breakdowns. Furthermore, a production line can be classified according to the number of configurations of the same product, for which we have: dedicated lines, multi-model lines and mixed-model lines. While the most common lines are: two-sided lines, serial lines, U-lines, parallel lines.

3. BALANCING OF PRODUCTION LINES

Balancing a production line means minimize the number of work stations, reducing the maximum time of the work stations. We must consider constraints such as:

- cycle time;
- space;
- allocation of operations;
- incompatibility between operations;
- management of material flows;
- precedence relations;
- number of operators requested in the same workstation.

Balancing of workstations influences the productivity of assembly lines, for this reason it is important to address this problem before the design of the line, or in the case of an existing line to improve performance. The balancing objectives are divided into technical and economic goals.
The technical objectives concern:

- the minimization of the number of stations, fixed the cycle time. In this case cycle time is determined by the production volume to reach;
- the minimization of cycle time, fixed the number of stations. To increase productivity, the cycle time must be low;
- the minimization of no-saturation of a machine/operator, it is the rate of time in which no activity is carried out and represents a damage in terms of productivity for the company;
- the minimization of the probability of exceeding the cycle time.

Regarding economic objectives, it is important:

- to minimize the expected cost. The total expected cost is the sum of the line cost (cost of labor and the cost of setting up the stations) and the expected cost of non-completion;
- to minimize the costs of duplicating stations;
- to maximize profit. The profit function is expressed by:

$$p(n, TC) = \frac{f_1}{TC} - \sum_{k=1}^{n} f_2(k)$$

where:

- f_1 is the contribution margin for unit of product;
- $f_2(k)$ is the fixed cost for unit of time tied to station k;
- n is the number of stations;
- TC is the cycle time.

3.1 Classification of balancing methods

Methods used to balance an assembly line are various, and can be classified according to the number of models realized (single model - multi model), the deterministic or probabilistic nature of the task times and the nature of the flow.A literature review was performed based on the nature of the task times, so the hierarchy of classification of the ALB problems is shown in Fig.3:

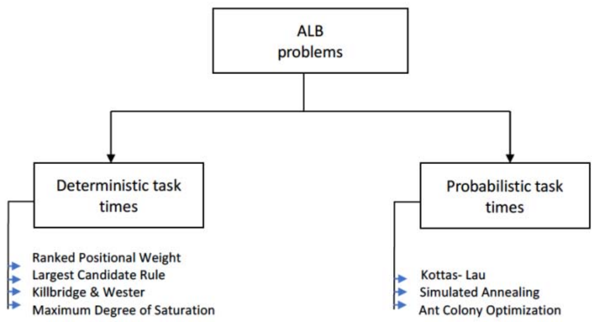


Figure 3: Classification of ALB problem

So, the most common methodologies, analyzed in this paper are:

- Ranked Positional Weight;
- Largest Candidate Rule;
- Kilbridge and Wester;
- Maximum Degree of Saturation;
- Kottas-Lau;
- Simulated Annealing;
- Ant Colony Optimization.

4. METHODS WITH DETERMINISTIC TASK TIME

4.1 Ranked positional weight

The RPW method is a balancing technique of line productions developed by Helgeson and Bernie. This method is based on the "positional weights", that is the coefficients attributed to each operation, so that the operation is assigned to the stations on the line.
The positional weight of an activity is given by the sum of its task time and of all the operations that follow it on the precedence diagram. So:

$$RPW_I = \sum_{K=1}^{n} t_k$$

where:

- i: the i-th operation;
- n: total number of operations;
- t_k: the execution time of each operation.

After calculating the RPW, the operations are entered in a list in descending order according to their weight. Then the activities with the highest positional weight are assigned to the open stations, respecting the maximum cycle time that it can support. If the cycle time is exceeded, a new station opens.
Advantages: thanks to this methodology, most of the cycle time of each workstation is exploited, obtaining the lowest number of stations and minimizing inactivity.
Disadvantages: the methodology provides a single solution that isn't the best solution to the problem.

4.2 Largest candidate rule

The LCR method is a simple application used to solve uncomplicated balancing problems, which considers the task time as the only parameter in order to assign the task to the station
In LCR a table is constructed, which is characterized by operations, their task times and predecessors. The tasks are ordered according to the major task time and the operations are assigned in the order obtained, respecting the precedence relationships and the maximum cycle time that the station supports. This procedure is repeated until there aren't operations available.
Advantages: simple to apply. In particular, the methodology uses only one parameter as a reference to assign the tasks to the stations.
Disadvantages: low efficiency in the search for the optimal solution, with a very high balancing delay coefficient.

4.3 Kilbridge & Wester method

The KWM methodology consists in selecting the operations to be assigned to the stations, considering their position in the precedence diagram. So, the activities are positioned within columns, and sorted in a table according to the column they belongs and in descending order of the task time. After completing the table, the tasks are assigned starting from the first column until the maximum cycle time is reached. If the maximum allowed cycle time is exceeded and unassigned operations remain, a new station opens. The application of this method ends when the list of transactions is empty.
Advantages: simple execution method.
Disadvantages: provides a single solution, and the stations have low saturation.

4.4 Maximum degree of saturation

The application of this algorithm results to be quite simple, the method starts by opening a new station, since the set of operations already assigned is empty. Respecting the relationships of precedence, we find a second set consisting of sequences of possible activities. Operations are available when the foregoing activities have already been assigned to a station. At this point, the saturation degree is calculated for each identified combination. This parameter is compared with the degree of saturation imposed α . Combinations of operations that do not satisfy the relationship:

$$GS \leq \alpha$$

are rejected, while the others are taken into consideration.
Advantages: this methodology outputs a single solution that turns out to be optimal, because it aims to minimize the delay coefficient.
Disadvantages: it doesn't determine a unique solution.

5. METHODS WITH STOCHASTIC TASK TIME

5.1 Kottas-Lau method

The Kottas-Lau algorithm aims to optimize total production costs and allows a good level of line balancing to be achieved. This method considers stochastic execution times belonging to a normal distribution with mean value and standard deviation assigned. To describe the functioning of the algorithm, it is necessary to introduce the following parameters: the production capacity of the line, the cycle time and the cost of labor. Moreover, the average duration and the standard deviation of the k-th operation are calculated respectively:

$$m_k = \frac{\sum_{i=1}^{n} T_i}{n}; \sigma_k = \sqrt{\frac{\sum_{i=1}^{n} (T_i - m_k)^2}{n-1}}$$

Moreover, it defines:

- available operation: an operation released from previous ones (already assigned)
- desirable operation: an operation available for which:

$$L_k = P_k * I_k.$$

Where:

- L_k is the average cost of execution on the k-th transaction line:

$$L_k = m_k * \frac{C}{60}$$

that is, the product between the cost of labor for unit of time and the average time of completion of the operation.

- P_k is the percentage of failure to complete the k-th operation.
- I_k is the total off-line completion cost of all operations that cannot be performed if operation k-th has not been completed.
- critical operation: it is a type of operation available for which it results:

$$L_k < P_k * I_k$$

even in "empty" stations, so it can be defined as an undesirable operation.

- Sure operation: an operation for which it results:

$$P_k < 0{,}005$$

that is, the probability of failure to complete the transaction must not exceed 0.5%.
Advantages: very valid in the industrial sector. Greater accuracy of the solution compared to other methods.
Disadvantages: long and complex calculation procedure. It provides a single solution that will be feasible, but not necessarily optimal.

5.2 Simulated Annealing

The simulated annealing methodology starts by generating an initial solution, used as the first current solution. The control parameter T, which indicates the "temperature", is decremented according to a cooling rate CR. As the temperature decreases, solutions adjacent to the current solution are found. The adjacent solution becomes the new current solution when the value of the objective function is greater than the value of the current solution. Otherwise, the inferior solutions can be accepted as current solutions if a certain acceptance criterion is satisfied, to avoid being trapped in a local perspective and achieving the global optimum.

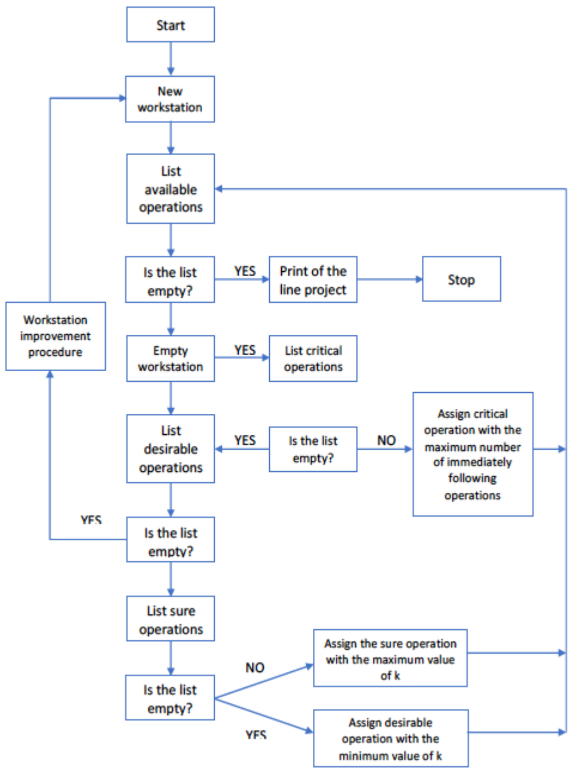


Figure 4:Flowchart of Kottas-Lau algorithm

The function that indicates the probability of accepting a bad solution is given by:

Exp (-(F [candidate solution]-F [current solution])/t)

where F is a function used to evaluate a solution.
At the beginning of the algorithm the value of T is higher and decreases through the cooling function, this involves a long search for the initial optimal solution before intensifying on good areas.
The algorithm stops when the termination criterion is satisfied. This criterion is represented by the number of iterations, the execution time or the final value of the control parameter T.
This methodology, unlike other assembly line techniques in which an operator can occupy a non-parallel work center, addresses the problem of parallelism of tasks in the workstations.
So, if the total time required by the tasks doesn't exceed the specified cycle time, the activities can be placed in a work center.
The logic of this algorithm can be represented by the following flowchart (Fig.5):
where:

- N: current iteration number;
- Nmax: maximum number of iterations;
- Tmin: arrest criterion;

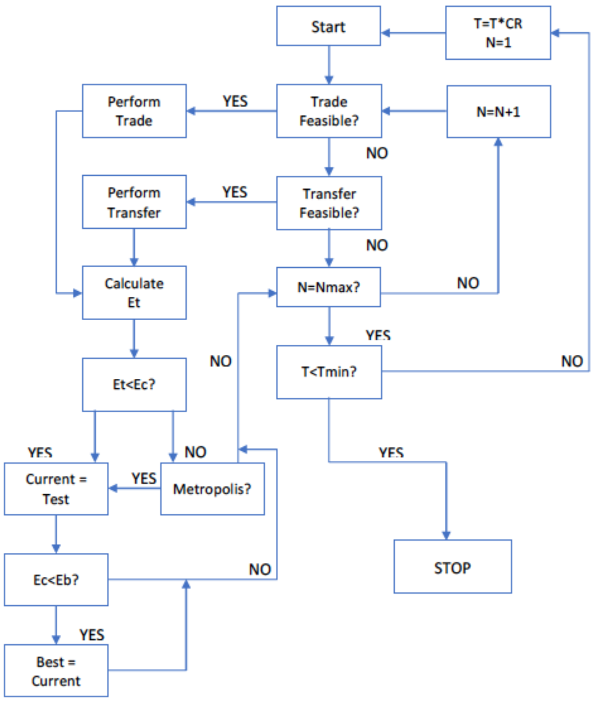


Figure 5: Flowchart of Simulated Annealing algorithm

- E_t: value of the objective function of the tested solution;
- E_c: value of the objective function of the current solution;
- E_b: value of the objective function of the best solution;
- CR: cooling speed.

Advantages: simple to be implemented. “Robust” methodology, in the sense that if left running long, it produces good quality solutions.
Disadvantages: the local search algorithm ends at a local minimum, and there is no information on how much this minimum differs from the global minimum

5.3 Ant Colony Optimization

The Ant Colony Optimization (ACO), introduced in 1992 by Marco Dorigo, is a meta-heuristic approach to solve combinatorial optimization problems. This methodology simulates the natural behavior of ants and develops cooperation mechanisms. As known, ants are insects that live in colonies and have the ability to find the shortest route between their nest and food sources.
Moreover, the ants are blind, but returning to the nest they leave behind the pheromones so that other ants can follow the same path. This process is called “stigmergy”. In this way, ants can follow an already existing path by not moving casually. Therefore, the ACO algorithm simulates the travel of artificial ants, moreover the "ant" selects the activity that must be added to the current work center thanks to a probabilistic mechanism. The probability of selection of the activity is determined by the level of "pheromone" present on the path between the ant and the candidate activity.
This pheromone quantity indicates a measure of the relative desirable solution and the procedure continues until all the activities have been inserted in the work centers.
Below, the flowchart related to the algorithm procedure is shown (Fig.6):

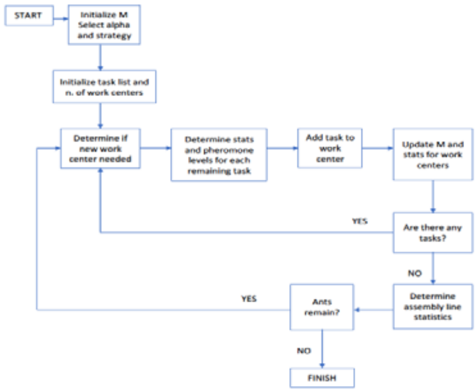


Figure 6: Flowchart of Ant Colony Optimization algorithm

Advantages: ability to build better solutions by exploiting a parameterized probability on a model representing the path of pheromones.
Disadvantages: it doesn’t provide an excellent solution in the finite time in case of optimization problems, or an admissible solution in case of constraint satisfaction problems.

6. CONCLUSIONS

In this literature review, the problem of assembly line balancing has been addressed, which consists in assigning the various operations to the work stations in order to obtain advantages in terms of cost, time and resources.
The study of literature has shown that to balance an assembly line, the most used techniques are heuristics and metaheuristics.
Furthermore, the research has been divided into two categories:

- deterministic problems;
- probabilistic problems.

Heuristic algorithms are able to produce a solution in relatively short time and it is possible to design these techniques for any balancing problem.
Whereas metaheuristics techniques tend to be versatile as they can be applied to different types of problems.
The heuristic techniques considered in this paper to solve assembly line balancing problems, most present in the literature, are the Positional Weight Method, Largest Candidate Rule, Kilbridge & Wester. These methods output a single solution that does not represent the optimal solution for the problem and therefore they are not very efficient to solve the problem of balancing. While the Maximum Saturation Degree algorithm aims to

maximize the saturation degree of each location and minimize the number of stations on the line. This methodology is more efficient than the algorithms mentioned above.
The heuristic method, mostly used in literature, which takes into account the variability of the times, is Kottas-Lau Method, which however has a long calculation procedure as a disadvantage.
While the two metaheuristic techniques reported in the work are Simulated Annealing and Ant Colony Optimization. These techniques are used to solve problems of parallelism between work stations and find optimal solutions to solve the problem of balancing.

REFERENCES

Bagher M., Zandieh M., Farsijani H., 2010. Balancing of stochastic U-type assembly lines: an imperialist competitive algorithm. International Journal of Advanced Manufacturing Technology.54:271-285.

Damiani, L., Giribone, P., Guizzi, G., Revetria, R., Romano, E., 2015 Different approaches for studying interruptible industrial processes: Application of two different simulation techniques (Book Chapter), Handbook of Research on Computational Simulation and Modeling in Engineering. 69-104.

De Felice, F., Petrillo, A., Silvestri, A., Offshoring: Relocation of production processes towards low-cost countries through the project management & process reengineering performance model.Business Process Management Journal. 21(2):379-402.

Di Bona G., Duraccio V., Silvestri A., Forcina A., 2014. Productive line reengineering through simulation techniques. Proceedings of the IASTED International Conference on Modelling, Identification and Control2014, Pages 291-296

Falcone D., De Felice F., 2007. Gestione e progettazione degli impianti industriali. Milano: Hoepli.

Gamberini R., Grassi A., Rimini B., 2006. A new multi-objective heuristic algorithm for solving the stochastic line re-balancing problem. International Journal of Production Economics. 102:226–243

Greco M., Cricelli L., Grimaldi M., 2013. A strategic management framework of tangible and intangible assets", European Management Journal. 31(1):55-66.

Hakansson J., Skoog E., Eriksson K. A review of assembly line balancing and sequencing including line layouts. In Proceeding of PLANs forsknings och ings conference. 69-84.

Hazir O., Dolgui A., 2013. Assembly line balancing under uncertainty: robust optimization models and exact solution method. Computers and Industrial Engineering. 65:261–267.

Kottas J.F., Lau H.S., 1973. A cost oriented approach to stochastic line balancing. AIIE Trans. 5:164–171.

Manna A., Talwar R., Bhushan B., Gupta R., 2012. A review of Ant Colony Optimization. International Journal ESS ISSN: 2249-9482

McMullen P.R., Frazier G.V., 1998. Using simulated annealing to solve a multi-objective assembly line balancing problem with parallel workstations. International Journalof Production Research. 36(10):2717–2741.

McMullen P.R., Tarasewich P., 2003. Using ant techniques to solve the assembly line balancing problem. IIE Transactions. 35:605–617.

Panneerselvam R., 2005. Production and operations management, 2nd edn. New Delhi, PHI Learning.

Petrillo, A., De Felice, F., Zomparelli, F., 2017. Performance measurement for world-class manufacturing: a model for the Italian automotive industry. Total Quality Management and Business Excellence. 1-28.

Petrillo, A., Di Bona, G., Forcina, A., Silvestri, A., 2018.Building excellence through the Agile Reengineering Performance Model (ARPM): A strategic business model for organizations.Business Process Management Journal. 24(1):128-157.

Portiolo Staudacher A., Pozzetti A., 2003. Progettazione dei sistemi produttivi. Milano, Hoepli.

Revetria, R., Guizzi, G., GiriboneP., 2016. Management redesign of a multi-product line through discrete event simulation and response surface methodology. Proceedings of the IASTED International Conference on Modelling, Identification and Control. 1-7.

Scholl A., Becker C. State of the art exact and Heuristic solution procedures for simple assembly line balancing. European Journal of Operations Research. 168(3):666-693.

Sivasankaran P., Shahabudeen P., 2014. Literature review of assembly line balancing problems. International Journal of Advanced Manufacturing Technology. 73:1665-1694.

GUIDELINES FOR RELIABILITY ALLOCATION METHODS

Luca Silvestri[(a)], Domenico Falcone[(b)], Giuseppina Belfiore[(b)]

[(a)] Niccolò Cusano University 00166 Roma (RM) – Italy
[(b)] Department of Civil and Mechanical Engineering, University of Cassino and Southern Lazio - Italy

[(a)]luca.silvestri@unicusano.it, [(b)]falcone@unica.it, [(c)]g.belfiore@unicas.it

ABSTRACT
The purpose of this paper is to review the literature on reliability allocation and to propose a guideline to choice the most suitable allocation method, in respect to desired goals and available resources. The proposed review systematically analyzes allocation methods in literature, determining main features and the area of application. The choice of a reliability allocation method depends on several factors and, moreover, the considered design stage. However, in literature there is a lack of a comprehensive methods' summarization, making it difficult to choose the most suitable strategy. New criteria for the classification are also highlighted. Furthermore, reliability allocation methods become ever more complex and accurate, in particular trough the employment of advanced mathematical tools.
The paper contains a comprehensive analysis of frequently-used allocation methods and a proper summarization through various criteria. Additionally, it points out opportunities for future research.

Keywords: reliability allocation, literature review, reliability guidelines,

1. INTRODUCTION

Reliability allocation is a crucial step for each product development process, it allows to assign level failure rate target to different system units and then to reach the desired reliability goals for the whole system. An efficient allocation strategy leads to the optimal target in agreement with the current technical state of the art, costs, plant layout, and the interaction with the environment.
Each allocation methodology is characterized by the definition of different factors (e.g., system reliability target and its properties, units fault and their relationship with the system fault etc.). In particular, system properties involve technology, operation cycles and times.
Considering the top-down nature of the allocation process, it is necessary to define reliability characteristics for each component from a given system target: mean fault rate, mean life or reliability for an established time period.
System may have its components arranged in series, redundant or in mixed configurations according to the different functional duplications.
An important aspect to take into consideration is the modal design adequacy (Aggarwal 1993, Dai and Wang 1992); it is defined as the probability that a system successfully ends a mission within its design specifications; it is important to investigate the reliability performance of a system in advance (Tucci et al. 2014, Duraccio et al. 2016)
According with the operative time that refers to operation cycles (low cycles and high reliability), it is possible to affirm:

- the reduction of the technology, of the operative time and the increase of the factor of importance lead to the increase of the allocated reliability for each component; and
- the same allocated reliability is expected for units with the same importance, operative time and technology.

The mathematical models behind the allocation procedure are fairly easy if the following hypotheses are respected:

- each unit has independent fault probabilities; and
- the unit state is denoted by binary values (on and off).

The allocation of a reliability requirement to each unit involves the resolution of the following inequality:

$$f(R_1^\circ, R_1^\circ, \ldots, R_N^\circ) \geq R^* \qquad (1)$$

Where R_i° is the allocated reliability to the unit, $R^*$is the system reliability goal and f() is functional relation between the unit and the system itself.
During last decades there was a significant progress in development of several reliability allocation methods, considering available information of the system and subsystems.
The Advisory Group of Reliability of Electronic Equipment (AGREE 1957) introduced a widely used allocation method. This method assigns to the unit or subsystem complexity and criticality a more important role than failure rates. In contrast to AGREE, Aeronautical Radio Inc. (Alven 1964) proposed the ARINC technique, in which failure rates of units or

subsystems are considered as key factors. Bracha (1964) presented a method that uses four factors: state-of-the-art, sublevel complexity as determined by number of parts, environmental conditions and operating time. In order to allocate the reliability to each subsystem, Karmiol (1965) relates the mission objectives to complexity, state-of-the-art, operational profile, and criticality of the system. Then, the engineering design guide, Reliability Design Handbook. (Anderson 1976), includes the feasibility-of-objectives (FOO) technique, which is part of Mil-hdbk-338B handbook (United States Department of Defense 1988), an established standard for military reliability design. This approach provides a detailed reliability allocation procedure for mechanical-electrical systems and Smedley (1992) applied this apportionment strategy for a low energy booster ring magnet power systems in a superconducting supercollider. Kuo (1999) proposed an average weighting allocation method as a guide for commercial applications, while Di Bona et al. (2015) used the Integrated Factors Method (IFM) for reliability and safety allocation for an aerospace prototype project and for industrial applications (Falcone et al., 2007).
More recently, the risk priority number (RPN) (Bowles 2003, Itabashi-Campbell and Yadav 2009) has been used by Kim et al. (2013) and Yadav et al. (2006, 2007, 2014) as a result of qualitative criticality analysis FMEA (Department of the *Army*. TM 5-689-4 2006). To make the reliability study more complex, human error might be considered (Di Pasquale 2015)
All conventional reliability allocation methods are based on the assumption that each considered factor has the same weight and many factors are subjectively evaluated. Considering these weaknesses, Cheng et al. (2006) and Chang et al. (2008, 2009) proposed a reliability allocation weight based on an ordered weighted averaging operator introduced by Yager (1988) (ME-OWA); in addition, Liaw et al. (2011) used different weights for each factors, basing on the DEMATEL decision making method together with ME-OWA.
Although each factor weight may vary depending on the unit in question, all these methods assign the same weight to the same unit or subsystem. This aspect suggests an allocation method able to assign appropriate weights for both factor and unit considered. In particular Di Bona et al. (2016) presented the analytic IFM method (A-IFM), which add to the IFM reliability apportionment the well-known analytic hierarchy process (AHP). This mathematical support allows the reliability allocation to be controlled by appropriately assigned weights to its components. Furthermore, Di Bona et al. proposed the Critical Flow Method, CFM (2015) and Analytic Critical Flow Method, ACFM (2017), as innovative reliability allocation methods well suited for series-parallel configurations. This feature distinguishes these methods from conventional ones.
In literature a general classification of reliability allocation methods is lacking. In particular, there seems to be no general guidelines in order to choose the most suitable method according to the specific application and different phase studying.
Therefore, this paper is a literature review that aims to investigate allocation methodologies in an attempt to classify them with precise and useful criteria.
The organization of this paper is as follows:
Section 2 gives a brief overview of conventional reliability allocation methods, Section 3 presents the chosen classification methodology, finally in Section 4 results and conclusions are analyzed.

2. RELIABILITY ALLOCATION METHODS

This section presents an analysis of principal procedures and techniques of reliability allocation. According to the specific phase or complexity of the system in studying, it is also possible to use simultaneously different methodologies. Let $R^*(t)$ be the reliability goal for a general series system that operates for a time $t > 0$. Let $R_i^*(t)$ be the reliability apportioned to the ith unit for $i=1,\dots,k$. The system reliability is:

$$R^*(t) = \prod_{i=1}^{k} R_i^*(t) \qquad t > 0 \qquad (2)$$

The allocation procedure is performed through an iterative process. The first step starts from the initial plan, when few data are available concerning components.
In this phase few data concerning components are still available, for this reason a common practice is to consider the subsystem in series and to adopt an allocation strategy designed for such systems.
Furthermore, when no precise information about the unit can be obtained, an Equal Method can be used. Let w_i be a reliability allocation weight for the general unit i, the failure rate target λ^* is apportioned equally among all units as follow:

$$\lambda_{i^*}^* = w_i \lambda^* \qquad (3)$$

where λ_{i^*} is the failure rate of the unit i and:

$$w_{i^*} = \frac{1}{k} \qquad i = 1, \dots, k \qquad (4)$$

It is possible to predict the failure rate of the subsystem i when $\lambda_{i^*}^*$ is available from databanks or background experience on comparable subsystem, the ARINC method provides:

$$w_{i^*} = \frac{\lambda}{\sum_{i=1}^{k} \lambda_i} \qquad (5)$$

A weighted sum of Equations (4) and (5) is used by the Boyd Method:

$$w_{i^*} = K\left[a\frac{1}{k} + (1-a)\frac{\lambda_i}{\sum_{i=1}^{k} \lambda_i}\right] \qquad (6)$$

where the a weight allows to couple the ARINC method with the Equal method and K is a safety margin.

When many parameters about the number of components and their connections are known, we can use more accurate methodologies. The AGREE method is based on the mathematical formulation:

$$R_{i^*}^* = R^*(t)^{w_i} \tag{7}$$

where:

$$\sum_{i=1}^{k} w_i = 1 \qquad and \qquad w_{i^*} = \frac{n_i}{\sum_{i=1}^{k} n_i} \tag{8}$$

$i=1,\ldots,k$ where the number of components in unit i is n_i. According to the growth of the available data about the number of components and their interconnections, it is possible to use more complex methods, based on different factors: system criticality, technology, mission time etc. The Bracha method considers the following four factors:

1. A_{i1}: state-of-the-art of unit *i* technology.
2. A_{i2}: complexity of unit *i*, given by the number of components in unit *i*.
3. A_{i3}: operating time of unit *i*, given by the operative time of unit *i* with respect to the total mission time.
4. A_{i4}: environmental condition, evaluated in terms of the applied stress to the unit *i*.

The allocation weight w_i is:

$$w_i = \frac{A_{i1}(A_{i2} + A_{i3} + A_{i4})}{\sum_{i=1}^{k}[A_{i1}(A_{i2} + A_{i3} + A_{i4})]} \quad i = 1, \ldots, K \tag{9}$$

In a similar manner, the FOO method is based on the equation (2) and it considers four factors:

(1) A_{i1}: state-of-the-art of unit i technology.
(2) A_{i2}: intricacy of unit *i*.
(3) A_{i3}: operating time of unit *i*.
(4) A_{i4}: environmental condition.

A detailed evaluation gives a ten-point numerical scale, and the final allocation weight is:

$$w_i = \frac{(A_{i1}A_{i2}A_{i3}A_{i4})}{\sum_{i=1}^{k}[A_{i1}A_{i2}A_{i3}A_{i4}]} \qquad i = 1, \ldots, k \tag{10}$$

Using the same basis of the FOO method, Chang *et al.* and Liaw *et al.* consider a reliability allocation weight based on an operator introduced by Yager. This technique aggregates n factor ratings together and assigns the weight to subsystems, according to the following rules:

1. high weight: when all factors have high ratings; and
2. high weight: when at least one of the factors has a high rating.

In relation to this rule, the operator named "orness" (a situation parameter) can assume a value between 0 and 1. Using the maximal entropy of the weight ME-OWA proposed by O'Hagan (1988) and Fuller and Majlender (2001), it is possible to determine an ordered position of a factor rating within the subsystem and finally to calculate the related weight b_j. Then, the allocation weight is:

$$w_i = \frac{\sum_{j=1}^{n} b_j A_{ij}}{\sum_{i=1}^{k}\sum_{j=1}^{n} b_j A_{ij}} \qquad i = 1, \ldots, k \tag{11}$$

Where A_{ij} is the ordering of A_{ij} within a subsystem *i* and n is the number of factors. As mentioned before, the weight b_j only depends on the specific position in a subsystem of the considered factor and not on its importance.
We now consider the case when the failure rate $\lambda_{i^*}^*$ is constant for all subsystems and reliability of the system and all the units is close to one. Using an approximation, $1 - e^{-\lambda t} = \lambda t$, the equation (3) becomes:

$$F_{i^*}^*(t) = w_i F^*(t) \quad i = 1, \ldots., k \tag{12}$$

Where $F_{i^*}^*$ is the unreliability of the unit i and F^* the unreliability of the whole system.
Karmiol method uses the Equation (12) and the following four factors:

1. A_{i1}: state-of-the-art of unit i technology.
2. A_{i2}: intricacy of unit *i*.
3. A_{i3}: operating time of unit *i*.
4. A_{i4}: criticality of unit i, evaluated as the effect of the unit i failure on mission success.

The allocation weight is:

$$w_i = \frac{(A_{i1} + A_{i2} + A_{i3} + A_{i4})}{\sum_{i=1}^{k}[A_{i1} + A_{i2} + A_{i3} + A_{i4}]} \quad i = 1, \ldots k \tag{13}$$

where A_{ij} are numerical rating on a 10-point scale.
Some authors recently proposed the failure effect in reliability allocation using Risk Priority Number (RPN) generated during the FMEA analysis. Let unit *i* have N_i failure modes with severity ranking S_{ij}, occurrence rating O_{ij} and detection ranking D_{ij} for j=1,….,N and *i=1,..,k*. In order to calculate the three factors, an ordinal scale from 1 to 10 is used. The RPN of failure mode *j* in unit *i* is:

$$RPN = S_{ij} \times O_{ij} \times D_{ij} \tag{14}$$

The reliability allocation weight is:

$$wi = \frac{\omega_i}{\sum_{i=1}^{k} \omega_i} \tag{15}$$

Where:

$$\omega_i = 1 - \frac{C_i}{\sum_{i=1}^{k} C_i} \quad (16)$$

and

$$C_i = \frac{1}{N_i}\sum_{j=1}^{N_i}\left(S_{ij} \times O_{ij} \times D_{ij}\right) \quad (17)$$

The IFM method considers the six factors:

1. A_{i1}: criticality of unit *i* (the ratio between the number of subsystem functions able to generate an undesirable event and the total number of the unit functions).
2. A_{i2}: intricacy of unit *i* (the ratio between the number of parts that form the unit and those that form the whole system).
3. A_{i3}: functionality of unit *i* (the ratio between all unit functions and the number of functions of the whole system).
4. A_{i4}: Effectiveness index (the ratio between the unit effectiveness time and the mission total time).
5. A_{i5}: Technology index (A_{i5}=0,5: traditional components; A_{i5}=1: innovative components).
6. A_{i6}: Electronic Functionality index (A_{i6}=1: completely electronic system; A_{i6}=0,1: completely mechanical system). With the purpose of distinguishing between electronic systems and mechanical ones, characterized by the same complexity.

The allocation weight is given by:

$$w_i = GI\% = \frac{(A_{i1}^{-1}A_{i2}A_{i3}A_{i4}A_{i5}A_{i6}^{-1})}{\sum_{i=1}^{k}[A_{i1}^{-1}A_{i2}A_{i3}A_{i4}A_{i5}A_{i6}^{-1}]}$$

$$\text{i}= 1, \dots, k \quad (18)$$

where A_{ij} are numerical ratings of factor.

The A-IFM method adds the analytic hierarchy process to the IFM reliability allocation method. This technique is able to address some weaknesses common to conventional allocation methods. In fact, through the AHP support, it is possible to perform a pairwise comparisons between the units and factors and consequently to assign a greater unreliability allocated to units less responsible for the mission failure. This technique allows to give a more accurate ranking to different A_i factors and to different units, in reason of their importance during the allocation process reliability.

Finally, the CFM is a recent reliability allocation method designed for complex systems. It can be used for both series and parallel systems, using analysis only of significant units (according to experience). This innovative aspect allows to limit the analysis to a very low number of elements, depending on the particular Top Event. In this way, there are less dispersed results and it is possible to create a scale of criticality and identifying priorities and hierarchies. The method considers five factors:

1. A_{i1}: criticality. It depends on system configuration: $A_{i1} = \frac{1}{n}$ where n corresponds to the number of "buffer elements". It varies between 0 (infinite units in a parallel configuration) and 1 (system series).
2. A_{i2}: state-of-the-art of unit *i* technology (A_{i2}=0 for old design and A_{i2}=1 for newly developed units).
3. A_{i3}: complexity. It considers the complexity of the structure, assembly and interactions through three possible values: A_{i3}=0,33 Not Complex subsystem; A_{i3}=0,66 Normal Complexity; A_{i3}=1 Complex subsystem.
4. A_{i4}: running time. It is given by the operative time of unit *i* with respect to the total mission time.
5. A_{i5}: operation profile. It considers environmental condition and it is evaluated in terms of the external applied stress to the unit *i*, through three possible values: A_{i5}=0,33 Easy operative conditions, A_{i5}=0,66 Normal operative conditions, A_{i5}=1 Difficult operative conditions.

The allocation weight is:

$$w_i = \frac{(A_{i1}A_{i2}A_{i3}^{-1}A_{i4}A_{i5})}{\sum_{i=1}^{k}[A_{i1}A_{i2}A_{i3}^{-1}A_{i4}A_{i5}]}$$

$$i = 1, \dots, k \quad (19)$$

Then, it is possible to allocate the reliability target using the equation (7).

Each analyzed reliability allocation method shows different characteristics and own strengths and limitations. Some methods can be used when very few data are available about the system, as the number of components and their connections, at the same time more accurate methods have been proposed for an advanced phase of the product development process. Finally, more recent methodologies can be successfully applied in very complex systems and often through the support of appropriate mathematical models. All these aspects will be dealt with in detail in Section 3, providing a general classification in order to help researchers and professionals in the choice of the most suitable method for each specific application.

3. CLASSIFICATION OF RELIABILITY ALLOCATION METHODS

In this section, a detailed comparison is performed between frequently-used reliability allocation methods present in literature and analyzed in the Section 2. As was mentioned above, the classification of methods can be studied from different aspects, such as advantages and disadvantages, the phase of the product development process, system complexities, mathematical algorithms used and so on. However, it is interesting to note how the choice of a proper allocation method always depends on

the level of available information about the system. Generally, the reliability allocation should occur in the design phase of a product or before the design of major system upgrades. It is reasonable to assume that the reliability prediction becomes more accurate along with the development of the design process.
As mentioned by Guangyan (2011), we can firstly subdivide the design process in:

- a demonstration design stage, when the product definition is not very clear;
- a preliminary design stage, when a very limited reliability data exist; and
- a detailed design stage: when specific details concerning the failure rate behavior of components and system configurations arise.

Furthermore, as indicated in the Section 1, a proper allocation method for each application can be chosen according to the level of knowledge of the system complexity, as the number of components and their assembly, or more operative information, i.e. maintenance policy (Andriulo et al. 2015). In addition, some methods, such as the already mentioned A-IFM and ME-OWA, add useful mathematical tools for a more accurate ranking of weight factors, according to the equation 3 or, in a similar manner, the equation 12. For example mathematical tools are: Fuzzy Decision Method (Zhang and Jia 2009, GU and Huang 2009), AHP (Di Bona et al. 2015, Lee et al. 2008), the maximal entropy ordered weighted averaging method (Chang et al. 2009), Analytical Target Cascading Method (Guo and Bai 2009), sensitivity evaluation method (SEM) (Ali 2009) etc.. Mathematical tools can help to address some weakness of conventional allocation methods, in terms of subjectivities, comparisons and general uncertain information.
As a guideline for the method choice, we opted for the realization of a comprehensive table that summarizes each method with its features.
Each row of the Table 1 contains a different reliability allocation method, the rows are ordered according to the level of the system information, in particular it increases from the top to the bottom of the list. First two columns show advantages and disadvantages, the third column an index that refers to the design stage considered (I: demonstration stage, II: preliminary design stage, III: detailed design stage) and the last column the possibility of a mathematical tool employment.

4. CONCLUSIONS

This paper analyzes state-of-the-art of the reliability allocation strategies present in literature determining the area of their application. The review of existing literature shows that there is a continuous improvement of conventional methods and several new ones have been developed in recent years. In general, each method requires its own level of system information and can be appropriate only for some phases of the product development. The main trend observed is the adoption of ever more advanced mathematical tools, in order to overcome some weakness of conventional allocation methods, in terms of subjectivities, comparisons and general uncertain information. In particular, mathematical tools have been used to improve the accuracy of some allocation methods with excellent results in various applications, but we believe that these tools can also be combined with other analyzed techniques.
In general, starting from the substantial lack of a general classification of reliability allocation methods, the present work aims to provide a practical guide for the choice of the most suitable method for each specific application.
As a result, we propose a table highlighting features of each frequently-used reliability allocation methods (Table 1).

Table 1. Classification of frequently-used reliability allocation methods

Increasing level of the system information	Methods	Advantages	Disadvantages	Design stage	Mathematical tool possibility
	Equal	Application semplicity	Not giving considerations to the actual differences between various subsystems	I	X
	ARINC	Application semplicity Objectivity	Only applicable to systems in series Only applicable in initial phases Fault rate knowledge of similar systems	II	✓
	Boyd	Versatility	Only applicable to systems in series Only applicable in initial phases	II	X
	AGREE	Good detail	Only applicable to systems in series Applicable in the advanced phase Partial subjectivity of the analyst	III	X
	Bracha	Exact analytical treatment	Not easy determination of stress factors Component criticality not considered	III	✓
	FOO	Application simplicity	Only applicable to systems in series Only applicable in initial phases	II	✓ example: ME-OWA or FOO+SEM
	Karmiol	Very good detailed Applicable to innovative systems	Subjectivity of the analyst	II	✓
	IFM	Very good detailed Applicable to innovative systems Applicable in every phase	Only applicable to systems in series Detailed information required	III	✓ example: A-IFM
	CFM	Applicable to complex system (parallel systems included)	Detailed information required	III	✓ example: AHP-CFM

Methods are listed according to the level of available information about the system (it increases from the top to the bottom of the list), the first two columns show advantages and disadvantages for each method, the third column provides an index that refers to the considered design stage and the last column shows the possibility of a mathematical tool enhancement.

This study represents the first step toward the realization of a comprehensive and useful guide for researchers and professionals interested in the choice of the most suitable method for the specific application, considering the development phase, level of information about the system, level of desired accuracy, the system complexity and so on. For further research we plan to elaborate a more systematic comparison between reliability allocation methods, with particular attention to their application typology and costs. A such literature analysis can contribute to more comprehensive knowledge of allocation strategies and lead to an appropriate choice of the method.

REFERENCES

Advisory Group of Reliability of Electronic Equipment (AGREE), 1957. Reliability of military electronic equipment. Office of the Assistant Secretary of Defense Research and Engineering, Washington, DC.

Aggarwal K.K., 1993. Reliability engineering. Dordrecht, Netherlands: Kluwer Academic Publishers.

Ali J., 2009. Calculation of reliability allocation factor using sensitivity evaluation method, Proceedings of 8th International Conference on Reliability, Maintainability and Safety, pp. 83-86.

Alven W.H., 1964. Reliability engineering: Prepared by ARINC research corporation. Englewood Cliff, NJ: Prentice Hall, Inc.

Anderson R.T., 1976. Reliability design handbook. Chicago: ITT Research Institute.

Andriulo S., Arleo M.A., De Carlo F., Gnoni M.G., Tucci M., 2015. Effectiveness of maintenance approaches for high reliability organizations, IFAC PapersOnLine, 28(3), 466-471.

Bowles J.B., 2003. An assessment of RPN prioritization in a failure modes effects and criticality analysis. In Processing annual reliability and maintainability symposium, pp. 380–386.

Bracha V.J., 1964. The methods of reliability engineering. Machine Design, 70–76.

Cheng C.H., Chang, J.R., 2006. MCDM aggregation model using situational ME-OWA and ME-OWGA operators. Fuzziness and Knowledge Based Systems, 14(4), 421–443.

Chang K.H., Cheng C.H., Chang Y.C., 2008. Reliability assessment of an aircraft propulsion system using IFS and OWA tree. Engineering Optimization, 40(10), 907–921.

Chang Y.C., Chang K. H., Liaw C.S., 2009. Innovative reliability allocation using the maximal entropy ordered weighted averaging method. Computers & Industrial Engineering, 57, 1274-1281.

Dai S.H., Wang M.O., 1992. Reliability analysis in engineering applications. New York: Van Nostrand Reinhold.

Department of Defense of USA, 1988. MIL-HDBK-338B, Electronic design reliability handbook

Department of the Army. TM 5-689-4., 2006. Failure modes, effects and criticality analysis (FMECA) for command, control, communications, computer, intelligence, surveillance, and reconnaissance (C4ISR) facilities.

De Felice F., Petrillo A., 2014. Proposal of a structured methodology for the measure of intangible criteria and for decision making. International Journal of Simulation and Process Modelling, 9(3), 157-166.

De Felice F., 2012. Editorial Research and applications of AHP/ANP and MCDA for decision making in manufacturing. International Journal of Production Research, 50(17), 4735–4737.

Di Bona D., Duraccio V., Silvestri A., Forcina A., Falcone D., De Felice F., 2014. Validation and application of a reliability allocation technique (advanced integrated factors method) to an industrial system. Proceedings of the IASTED International Conference on Modelling, Identification, and Control, pp.75-79. Innsbruck; Austria.

Di Bona G., Forcina A., Petrillo A., De Felice F., Silvestri A., 2016. A-IFM reliability allocation model based on multicriteria approach. International Journal of Quality & Reliability Management, 33(5), 676–698.

Di Bona G., Silvestri A., Forcina A., Falcone D., 2017. AHP-IFM Target: An innovative method to define reliability target in an aerospace prototype based on analytic hierarchy process. Quality and Reliability Engineering International, 33(8), 1731-1751.

Di Bona,G., Falcone D., Silvestri A., Forcina A., 2017. IFM target 2.0: an innovative method to define reliability target for prototype systems. International Journal of Advanced Manufacturing Technology, 95(9-12), 3349-3367.

Di Bona G., Forcina A., Silvestri A., 2015. Critical Flow Method: A New Reliability Allocation Approach for a Thermonuclear System. Quality and Reliability Engineering International, 32(5), 1677–1691.

Di Bona G., Forcina A., 2017. Analytic Critical Flow Method (ACFM): A Reliability Allocation Method Based on Analytic Hierarchy Process. Journal of Failure Analysis and Prevention, 17(6), pp. 1149-1163.

Di Pasquale V., Miranda S., Iannone R., Riemma S., 2015. A Simulator for Human Error Probability Analysis (SHERPA), Reliability Engineering and System Safety, 139, 17-32.

Duraccio V., Compagno L., Trapani N., Forcina A., 2016. Failure Prevention Through Performance Evaluation of Reliability Components in Working Condition. Journal of Failure Analysis and Prevention, 16(6), 1092-1100.

Falcone D., Di Bona G., Duraccio V., Silvestri A., 2007. Integrated hazards method (IHM): A new safety allocation technique. Proceedings of the IASTED International Conference on Modelling and Simulation, pp. 338-343

Fuller R., Majlender P., 2001. An analytic approach for obtaining maximal entropy OWA operator weights. Fuzzy Sets and Systems, 124(1), 53–57.

GU Y.K., Huang K.Q., 2009. Fuzzy Reliability Allocation Method for Engine based on the Experts Knowledge. Journal of Aerospace Power, 24, 1143-1149.

Guangyan Z., Tong Q., Yufeng S., Weiwei H., 2011. The Method of Mission Reliability Allocation for Complex System Based on Simulation. School of Reliability and System Engineering, Beihang University, Beijing, China

Guo Z.W., Bai G.B., 2009. Reliability Allocation Using Analytical Target Cascading Method. Proceedings 2009 IEEE 16th International Conference on Industrial Engineering and Engineering Management, pp. 1195-1199.

Itabashi-Campbell R.R., Yadav O.P., 2009. System reliability allocation based on FMEA criticality. Proceedings of SAE World Congress & Exhibition. Detroit, (Michigan USA).

Karmiol E.D. 1965. Reliability apportionment. Preliminary Report EIAM-5, Task II, General Electric, Schenectady, NY pp. 10–22.

Kim O., Yang Y., Zuo M.J., 2007. A new reliability allocation weight for reducing the occurrence of severe failure effects. Reliability Engineering and System Safety, 117, 81-88.

Kuo H.E., 1999. Reliability assurance: Application for engineering and management (2nd ed.). Chinese Society for Quality

Lee G.L., Lin H.J., Yu T.W., Ma, C.C., 2008. Optimal Allocation for Improving System Reliability Using AHP. IEEE International Conference on Sustainable Energy Technologies, pp. 159-163.

Liaw C. S., Chang Y.C., Chang K.H., Chang T.Y.. 2011. ME-OWA based DEMATEL reliability apportionment method. Expert Systems with Applications, 38, 9713–23.

NASA Report, 2007. Designing for Dormant Reliability, Johnson Space Center (JSC) Guideline No. GD-ED-2207.

O'Hagan M., 1988. Aggregating template or rule antecedents in real-time expert systems with fuzzy set logic. Proceedings of 22nd annual IEEE Asilomar conference on signals, systems, computers, pp. 681–68. Pacific Grove, CA, Piscataway. NJ.

Saaty T.L., 1980. The Analytic Hierarchy Process. Third ed. McGraw-Hill, New York.

Saaty T.L., 1990. How to make a decision: The Analytic Hierarchy Process. European Journal of Operations Research, 48, 9–26.

Smedley K.,1992. Reliability analysis for LEB ring magnet power system in SSC. IEEE Transactions on Nuclear Science, 39(4), 1170–1174.

Tucci M., De Carlo F., Borgia O., Fanciullacci, N., 2014. Accelerated life tests as an integrated methodology for product design, qualification and production control: a case study in household appliances. Production & Manufacturing Research, 2(1), 112-127.

Yadav O.P., Singh N., Goe P.S., 2006. Reliability demonstration test planning: a three dimensional consideration. Reliability Engineering and System Safety, 91:882–93.

Yadav O.P., 2007. System reliability allocation methodology based on three-dimensional analyses. International Journal of Reliability and Safety, 1, 360–75.

Yadav O.P., Zhuang X., 2014. A pratical reliability allocation method considering modified critically factors. Reliability Engineering and System safety, 129, 57-65.

Yager R.R. 1988. On ordered weighted averaging aggregation operators in multicriteria decision making. Systems, Man and Cybernetics, 18(1), 183-190.

Zhang H.B., Jia Z.X., 2009. Complex System Reliability Allocation based on Fuzzy Decision Method, International Workshop on Intelligent Systems and Applications.

A METHODOLOGICAL FRAMEWORK TO IMPLEMENT LEAN IN DYNAMIC AND COMPLEX SOCIO-TECHNICAL SYSTEMS

Joshua Behr[(a)], Rafael Diaz[(b)], Francesco Longo[(c)], Antonio Padovano[(d)]

Virginia Modeling, Analysis & Simulation Center, Old Dominion University, USA[(a)]
Zaragoza Logistics Center, Spain[(b)]
DIMEG, University of Calabria, Italy [(c)(d)]

jbehr@odu.edu[(a)], rdiaz@zlc.edu.es[(b)], f.longo@unical.it[(c)], antonio.padovano@unical.it[(d)]

ABSTRACT
Dynamic and complex socio-technical systems notoriously hinder the application of lean thinking and tools and, as a result, any improvement opportunity. Emblematic is the case of a post-disaster recovery, whose processes and performance are unpredictable due to many sources of complexity, randomness and waste, both for industrial organizations and communities.
This article proposes an innovative framework, named LOOP, which permits to implement lean thinking even in dynamic and complex scenarios. It consists of an iterative technology-driven methodology based on four steps: Internet of Things-driven Real-Time Knowledge Creation, Waste Identification & Value Focusing, Modeling & Simulation driven Set-Based Concurrent Engineering and Knowledge Management & Lean Implementation. The contribution of this article is then mainly methodological: this framework is proposed as the guidance to lead socio-technical systems, such as a post-disaster recovery, through a value-oriented continuous improvement process with a systemic perspective, which concurrently consider uncertainties, dynamic factors and interdependencies as sources of wastes in a post-disaster recovery.

Keywords: Lean Management, Waste, Modeling & Simulation, Internet of Things, Post-Disaster Recovery

1. INTRODUCTION

Stability creates a basis for lean-based process improvement in modern organizations, which deliver tangible artefacts and intangible services to cope with a growing market competitiveness and to meet customer needs better (Wang et al., 2011). However, very few environments and systems can be considered stable because most systems are characterized by high levels of complexity, uncertainty, non-linear behavior, diversity, and dynamic interactions, among other features (Cilliers, 1998).
Although lean thinking has been more and more used in highly complex systems (Azadegan et al., 2013), Kolberg & Zühlke (2015) argue that it may have today reached its limit, since modern systems' level of complexity is in conflict with the required stability praised by lean and its blindfolded application can have unintended consequences that can make the whole system more vulnerable to unexpected variability (Soliman & Saurin, 2017). Considering the aforementioned context, the research question addressed by this study is stated as follows: how can lean thinking be reliably applied to unstable and dynamic complex environments and systems?
This study answers by proposing an innovative iterative methodological framework, named LOOP and developed by the authors, which leverages on the new-fashioned enablers of the Digital Transformation that are offering ways to rethink and handle system complexity (Kenney, 2015), namely Big Data, Process Digitalization, Internet of Things (IoT), Modeling & Simulation (M&S), Artificial Intelligence. Given the dual (product and service oriented) nature of modern organizations, the LOOP framework combines the principles of lean manufacturing and lean services and consists of 4 steps that guide planners and managers in the proper evaluation, design and implementation of lean management practices in complex socio-technical systems: (i) an IoT-driven Real-Time Knowledge Creation; (ii) a Waste Identification & Value Focusing; (iii) a M&S-driven Set-Based Concurrent Engineering; (iv) a Knowledge Management & Lean Implementation. The framework has been designed to be iterative to continuously adapt to the unpredictable behavior of a complex socio-technical system and react immediately to any change that may deviate the processes from their optimum, thus establishing a foundation for true improvement.
An emblematic application study of topical interest is represented by a post-disaster recovery scenario. Since it is nearly impossible to develop a meaningful one-size-fits-all pre-disaster recovery plan because of the dynamic factors that intervene and it is hard to track the progress of recovery operations in the aftermath of a disaster (Zobel, 2014), this article opens the discussion on the potential benefits of a lean post-disaster recovery.
The LOOP framework is finally proposed as a guidance lead a complex socio-technical system through a value-oriented continuous improvement process and as the

key basal architecture of a decision support ecosystem (whose development is the focus of further work and research by the authors) to drive the implementation of lean thinking and practices in a post-disaster recovery.

2. BACKGROUND ANALYSIS

Organizations worldwide are seeking ways to focus on what is valuable for their customers, to optimize their value creating processes by cutting out waste and to perfect the entire operation so that the service flows smoothly, thus acquiring a strong position in a more and more competitive market. Management literature has suggested that contextual factors may present strong inertial forces that inhibit practices that appear technically rational (Nelson & Winter, 1982). Lean has been identified as an organizational change and improvement method, particularly as a time and cost reduction mechanism (Achanga, 2006; Bicheno, 2008), and is still today one of the most common initiatives in Operations Management that organizations adopt to boost their competitiveness and performance (Abreu-Ledón et al., 2018). The lean philosophy has been coined with reference to manufacturing systems, where it soon became a multi-dimensional approach that encompasses a wide variety of management practices, including just-in-time, quality systems, work teams, cellular manufacturing, supplier management, etc. aimed at creating a streamlined, high quality system that produces finished products at the pace of customer demand with little or no waste (Shah & Ward, 2003). Over the years, lean thinking has been defined as a "dynamic, knowledge-driven, and customer-focused process through which all people in a defined enterprise continuously eliminate waste with the goal of creating value" (Murman et al., 2002).

This definition gives space to the application of lean principles also in the service industry. The desirability of transferring lean manufacturing logic and practices to service operations was already defended by Bowen & Youngdahl (1998), who have still a production-line perspective of lean services though. After decades of lean principles implementation in manufacturing companies, lean principles have been recently applied successfully especially in the healthcare sector, where the Toyota Lean manufacturing principles improved waiting times and quality of care in emergency departments (Ng et al., 2010). A lean service philosophy has established itself next to the lean manufacturing (Andrés-López et al., 2015), though the research is still at nascent stage. Gupta et al. (2016) suggests to focus on process difference between services and manufacturing and on the respect for people and employment engagement (among others) in order to benefit the most from the implementation of lean principles in the service industry. Today is clear that lean service has a positive impact on firm operational and financial performance but a systematic approach on the system as a whole is needed when implementing lean service practices (Hadid et al., 2016) because the interplay between manufacturing and service operations may conceal pitfalls in the adequate lean implementation.

2.1. Limits to lean: unstable and dynamic complex systems

When dealing with modern complex systems and organizations, what immediately emerges is that they cannot be described and modelled mathematically and their behavior is unpredictable to some extent (Dekker, 2016) mainly because (among other causes) of the dynamic interactions among a large number of elements and operations far from equilibrium (and then instability). In these environments, lean thinking has been proved to have ambiguous impacts on the system's performance. For example, Azadegan et al. (2013) reported on the basis of a survey with 126 US manufacturing companies that when the rate of environmental change, complexity and dynamism was high, performance was not improved by lean operations. A similar finding was obtained by Chavez et al. (2013), with 228 manufacturing Irish companies, in which lean improvement has been achieved only when dynamism was low. Complexity is generally regarded as a barrier to lean, since it creates uncertainty and variability that makes it difficult the implementation of core lean practices such as standardization and pull production. Some studies recommend to reduce complexity to the possible extent: "lean works best in predictable environments" (Summers & Scherpereel, 2008) and "the higher levels of unpredictability and instability in dynamic environments […] undermines the effectiveness of lean operations" (Azadegan et al., 2013). Barriers to lean practices have been identified organically only recently in Soliman & Saurin (2017) that list them as follows: (i) limited suitability of lean principles and practices (e.g. the kanban system cannot be implemented in environments in which the demand is unstable); (ii) difficulty to establish cause-and-effect relationships; (iii) no clear-cut boundaries between waste and value; (iv) dynamic contextual factors that are hard to manage. Therefore, the research question that guided this study had been stated as follows: how can lean thinking be reliably applied to unstable and dynamic complex environments and systems?

2.2. Lean insights into a post-disaster recovery

The very few attempts to bring more insight into the post-disaster recovery scenario are limited to the reconstruction process (Mojtahedi & Oo, 2012; Diaz et al., 2015) mainly because of its manufacturing-oriented nature. A remarkable attempt has been performed by the nonprofit association St. Bernard Project (SBP) which is the first recovery organization recognizing the potential of lean principles in post-Katrina recovery in New Orleans (Marchwinski, 2015). Despite the spotlight of this visionary attempt is essentially the reconstruction process, it proves that lean practices yield significant benefits, such as increased resources utilization, fewer defects, reduced waiting and lead time and ultimately, a higher satisfaction of the affected population. However, lean thinking can remarkably boost not only the reconstruction process but the whole post-disaster scenario with industries and communities as sides of the same coin. The strategic connection between the post-disaster recovery of industrial organizations and communities is clear to Toyota, pioneer in the lean management, which considers trust between businesses and the well-being of employees, their families and the community they live in as the first factor for a successful recovery from a disaster. The company is aware that in case of a disaster a hasty “production first” attitude will end up destroying the trust employees, suppliers and communities have developed over time and that if the company switches to an alternative supplier as soon as things go south, they won’t trust the company anymore. Since the interplay among different factors that characterize a post-disaster recovery scenario limit the ability to predict the outcomes of the process (Azadegan et al., 2012) and therefore the applicability of lean practices, a theoretical and methodological framework capable to support lean-oriented decision-making in dynamic and complex socio-technical systems, such as a post-disaster recovery process, is of vital importance.

3. THE LOOP FRAMEWORK

Typically, the first step in creating lean-compliant processes is to reach a consistent level of capability. But this is not always possible because tools, materials and people may be unavailable or too expensive to acquire. Therefore, a lean based improvement may be the only viable solution to get the maximum from the available resources and avoid wastes that causes inefficiencies.

An innovative iterative methodological framework, named LOOP, has been developed by the authors. This framework (depicted in Figure 1) embraces the kaizen philosophy for a continuous improvement spiral (hence the name of the framework) in complex dynamic socio-technical systems which leverages on the new-fashioned enablers of today’s Digital Transformation that is offering ways to rethink and handle complexity, including Process Digitalization, Internet of Things (IoT), Big Data, Modeling & Simulation (M&S), Artificial Intelligence, etc. The LOOP framework has been designed to be iterative to continuously adapt to the unpredictable behavior of a complex socio-technical system and react immediately to any change that may deviate the processes from their optimum. The primary objective of this methodology is to create the basis for consistency in the application of lean management practices to unstable environments, such as a post-disaster recovery, so that the “reality” can be seen and random activities removed, thus establishing a foundation for true improvement.

The remaining of this section discusses the four steps of this framework that guide planners and managers in the proper implementation of lean management practices in dynamic and complex socio-technical systems:

- **Step 1: IoT-driven Real-Time Knowledge Creation** (Planning, Control & Lean Implementation);
- **Step 2: Waste Identification & Value Focusing** (System State Analysis);
- **Step 3: M&S-driven Set-Based Concurrent Engineering** (System State Analysis);
- **Step 4: Knowledge Management & Lean Implementation** (Planning, Control & Lean Implementation).

The two macro-areas indicate that a good decision-making is guaranteed only when the System Control & Lean Management relies upon a consistent System State Analysis based on reliable knowledge. It is worth mentioning that the successful use of the present framework requires the design and development of a decision support ecosystem (as a combination of distributed IoT devices and a central simulation-based inference engine) as discussed in the following.

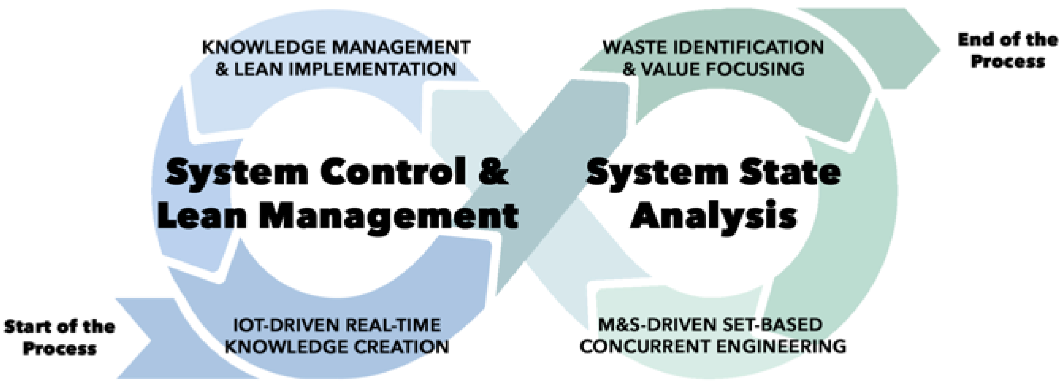


Figure 1: The LOOP framework

3.1. IoT-driven Real-Time Knowledge Creation

To fulfill the framework objectives, the system should have the ability to perceive the context. While the use of IoT is consolidated in most of today’s manufacturing systems, it is a little bit disregarded in many other complex socio-technical systems (e.g. a post-disaster recovery). However, they could hugely benefit from IoT especially when they present a strong level of dynamism and complexity which makes hard monitoring their evolution over time and behavior. Internet of Things (IoT) increasingly gives the ability to sense and monitor the environment in which it is operating. The most significant advantage of this step is therefore the real-time visibility over the system, thus meeting the need for real-time information transparency. This is achieved by considering that the only way to manage uncertainty is a rapid data collection by devices and sensors (e.g. RFIDs, GPS, raw sensors) and exchange over an information network. Obviously, this step requires a preliminary ad-hoc design and implementation of an IoT-based ecosystem that is tailored on the socio-technical system under study. Devices and sensors are today more affordable than ever and very flexible which makes them suitable for every environment.

The process control is not (and cannot be) limited to the collection of data. We should not stop at just obtaining real-time visibility, but should continue on to manipulate that visible big amount of real-time data (retrieved by using IoT devices) to create a synthetic situational knowledge about the system state to be used for the next analysis. Thanks to this knowledge, the wastes can be easily identified in the current state of the system, thus focusing only on those processes that provide value added. The knowledge creation is therefore a crucial point of the methodology to get valuable information to be used in the next steps of the LOOP framework. The knowledge is created and updated by processing the context information, such as the stream of collected data from the different sensors. Data need to be aggregated and then reasoned with an inference engine that recognizes the current context to classify the context status. The main work here is then to use real-time ontologies, statistical analysis, big data management algorithms and artificial intelligence to generate real-time knowledge about the current state of the system.

An **'Internet of Things driven real-time knowledge creation'** is therefore the Step 1 that gives the start to this framework and will feed the next step in the framework.

3.2. Waste Identification & Value Focusing

In the case of modern socio-technical systems, a dual nature can be recognized. Since a complex socio-technical system is generally neither a manufacturing nor a service system but rather a combination of them, the LOOP framework continue with a **'waste identification & value focusing'** step, which combines the principles of lean manufacturing and lean services. Thanks to the deep real-time knowledge of the current state of the investigated system provided by the previous step, it is easier to identify the wastes in the system. Starting from the analysis of traditional wastes in manufacturing and service industries, 8 sources of waste can be identified for socio-technical systems as a combination and reinterpretation of the wastes discussed by Ohno (1988), Seddon (2003), Womack & Jones (2005) and Bicheno (2008), namely:

- **defects** refer to errors, mistakes, low-quality products or incomplete work that require additional resources to be fixed;
- **excess processing** refers to the process of adding more value to a product than the customer really requires, to effort duplication, and to the execution of multiple versions of the same task;
- **excess movement** refers to any movement of people, tools or material that does not add any value to the service or product delivered;
- **waiting** refers to delayed or stopped tasks causing queues of materials or customers;
- **incorrect inventory** refers to materials that are needed but are out-of-stock, kept in duplicates, or kept for too long as well as to ordering errors;
- **not/low-utilized talent** refers to the people that are currently idle and inactive or that are not being utilized to their full capability or that are engaged in tasks that would be more efficiently done by someone else;
- **miscommunication** refers to the unclear verbal and written jargon and communication resulting into an additional time for clarifications and search for information;
- **failure demand** refers to the demand caused by a failure to do something or do something right, thus provoking bad relationships with stakeholders because of a lack of value focusing.

It can be observed that the iterative process finishes when every process in the system is over and wastes are no longer identified.

3.3. M&S-driven Set-Based Concurrent Engineering

Knowledge about the system state and wastes are the the key input for the Step 3 of the LOOP framework. This step is aimed at (i) analyzing the expected long-term behavior of the system based on its current state and its history in order to carry out an in-depth analysis of the sources of waste, and (ii) compare this scenario in the long term with other potential future scenarios, characterized by the implementation of specific lean strategies, by carrying out what-if analysis aimed at understanding which scenario would perform better. The first objective is fundamental because most of the times analysts are able to identify wastes in complex socio-technical systems but not their sources, hence an accurate analysis of the system's processes is needed. The second objective is strategic because it will drive the implementation of the best performing set of lean strategies to be applied in the real-world system. To this end, the LOOP framework proposes a **'M&S-driven Set-Based Concurrent Engineering'**, which begins by analyzing the current scenario on the long-term, broadly considering sets of possible future scenarios (where one or more lean practices are virtually implemented in the simulated system) and gradually narrowing the set of possibilities to converge on a final solution, thus producing useful insights that can be used by planners and managers to implement the correct set of lean management practices.

This step requires the design and development of a simulation-based inference engine to enable engineers exploring a number of alternatives in a short time interval based on current available knowledge about the system state, making rapid changes (as a sculptor does) to the current system state and evaluating the differences among different future potential scenarios. Modeling & Simulation (M&S) has widely demonstrated over the past years and thanks to the latest technological advances to significantly support the analysis of complex dynamic socio-technical systems thanks to its ability to compress and expand time (Banks, 1998) and evaluate different planning alternatives in different contexts, including disaster management (Bruzzone et al., 2014). An interesting example is provided in Longo et al. (2019) which shows how managers can model possible future state of a complex system (where one or more lean practices are virtually implemented in the simulated system) by using a digital value stream mapping approach and rapidly choose the best performing dynamic and uncertain business diversification and supply chain (re)design scenario.

3.4. Knowledge Management & Lean Implementation

The last step of the LOOP framework consists of harnessing the available knowledge created in the three previous steps and implement in the real-world system the best performing set of lean strategies and practices tested in the Step 3. This last step is then oriented to removing the process wastes that emerged dynamically and were identified in Step 2 and drive the whole system through a better performance as showed by the Step 3.

However, the dynamic nature of the complex socio-technical systems requires continuous iterations of the LOOP framework because it is not unreasonable to expect the implemented lean strategies to behave differently from what has been predicted and new wastes to emerge unexpectedly. A continuous control of the system variables and adjustments to the lean strategies are needed: the new system state deriving from the i-th iteration of the LOOP framework will be used therefore as input for the Step 1 of the $(i+1)$-th iteration of the LOOP framework, thus giving new fuel to the methodology and "keeping under control the system dynamism and instability" that may generate new undesired wastes.

The continuous improvement spiral (or loop) can be executed with a variable recurrence, which basically depends on: the level of instability of the system; the time needed to a specific set of lean strategies to produce statistically significant (negative or positive) results; on the gravity of the consequences that may arise if no mitigation actions or adjusted lean strategies are implemented when the system's behavior diverges.

4. CONCLUSIONS AND FURTHER WORK

This study is intended to raise attention about the need to support methodologically and technologically planners and managers in the implementation of lean management practices in complex socio-technical systems, such as a post-disaster recovery scenario. The presence of several uncertainties, dynamic factors and strong interdependencies among the difference aspects of a post-disaster recovery make it a fertile but rough soil for lean thinking to be applied without adequate measures that ensure a true improvement. The main contribution of this paper is the development of a methodological framework, named LOOP, which leverages on the new-fashioned enablers of a Digital Transformation, namely Big Data, Process Digitalization, Internet of Things (IoT), Modeling & Simulation (M&S), Artificial Intelligence, to establish a valid foundation for the evaluation of sets of lean practices in a dynamic and unstable environment. To overcome this issue and adapt to the changing behavior of these systems, the framework is iterative and provides a real-time reliable support both to the planning and to the management and implementation of the lean practices thanks to a ubiquitous knowledge about the current state of the system and its potential behavior in the long-term.

This framework is proposed specifically in the context of a post-disaster recovery, where managers must work on both short- and long-term perspectives, be able to coordinate a long list of expectations and potentially conflicting requests from all stakeholders under extremely stressful conditions and lead the whole recovery through an efficient and effective process. As suggested by this study, the management team must fight the tendency to consider the post-disaster recovery actions as isolated processes and adopt a systemic view, which consider all the possible interdependencies among the aspects of a post-disaster recovery. This approach will allow the clear identification of all the possible wastes (muda) in a post-disaster recovery and the LOOP framework will serve as basis for further work on this topic. Future research activities will be devoted to the analysis of the post-disaster recovery processes and the development of a decision support ecosystem able to support decision-making in complex socio-technical systems, such as a post-disaster recovery scenario, and streamline the implementation of lean management practices.

REFERENCES

Abreu-Ledón R., Luján-García D.E., Garrido-Vega P., Escobar-Pérez B. (2018). A meta-analytic study of the impact of Lean Production on business performance. International Journal of Production Economics, 200, pp. 83-102.

Achanga, P. (2006). Critical success factors for lean implementation within SMEs. *Journal of Manufacturing Technology Management*, 17 (4), pp. 460-71.

Andrés-López, E., González-Requena, I., & Sanz-Lobera, A. (2015). Lean service: reassessment of lean manufacturing for service activities. Procedia engineering, 132, 23-30.

Azadegan, A., Patel, P. C., Zangoueinezhad, A., & Linderman, K. (2013). The effect of environmental complexity and environmental dynamism on lean practices. Journal of Operations Management, 31(4), 193-212.

Azadegan, A., Patel, P.C., Parida, V., 2012. Operational slack and venture survival: the effects of environ-

mental uncertainty. Production and Operations Management Journal 22 (1), 1–18.
Banks, J. (1998). Handbook of simulation: principles, methodology, advances, applications, and practice. John Wiley & Sons.
Bicheno, J. (2008). The Lean Toolbox for Service Systems. Buckingham: Picsie Books.
Bowen, D. E., & Youngdahl, W. E. (1998). "Lean" service: in defense of a production-line approach. International journal of service industry management, 9(3), 207-225.
Bruzzone, A., Frascio, M., Longo, F., Chiurco, A., Zanoni, S., Zavanella, L., ... & Petrillo, A. (2014). Disaster and emergency management simulation in industrial plants. In Proceedings of the 26th European Modeling and Simulation Symposium, EMSS 2014, pp. 649-656.
Chavez, R., Gimenez, C., Fynes, B., Wiengarten, F., & Yu, W. (2013). Internal lean practices and operational performance: The contingency perspective of industry clockspeed. International Journal of Operations & Production Management, 33(5), 562-588.
Dekker, S. (2016). Drift into failure: From hunting broken components to understanding complex systems. Boca Raton: CRC Press.
Diaz, R., Kumar, S., & Behr, J. (2015). Housing recovery in the aftermath of a catastrophe: Material resources Perspective. Computers & Industrial Engineering, 81, 130-139.
Gupta, S., Sharma, M., & Sunder M, V. (2016). Lean services: a systematic review. International Journal of Productivity and Performance Management, 65(8), 1025-1056.
Hadid, W., Mansouri, S. A., & Gallear, D. (2016). Is lean service promising? A socio-technical perspective. International Journal of Operations & Production Management, 36(6), 618-642.
Kenney, M., Rouvinen, P., & Zysman, J. (2015). The digital disruption and its societal impacts. Journal of Industry, Competition and Trade, 15(1), 1-4.
Kolberg, D., & Zühlke, D. (2015). Lean automation enabled by industry 4.0 technologies. IFAC-PapersOnLine, 48(3), 1870-1875.
Longo F., Nicoletti L., Padovano A. (2019). A system for supply chains diversification and (re)design: supporting managers perspective in the face of uncertainty. International Journal of Logistics Systems and Management. DOI: 10.1504/IJLSM.2019.10008306.
Marchwinski C. (2015). Toyota partners with nonprofit to rebuild homes and lives. Retrieved from: https://www.lean.org/common/display/?o=3016.
Mojtahedi, S. M. H., & Oo, B. L. (2012). Possibility of applying lean in post-disaster reconstruction: An evaluation study. In Proc., 20th Annual Conf. of the Int. Group for Lean Construction: Challenging Lean Construction Thinking: Are We near a Tipping Point (Vol. 1, pp. 411-420).
Murman, E., Allen, T., Bozdogan, K., Cutcher-Gershenfeld, J., McManus, H., Nightingale, D., ... & Warmkessel, J. (2002). Lean Enterprise Value: insights from MIT's Lean Aerospace Initiative. New York: Palgrave.
Nelson R.R., Winter S.G. (1982). An Evolutionary Theory of Economic Change, Harvard University Press, Cambridge, MA.
Ng, D., Vail, G., Thomas, S., & Schmidt, N. (2010). Applying the Lean principles of the Toyota Production System to reduce wait times in the emergency department. Canadian Journal of Emergency Medicine, 12(1), 50-57.
Ohno, T. (1988), Toyota Production System: Beyond Large Scale Production, Productivity Press, Portland, Oregon.
Seddon, J. (2003). Freedom from command and control. Vanguard Education Limited.
Shah, R., & Ward, P. T. (2003). Lean manufacturing: context, practice bundles, and performance. Journal of operations management, 21(2), 129-149.
Soliman, M., & Saurin, T. A. (2017). Lean production in complex socio-technical systems: A systematic literature review. Journal of Manufacturing Systems, 45, 135-148.
Summers, G. J., & Scherpereel, C. M. (2008). Decision making in product development: are you outside-in or inside-out?. Management Decision, 46(9), 1299-1312
Wang, P. P., Ming, X. G., Li, D., Kong, F. B., Wang, L., & Wu, Z. Y. (2011). Status review and research strategies on product-service systems. International Journal of Production Research, 49(22), 6863-6883.
Womack, J. P., & Jones, D. T. (2005). Lean Consumption. Harvard Business Review, 83(3), 58-68.
Zobel, C. W. (2014). Quantitatively representing nonlinear disaster recovery. *Decision Sciences*, 45(6), pp. 1053-1082.

Author's Index