MAS 2009

Puerto de la Cruz, Spain
September 23-25

Contents

WELCOME TO MAS 2009

Analytical approaches rarely succeed in identifying solutions for real world systems problems. When considering systems complexity, a number of dynamic relationships, stochastic variables and non linear interactions between critical factors and system behavior come up providing challenging problems for researchers and practitioners in different application domains.

The International Workshop on Modeling & Applied Simulation (MAS 2009) has reached its 8th edition, becoming an outstanding international meeting point for Modeling & Simulation worldwide community. As in previous editions, MAS 2009 will be held within the high quality (technical and scientific) framework provided by the 6th International, Mediterranean and Latin American Modeling Multi-conference (I3M 2009), including the European Modeling & Simulation Symposium (EMSS 2009) and the International Workshop on Harbor, Maritime, Multimodal Logistics Modeling & Simulation (HMS 2009).

Over the years, MAS has served as a showcase for providing evidence on the potentials of the Simulation approach applied to a wide spectrum of practical problems. Such practical focus is every year reflected by the scientific and scholarly relevance of selected papers as well as the high professional profile of MAS attendees. In addition, MAS 2009 settles an ideal framework to start new international co-operations and projects.

We wish that Tenerife will offer an incomparable framework for taking the most from the event as well as provide the attendees with an ideal place for enjoying wonderful beaches crowned by Mount Teide and surrounded by the Atlantic Ocean.

MAS Proceedings Editors,

Osman Balci, Virginia Tech, USA
Iván Castilla, ISAATC, Universidad de La Laguna, Spain
Francesco Longo, MSC-LES, University of Calabria, Italy
Marina Massei, Liophant Simulation Club, Italy

I3M Organizers

Proceedings of the International Conference on Modeling and Applied Simulation, MAS 2009
ISBN 978-84-692-5417-2

Academic Sponsors

Sponsors

Proceedings of the International Conference on Modeling and Applied Simulation, MAS 2009
ISBN 978-84-692-5417-2

TEACHING SIMULATION BASICS THROUGH FLOWCHART SIMULATION THE EVENT SCHEDULING WORLD VIEW

Guilherme A B Pereira[(a)], Luís M S Dias[(a)], Hugo T C Rocha[(a)]

[(a)]Departamento de Produção e Sistemas, Escola de Engenharia, Universidade do Minho, Portugal

[(a)]gui@dps.uminho.pt, lsd@dps.uminho.pt, Hugo.TC.Rocha@azores.gov.pt

ABSTRACT
This paper refers to Event Scheduling World View, focusing on working explicitly with the foundational concepts of discrete event simulation, acting as an automatic generator of simulation programs, thus eliminating any programming effort and expertise. The main strategy is to enhance the utilization of flowcharts in modeling. Therefore, a graphical support tool (Microsoft Visio) is used to represent how the system really behaves and will also act as the source for the automatic generation of Visual Basic (VB) simulation programs. The software tool VBS (Visio Basic for Simulations) was created to read Visio flowcharts, interpret them and generate a VB simulation program.

Keywords: discrete event simulation, event scheduling world view, automatic generation of simulation programs, flowchart simulation

1. INTRODUCTION

(Banks 1998) defines simulation as "… the imitation of the operation of a real-world process or system over time..." and (Chase et al 2006) "…as a "…computer-based model used to run experiments on a real system…" These definitions call for the creation of a model that represents the behaviour of real processes or systems under analysis.

In this context, (Carson 2003) argues that a "…simulation model is a representation of a system or process … incorporates time and the changes that occur over time [and] … a discrete model is one that changes only at discrete points in time…"

However, (Schriber and Brunner 2008) convincingly support that "…A "black box" approach is often taken in teaching and learning discrete-event simulation software…". In fact, as far as discrete event simulation is concerned, teaching and learning approaches usually tend to neglect the full comprehension of simulation basic concepts. Again, (Schriber and Brunner 2008) state that "…the foundation on which the software is based is ignored or is touched on only briefly … The modeler therefore might not be able to think things through when faced with such needs as developing good approaches for modeling complex situations".

Most authors like (Pidd 1992), (Nance 1993), (Bennet 1995), (Zeigler 1976), (Ziegler et al 2000), (Brito and Teixeira 2001), (Guasch et al 2002), (Overstreet and Nance 2004) and (Sargent 2004) would refer three classical simulation approaches in Discrete Event Simulation - Event Scheduling World View, Process World View and Activities World View.

In the research here presented the authors concentrate their work on Event Scheduling World View, focusing on helping students construct their simulation "house", brick by brick, i.e., working explicitly with the foundational concepts of discrete event simulation – events, entities, resources, queues, randomness, future events schedule, simulation time advance, etc…

Event Scheduling World View essentially represents the behaviour of a system over time by means of defining specific events to occur at discrete points in time – these events, planned and executed, would mimic the real system.

Even though with a different approach as shown in (Dias et al 2008), in this work the authors again use Event Scheduling concepts, acting as an automatic generator of simulation programs, thus eliminating any programming effort and expertise. Previously (Dias et al 2005) have presented a similar procedure, but based in activity cycle diagrams.

2. EVENT SCHEDULING CONCEPTS

Under Event Scheduling paradigm one would define an event as an instantaneous action that might change the state of the system (Guasch et al 2002).

A change in the state of a system would refer to the state of the resources of that system. Each time a resource state changes from busy to free or otherwise, one could say that the state of the system would also change.

Therefore, every instant in time where the state of a system might change would be defined as an instant where a specific event has occurred.

At that time, the tasks involved in that event would have to be performed.

These tasks would reflect not only the implementation of the change in the system, as far as resources, queues and flow of entities are concerned (representing physical modifications in the real system) but also logical changes reflecting the planning of future events, recording statistics for future use and also

generating random variables to model stochastic behaviour of the real system.

The main strategy of this research work is to enhance the utilization of flowcharts in modeling, thus making it a great contribution to automatic generation of simulation programs, keeping it simple. Therefore, a graphical support tool is used to represent how the system really behaves and will also act as the source for the automatic generation of simulation programs.

Flowcharts are probably the most widely used graphical syntax in behaviour specification (Gilbreth and Gilbreth 1921). The first known mathematical formalization was made by (Nassi and Shneiderman 1973). It can be accepted as a universal visual language, and it can be easily assumed that every professional, in some technical work, has already used it.

(Pidd 1992) and (Tocher 1963), even support the view that when generic programming languages were replaced by specific purpose simulation languages the use of paper diagrams remained as a previous step to programming.

In fact, this paper emphasizes the importance of this step, by proposing a way for automatically translating "paper diagrams" (flowcharts) into a simulation program.

3. EVENT SCHEDULING IMPLEMENTATION

The Event Scheduling simulation philosophy, as previously mentioned, is based upon the identification of events that, together, would represent the mimic of the system under analysis. The identification of each event is complemented with the definition of the tasks to be performed each time an event occurs. These tasks, as far as a discrete simulation approach is concerned, would include:

1. Managing queues (removing/inserting entities from/in queues)
2. Managing resources utilization (either seizing or releasing resources)
3. Recording statistics (for future evaluation of performance indexes, i.e., average waiting time in queue, average queue length, average resource utilization, etc.)
4. Generating random variables
5. Managing future events schedule and simulation time advance

Basic Simulation Facility – BSF (Thesen 1978), constitutes a way of implementing such a representation model for the behaviour of a real system over time. BSF is based on the computer programming concept of managing files. Moreover a file, as far as a simulation systems is concerned could represent:

1. the behaviour of a queue, where

- Inserting an entity in a queue would be represented through the insertion of a record on that file
- Removing an entity from a queue would be represented through the removal of a record from that file

2. the state of a resource, where

- seizing a resource would be represented through the insertion of a record in that file
- releasing a resource would be represented through the removal of a record from that file

3. the future events list, where

- planning a future event would be represented through the insertion of a record in that file
- executing an event would be represented through the removal of a record from that file

This type of implementation, using an appropriate data structure to accomplish the above features would also be useful to record statistics, using the mentioned files – these tasks would simply involve using some fields of the records of those files in order to register the statistical information needed.

BSF includes the following four routines already developed:

1. INIT – essentially dedicated to the design and initialization of the data structure that supports the simulation
2. INSERT – basically dedicated to the insertion of a record into a file (e.g. the arrival of an entity to a queue, or seizing a resource or even the planning of a future event)
3. REMOVE – basically dedicated to the removal of a record from a file (e.g. the removal of an entity from a queue, or releasing a resource or even preparing the execution of a future event)
4. REPORT – essentially dedicated to the computation of simulation performance measures

These routines, and the associated philosophy, could be found (implemented) in various programming languages (Java, C, VB, Pascal, Excel VBA, etc.). Nevertheless, it is essential to develop a computer program, specifically dedicated to the system under analysis, which would invoke these routines, thus creating a mimic of the system. The development of this computer program, together with the correct utilization of the aforementioned routines, is usually better described (modeled) by the use of appropriate flowcharts for each event identified.

Next section presents a software tool based on flowcharts, built and founded on key issues of the Event

Scheduling philosophy, which automatically generates a simulation computer program to mimic the real system.

4. SOFTWARE TOOL DEVELOPED

The relevancy of creating a simulation software tool capable of generating a simulation program (VBS – Visio Basic for Simulations) is based on the following general principles/premises:

1. event scheduling world view definitely contributes to a better understanding of the foundations of simulation, i.e., the basic concepts of simulation
2. VBS would emphasize the utilization of a visual approach to deal with the representation of the real problem through the event scheduling paradigm
3. VBS would literally enable the automatic generation of a simulation program (VBA or Java program) thus with no programming effort at all

VBS would then use the well known graphical editor Microsoft Visio for incorporating the event scheduling flowcharts representing each event. This task would be accomplished through the creation of Visio Shapes that would reflect the different tasks involved in each event identified. Then VBA (Visual Basic for Applications) would interact with Visio, interpreting the shapes and tasks associated, as well as the sequence of shapes to be "executed". At the end, this means that the effort to run simulation experiments through this tool would be equivalent to building event scheduling flowcharts on a piece of paper.

This was the challenge of the work presented in this paper.

For this purpose, and according to the ideas presented in section 3, the software tool would have to implement the tasks for each event. In fact, through VBS a flowchart for each event will be created and a main flowchart will coordinate the execution of each event flowchart.

For each type of task, the software tool includes the following Visio Stencils and the corresponding Visio Shapes:

- For managing queues (Figure **1**) – Shape Insert (inserting an entity in a Queue) and Shape Remove (removing an entity from a queue). Additionally there is a shape for queue declaration.

  Parameters: Queue identification

Figure 1: Queue Stencil

- For managing resources utilization (Figure **2**) - Shape Seize (seizing a resource) and Shape Release (releasing a resource). There is also a shape for declaring each resource.

  Parameters: Resource identification

Figure 2: Resource Stencil

- For managing future events schedule and simulation time advance (Figure **3**) – Shape Plan (planning future event to occur, inserting information on the Future Events List), Shape Identify (identifying next event to occur, reading information from the Future Events List) and Shape Execute (transferring the execution of the program to the corresponding flowchart of the next event to be executed).

  Parameters: Identification of next event to occur and future instant for that event to occur

Figure 3: Event Stencil

- For generating random variables (Figure **4**) – Shape Random, generating a value from a statistical distribution with the respective distribution parameters.

  Parameters: Statistical Distribution and corresponding distribution parameters

Figure 4: Statistics Stencil

In Event Scheduling World View there is frequent need to identify if:

- a resource is free
- a queue is empty

The tool developed includes a shape for each of the situations above – see Figure **5**. The parameters include, respectively, the identification of the resource and the queue.

Figure 5: Decision Stencil

Additionally, every shape needs to be linked to other shapes. There are also a set of Link shapes for that purpose – see Figure **6**. These include the straight link between shapes, but also output decision links from decision shapes – True and False outputs following a decision point.

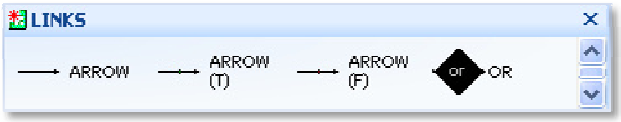


Figure 6: Links Stencil

Finally, and for software implementation purposes, there are also shapes for activating the main program as well as for activating each event. At the end of each event – after executing all the tasks for that event, a Return shape is used to direct program execution back to the main flowchart. This flowchart, at the end of the simulation, uses the shape Report to compute some statistical indicators for evaluating the system performance – Figure **7** below.

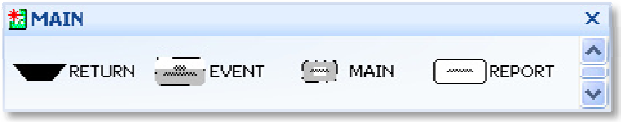


Figure 7: Main Stencil

The tool developed (VBS) promotes the interaction between Visio shapes and VBA (Visual Basic for Applications), that interprets each shape and corresponding parameters and generates computer code accordingly. Figure 8 and Figure **9** show general screenshots of the VBS application. At the end, a whole computer program is generated, compiled and executed.

Next section will present an application example, the corresponding interaction with the software tool and the computer code (VB) generated.

5. APPLICATION EXAMPLE

For a brief explanation of the tool developed, a simple example is presented.

This example represents a system that incorporates the arrival of entities. Once in the system, the entities would require the use of a first resource (*Resource_A*). Then, the entity would also require the use of a second resource (*Resource_B*). There is a first-in first-out type of queue associated with each resource. The entity will then leave the system.

The Event Scheduling philosophy would identify three events – Entity Arrival Event (Figure **10**); End of *Resource_A* Utilization (Figure **11**); End of *Resource_B* Utilization (Figure **12**). Figure **13** represents the main flowchart that integrates and coordinates the behaviour of each event occurring in the system.

Following each flowchart is the respective code, automatically generated by VBS, based on each shape and respective parameters.

VBA, by reading each flowchart, would recognize each shape and would interpret each parameter of each shape. Computer code generation would then follow in accordance.

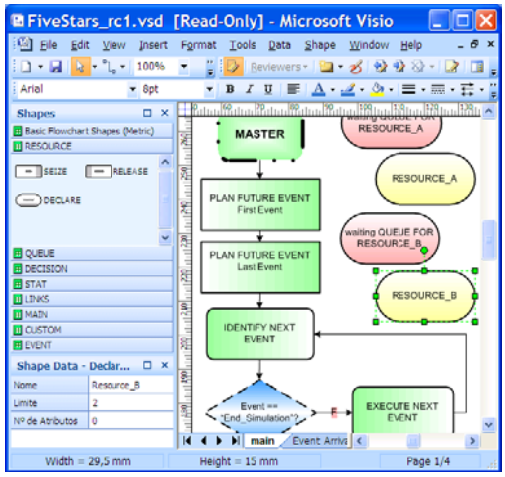


Figure 8: VBS Screenshot 1

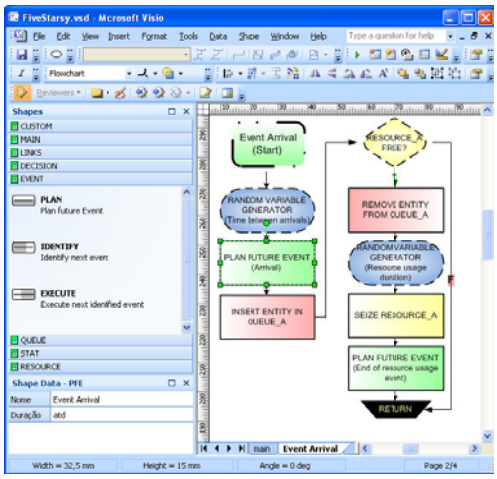


Figure 9: VBS Screenshot 2

For a better understanding of the relation between each shape and the corresponding computer code, let us follow the flowchart for Event Arrival (Figure **10**): first shape corresponds to the activation of the respective flowchart, meaning a call to the routine Sub Arrival. Then the shape Random generates line code *T=RVG("EXPO(5)")* – this means that the parameters of this shape were recognized as having an Exponential type of distribution with a mean of 5 (minutes). Shape Plan follows, indicating the planning of a future event – inserting a record in the future events list – line code *Insert("FutureEvents", Time+T, "Arrival")*. The entity that has just arrived in the system will be part of the

corresponding queue – line code *Insert ("QUEUE_A",Time)*.

Then, if *Resource_A* is free (next shape), the following actions need to be performed:

- removing entity from the queue
- generating duration of *Resource_A* utilization
- seizing *Resource_A*
- planning next event – end of *Resource_A* utilization
- end of flowchart

The above actions are equivalent to the following code generated by the application:

Table 1: Partial code for Event Arrival

```
If Size("RESOURCE_A")<Max_A Then
  Remove("QUEUE_A")
  T=RVG("normal(4,1)")
  Insert("RESOURCE_A",Time+T)
  Insert("FutureEvents", Time+T,
     "End of Resource_A Utilization")
End If
```

Having performed these steps, the full code for this flowchart is shown in Table **2**.

Table 2: Code for Event Arrival

```
Sub Arrival()
  T=RVG("EXPO(5)")
  Insert("FutureEvents", Time+T, "Arrival")
  Insert("QUEUE_A",Time)
  If Size("RESOURCE_A")<Max_A Then
    Remove("QUEUE_A")
    T=RVG("normal(4,1)")
    Insert("RESOURCE_A",Time+T)
    Insert("FutureEvents", Time+T,
       "End of Resource_A Utilization")
  End If
End Sub
```

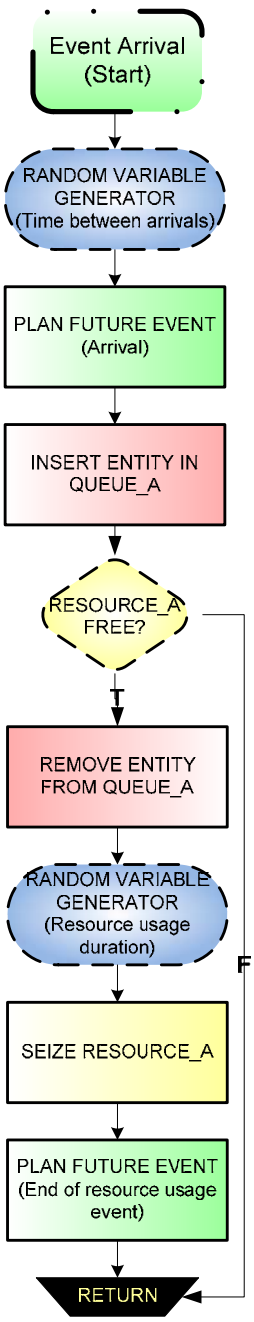


Figure 10: Entity Arrival Event Flowchart

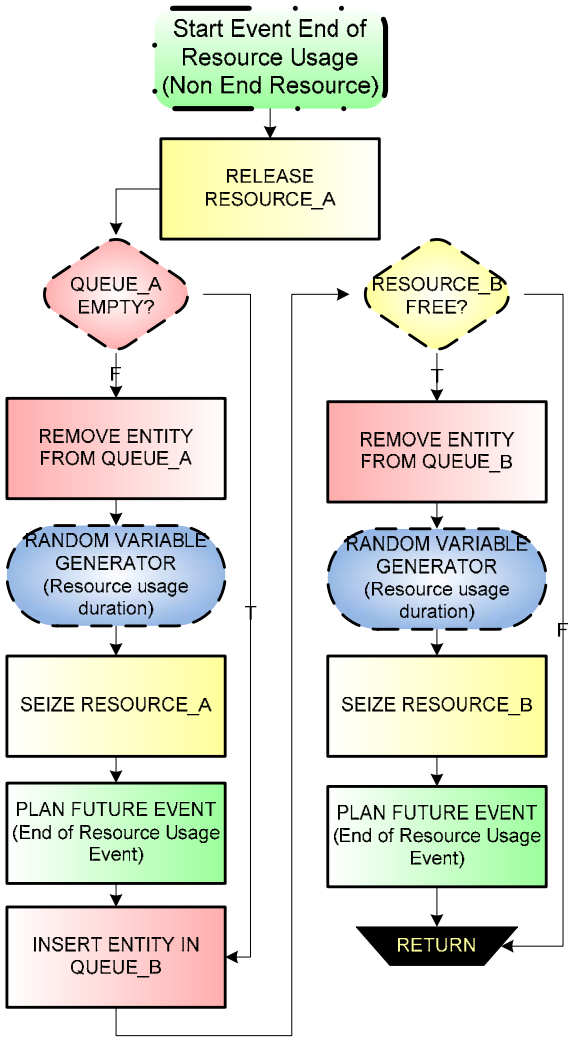


Figure 11: End of *Resource_A* Utilization Event Flowchart

Similar procedures would generate computer code of Table **3** for event represented in Figure 11.

Table 3: Code for Event End of *Resource_A* Utilization

```
Sub End_of_Resource_A_Utilization()
  Remove("RESOURCE_A")
  If Size("QUEUE_A")>0 Then
    Remove("QUEUE_A")
    T=RVG("normal(4,1)")
    Insert("RESOURCE_A",Time+T)
    Insert("FutureEvents", Time+T,
       "End of Resource_A Utilization")
  End If
  Insert("QUEUE_B",Time)
  If Size("RESOURCE_B")<Max_B Then
    Remove("QUEUE_B")
    T=RVG("normal(18,2)")
    Insert("RESOURCE_B",Time+T)
    Insert("FutureEvents", Time+T,
       "End of Resource_B Utilization")
  End If
End Sub
```

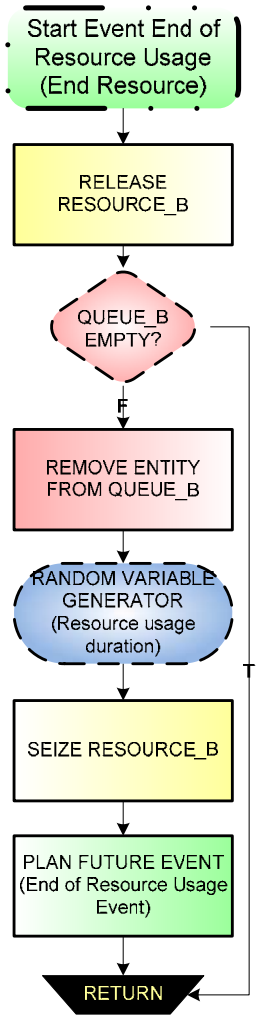


Figure 12: End of *Resource_B* Utilization Event Flowchart

Also for flowchart of Figure 12, the following code is generated.

Table 4: Code for Event End of *Resource_B* Utilization

```
Sub End_of_Resource_B_Utilization()
  Remove("RESOURCE_B")
  If Size("QUEUE_B")>0 Then
    Remove("QUEUE_B")
    T=RVG("normal(18,2)")
    Insert("RESOURCE_B",Time+T)
    Insert("FutureEvents", Time+T,
       "End of Resource_B Utilization")
  End If
End Sub
```

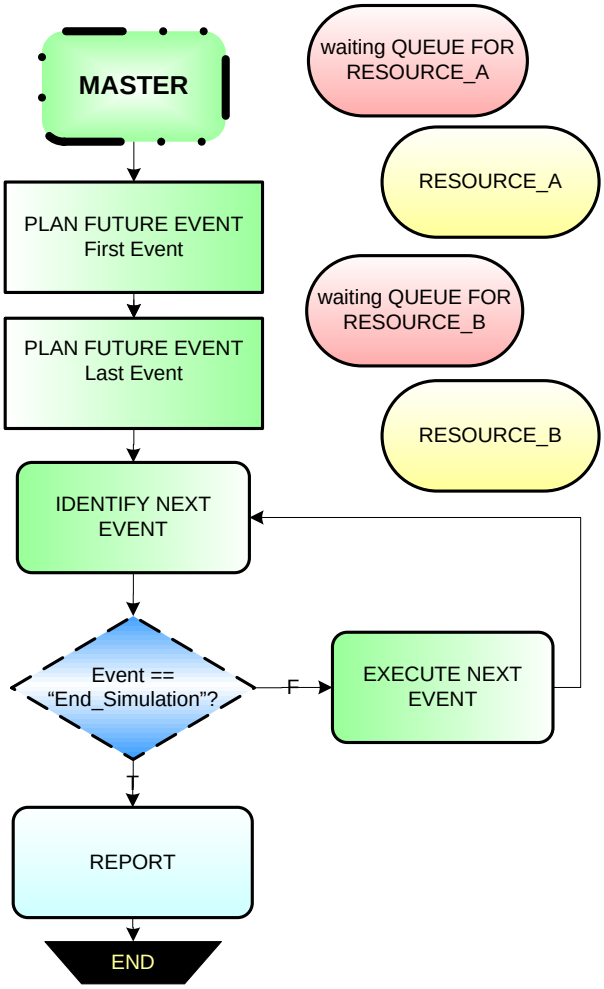


Figure 13: Main Flowchart

Finally for flowchart of Figure **13**, the code included in Table **5** is generated.

Table 5: Code for Main Flowchart

```
Main Program:
  INIT()
  Max_A=1
  Max_B=4
  Insert("FutureEvents", 0, "Arrival")
  Insert("FutureEvents",1000,"End_Simulation")
  DO
    Event=Remove("FutureEvents")
    If Event == "Arrival" Then Call Arrival
    If Event == "End of Resource_A Utilization"
       Then
           Call End_of_Resource_A_Utilization
    If Event == "End of Resource_B Utilization"
       Then
           Call End_of_Resource_B_Utilization
  UNTIL event == "End_Simulation"
  Call Report
```

This computer program is then ready for compiling and executing. At the end, the usual performance indicators are available – average time spent in a queue, average time spent in the system, average queue length, average number of resources busy, etc.

6. CONCLUSIONS

The software tool developed shows three particularly interesting features, namely:

- It is based on simple flowcharts that follow Event Scheduling Simulation philosophy
- It automatically generates a VB computer program to perform the mimic of the system under analysis
- It runs the model directly over the flowcharts, producing debugging trace files

These features, together, would contribute to

- the generalization and a better understanding of the use of simulation
- the comprehension of the foundations of simulation
- the automatic generation of simulation programs

In brief, it can be argued that the generalization and better understanding of the use of simulation would have been accomplished since the tool herein developed only requires i) expertise on a basic simulation approach: event scheduling, ii) incorporating simple flowcharts that define the system and its functioning rules.

Furthermore, these flowcharts, apart from providing an understanding of the system's behaviour, contribute to the comprehension of fundamental simulation concepts, such as entities, queues, resources and also to a very important simulation concept – the evolution of the state of the system over time.

Finally, these simple flowcharts are translated into the software tool by means of an automatic generation of a computer program that performs the mimic of the system and evaluates the corresponding efficiency measures. The simulation runs over the events' flowcharts, step by step, enabling the user to gradually assimilate concepts while validating his learnings.

REFERENCES

Banks, J., 1998. Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice, John Wiley and Sons, Inc.

Banks, J., 2000. Introduction to Simulation, in Proceedings of the 2000 Winter Simulation Conference, eds., J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, pp. 9-16.

Bennett, B. S., 1995. *Simulation Fundamentals*, Prentice Hall, ISBN 0-13-813-262-3

Brito, A. E. S. C. and Teixeira, J. M. F., 2001. *Simulação por Computador – Fundamentos e implementação de código em C e C++*, Publindústria - Edições Técnicas, ISBN 972-98726-2-7

Carson, J. S., 2003. Introduction to Modeling and Simulation, Proceedings of the 2003 Winter Simulation Conference, eds., S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, pp. 7-13

Chase, R. B., Jacobs, F. R. and Aquilano, N. J., 2006. Operations Management for Competitive Advantage, McGraw-Hill/Irwin.

Dias, Luís M. S., Rodrigues, António J. M. G. and Pereira, Guilherme A. B., 2005. An Activity Oriented Visual Modelling Language with Automatic Translation to Different Paradigms, Proceedings of the 19th European Conference On Modelling And Simulation (ECMS 2005), Riga, Letónia. Ed. Yury Mercuryev et al, pp. 452-461

Dias, Luis M. S., Pereira, Guilherme A. B. and Oliveira, José A. V., 2008. Event Scheduling Made Easy: Basic Simulation Facility Revisited, in Proceedings of the European Modeling and Simulation Symposium, EMSS 2008, Lamezia, p. 688-692.

Gilbreth F.B. and Gilbreth L.M., 1921. Process Charts - First Steps in Finding the One Best Way to do Work. Presented at the Annual Meeting of The American Society of Mechanical Engineers, New York, USA.

Guasch, A., Pierra, M., Casanovas, J. and Figueras, J., 2002. Modelado y Simulación – Aplicación a procesos logísticos de fabricación y servicios. Edicions de la Universitat Politècnica de Catalunya, Barcelona.

Nance, R. E., 1993. A history of discrete event simulation programming languages, in the Second ACM SIGPLAN Conference on History of Programming Languages (Cambridge, Massachusetts, United States, April 20 - 23). HOPL-II. ACM Press, New York, NY, pp 149-175. DOI=10.1145/154766.155368

Nassi and Shneiderman, 1973. Flowchart techniques for structured programming, ACM SIGPLAN Notices, 12.

Overstreet, C. M. and Nance, R. E., 2004. Characterizations and Relationships of World Views, in *Proceedings of the 2004 Winter Simulation Conference*, ed. R .G. Ingalls, M. D. Rossetti, J. S. Smith and B. A. Peters, pp 279-287, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Pidd, M., 1992. Computer Simulation in Management Science. 3rd Ed. John Wiley & Sons, Inc.

Sargent, R. G., 2004. Some recent advances in the process world view, *Winter Simulation Conference*, pp 293-299.

Schriber, T. J. and Brunner, D. T., 2008. Inside discrete-event simulation softe«ware: how it works and why it matters, in *Proceedings of the 2008 Winter Simulation Conference, eds. S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson and J. W. Fowles.*

Thesen, A., 1978. Computer Methods in Operations Research, New York: Academic Press.

Tocher, K. D., 1963. The Art of Simulation, UNIBOOKS – English Universities Press.

Zeigler, B. P., 1976. *Theory of Modeling and Simulation*, Wiley, New York.

Zeigler, B. P., Herbert, P. and Tag, G. K., 2000. *Theory of Modeling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems*, 2nd edition, Academic Press.

AUTHORS BIOGRAPHY

Guilherme A B Pereira was born in 1961 in Porto, Portugal. He graduated in Industrial Engineering and Management in the University of Minho, Portugal. He holds an MSc degree in Operational Research and a PhD degree in Manufacturing and Mechanical Engineering from the University of Birmingham, UK. His main research interests are Operational Research and Simulation.

Luís M S Dias was born in 1970 in Vila Nova de Foz Côa, Portugal. He graduated in Computer Science and Systems Engineering in the University of Minho, Portugal. He holds a PhD degree in Production and Systems Engineering from the University of Minho, Portugal. His main research interests are Operational Research, Simulation and Systems Visual Modeling.

Hugo T C Rocha was born in 1975 in Aveiro, Portugal. He graduated in Computer Engineering in the University of Minho, Portugal. During his degree he developed some research work in simulation. He currently works in the Center of Informatics for Finance of the Regional Government of Açores, Portugal and he is a professor of Information and Communication Technologies in the Professional School of Santa Casa da Misericórdia dos Açores. His main research interests are Informatics and Simulation.

BLOOD FLOW SIMULATION OF INTRACRANIAL ANEURYSMS

Wolfgang Fenz, Johannes Dirnberger[(a)]

[(a)] RISC Software GmbH, Research Unit Medical-Informatics, Softwarepark 35, 4232 Hagenberg, Austria

[(a)]{wolfgang.fenz | johannes.dirnberger}@risc.uni-linz.ac.at

ABSTRACT

In the course of the research project MEDVIS 3D (www.medvis3d.at), a clinical software application has been developed, capable of 3D reconstruction and visualization of intracranial aneurysms based on 2D medical image data. This system is now being extended with the functionality of a physically correct simulation of blood flow through aneurysms. It shall provide means to estimate rupture risks by calculating the distribution of pressure and shear stresses in the aneurysm, in order to support diagnosis and intervention planning. Due to the time-critical nature of the application, we are forced to use the most efficient state-of-the-art numerical methods and technologies. The elasticity equations for vessel walls and Navier-Stokes equations for blood flow are discretized via the Finite Element Method (FEM), and the resulting linear equation systems are handled by an Algebraic Multigrid (AMG) solver. The boundary conditions of both fluid and wall domains are coupled via Fluid-Structure Interaction (FSI) algorithms. In order to minimize computation time, the simulation will also be parallelized and distributed in the Austrian Grid network. First results using commercial Computational Fluid Dynamics (CFD) packages already show good medical relevance for diagnostic decision support. Our goal is to replace commercial modules step by step by our own implementations in order to end up with a license-free simulation system that is available for every hospital.

Keywords: computational fluid dynamics, flow simulation, fluid structure interaction, 3D visualization

1. INTRODUCTION

The various forms of vascular diseases are complex in their origins and their manifestations. In order to diagnose, prevent and treat such diseases a detailed knowledge of blood flow is essential. Vessels respond to stimuli and adapt to changes in blood flow and blood pressure. Hemodynamic (blood flow mechanic) factors are strongly correlated with the localization of atheriosclerotic plaque.

Atherosclerosis is primarily responsible for strokes, heart attacks and aneurysms, which are among the most frequent causes of death in western countries. In many cases the time of reaction is an important factor in the treatment of these diseases and is crucial for the convalescence of the patient. Due to atherosclerosis it is often necessary to dilate vessels to improve the blood flow. Stents are used there to support the vessel wall after dilation. We see a major application area in the simulation of blood flow in detected aneurysms and stent supported vessels.

The decision of the physician, which stent to use and where to place it, can be supported by the resulting flow patterns. As a future outlook also the stent design and stent optimization can be improved by simulating the hemodynamics. A simulation of the hemodynamics in a 3D geometry based on medical imaging should focus on the following:

- Correct modeling of the vascular system in 3D and an easy to handle tool for the physicians to segment the regions of interest and to insert and design stents and coils in the 3D environment.
- Fast and robust calculation of the hemodynamics in the relevant area.
- Visualization of the resulting flow patterns in a way that it can support the physician's decision without being an expert in mathematics and/or physics.

A decision support tool for the physician can therefore play a major role for recovery. To assure the relation to practice, a close participation of clinical partners during the whole project is intended.

2. METHODS

After having generated a new 3D reconstruction, segmentation and visualization library named REVOLTE (Baumgartner 2006; Thumfart 2007), the main focus of our project is the implementation of a realistic and physically correct simulation model of the blood flow through the visualized vessel geometry.

The geometry information gained by iso-surface extraction will be put into a computer simulation together with blood and tissue parameters, building up a virtual model of the patient's vascular system. Starting

with measured blood velocity, a simulation procedure will initiate and calculate pressure and vessel deformation of the whole system at discrete time steps. The results have to be visualized in 3D, in order to quickly provide precious diagnostic information for the assessment of the aneurysm's rupture risk.

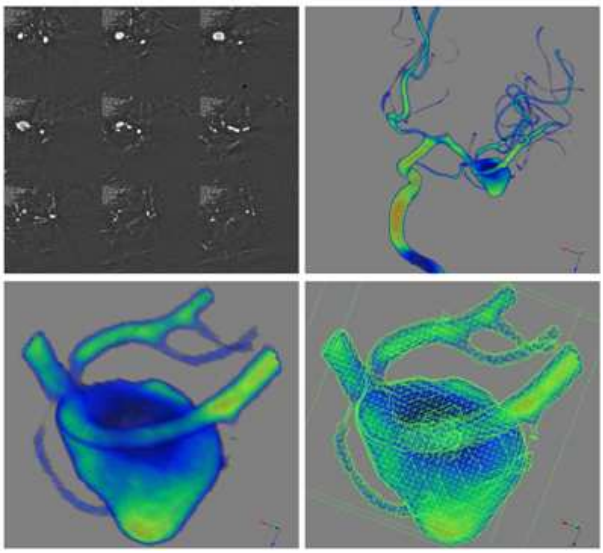


Figure 1: Biplane MRI data, 3D visualization, segmented aneurysm, surface triangulation generated by the skeleton climbing algorithm.

2.1. Mesh Generation

Commonly used methods in medical imaging are Magnetic Resonance Imaging (MRI), X-Ray Angiography and Computer Tomography (CT). All these methods provide series of cross-section pictures of the patient. The software *MEDVIS 3D* is able to reconstruct the 3D volume out of these cross sections and subsequently generate a finite element mesh of a semi-automatically selected region (volume of interest; see figure 1). For surface meshing the skeleton climbing algorithm is used, which is superior to commonly used marching cube schemes, whereas for the creation of tetrahedra we apply Delaunay triangulation.

The quality of both meshes will be further improved by applying area weighted Laplacian smoothers (2D) and iterative mesh improvement (3D).

2.2. Simulation

The simulation of a number of processes in engineering and nature requires the coupled solution of different physical problems. In the present research the fluid-structure interaction and its application to hemodynamics in the human vascular system is taken into account. The interaction between the fluid (blood) and the structure (vessel wall) takes place only at the interface. In large vessels, blood is usually modeled as a Newtonian fluid using Navier-Stokes equations. Due to the fluid pressure and shear stress, forces appear on the structural boundary. The vessel wall displacement due to these forces is small, therefore linear elasticity equations (Navier equations) are used for modeling the vessel wall. The vessel walls are homogeneous and we assume isotropic coefficients for simplicity of presentation. The solution of each subproblem is incorporated into the boundary conditions on the interface for the other subproblem. The fluid forces have to be taken into account within the structural dynamic subtask and the fluid domain has to be updated in a way that it fits to its new boundaries determined by the structural displacements. This interaction is non-linear and an analytical solution of Fluid-Structure-Interaction (FSI) problems is not possible at all in practical cases, therefore the FSI task has to be solved numerically.

Of course, the numerical solution of the coupled problem requires the numerical solution of the fluid and the structural subtasks. Although in the past decades a number of efficient solution methods have been developed for both subtasks, the development of efficient solvers for the FSI problem is still a challenge.

Our targeted method of choice is the stabilized loose coupling (Sieber 2001) which uses an iterative predictor-corrector scheme to guarantee the energy conservation and the equilibrium between the fluid and the structure. It acts as a compromise between fully implicit coupling algorithms that ensure exact energy conservation and mechanical equilibrium at the price of high computational demands and the so-called loose coupling or explicit methods that are faster but numerically unstable. All these coupling methods belong to the group of partitioned approaches, where the fluid and structure problems can be solved separately by specialized and highly efficient Computational Fluid Dynamics (CFD) and Computational Structure Dynamics (CSD) solvers, respectively, as opposed to the monolithic approach where the subproblems are reformulated into a coupled equation system and solved simultaneously.

2.2.1. Solving Subproblems

The vessel wall is described by a linear elasticity equation (Navier equation), which is given by

$$(\lambda + \mu)\nabla(\nabla u) + \mu\Delta u + f = \rho\frac{\partial^2 u}{\partial t^2} \qquad (1)$$

where $\mathbf{u}(\mathbf{r},t)$ denotes the structural displacement at coordinate $\mathbf{r}$ and time t, ρ is the density of the wall tissue, λ and μ are the so-called Lamé parameters describing its elastic properties, and $\mathbf{f}$ is an external force. As a first step, the wall is modeled as a shell with displacements only in radial direction, ignoring the viscous shear stress acting on the wall. This can later be replaced by a finite thickness model, which, however, requires the medical images to be of high quality (for example, MRI images do not include information on the vessel walls).

For the fluid domain, we consider the Navier-Stokes and the mass continuity equation,

$$\rho\left[\frac{\partial v}{\partial t} + (v\nabla)v\right] - \mu\Delta v + \nabla p = f, \quad (2)$$

$$\nabla v = 0, \quad (3)$$

describing the velocity **v**(**r**, t) and pressure field **p**(**r**, t) of an incompressible Newtonian fluid with density ρ and viscosity μ. Both the fluid and structure subtasks are solved numerically by application of the Finite Element Method, which provides a discretization of the corresponding fields for arbitrary geometries, and replaces the partial differential equations with large systems of linear equations of the general form

$$Ax = b, \quad (4)$$

where A is a sparse matrix. Such equations are usually solved by iterative methods, most notably the so-called Krylov Subspace Methods such as Conjugate Gradient. The use of a preconditioner which transforms (4) into an equivalent system that is easier to solve can speed up the calculation considerably.

One quite recently developed method is Algebraic Multigrid described in Stüben (2001). It can be used either as a standalone solver or as a preconditioner, is fast and robust, but it requires the coefficient matrix A to be symmetric and positive definite.

For the Navier equation, this is automatically fulfilled. In the case of the Navier-Stokes equation, however, a splitting algorithm has to be applied first. Using the technique of Haschke and Heinrichs (2001), instead of one non-linear equation coupling velocity and pressure, four uncoupled equations of Helmholtz and Poisson type have to be solved, all of which yield a symmetric and positive definite system matrix after discretization.

As boundary conditions to the simulated system we assume a no-slip condition of the blood on the vessel walls and take the velocity variation of the incoming blood stream during one pulse cycle from measurements performed by our medical partners via Doppler ultrasound.

2.2.2. Current Implementation Status

For the 3D blood flow simulation, work on a FEM Navier-Stokes solver has started. It takes as input an arbitrary unstructured tetrahedral mesh in a simple file format, which can for example be exported by the open source mesh generator NETGEN (Schöberl, 1997), and reads further parameters such as fluid density and viscosity from a separate data file. Boundary conditions for each domain of the mesh surface (as defined in the mesh file) have to be provided too.

Our code uses Taylor-Hood elements with quadratic shape functions (10 nodes per tetrahedron) for the velocity and linear shape functions (4 nodes per tetrahedron) for the pressure. The unsteady incompressible Navier-Stokes (NS) equations are solved either in their coupled form using Oseen iterations or via the splitting algorithm after Frochte and Heinrichs (2009). In the first case, the coefficient matrix of the system of linear equations obtained via the FEM discretization is asymmetric, and thus one has to use a Biconjugate Gradient method to solve it. In particular, we are applying the so-called BiCGStab algorithm. In the other case, the original partial differential equations are replaced by three Helmholtz equations for the velocity components and a Poisson equation for the pressure. All of them yield symmetric and positive definite coefficient matrices, and therefore the Conjugate Gradient method combined with some preconditioner is applicable. Apart from a simple diagonal preconditioner we are using the Algebraic Multigrid (AMG) method. Our project partner Prof. Gundolf Haase of the Institute for Mathematics and Scientific Computing at the Karl-Franzens-University Graz provided the so-called Parallel Toolbox which includes an AMG routine. We have incorporated a parallel version of it into our solver, but have only used it with one CPU so far. In order to efficiently parallelize the AMG/CG iterations, some partitioning of the mesh has to be provided, which we have not produced yet.

Gradients of pressure and velocity occurring in the Helmholtz and Poisson equations are calculated via the Taylor based gradient recovery technique which was also introduced by Frochte and Heinrichs (2009).

Other tasks that have already been started but are not yet finished are the implementation of Streamline Diffusion in order to improve numerical stability for convection-dominated problems and the integration of mesh smoothing algorithms.

2.3. Visualization

In contrast to the abstract mathematical results of vascular flow simulation the presentation of these results in a medical application should make the figures ready for interpretation and verification tasks. Therefore, the intuitive graphical visualization of flow patterns and turbulences within the reconstructed 3D structures of the patient's morphology plays an important role for the acceptance and deployment of computational simulation systems in medical environments. Together with our medical partners we will develop blood flow visualizations which allow physicians to easily explore a patient's vascular system and to locate parts of enhanced risk. Displaying results in real-time will require fast and flexible strategies in the field of computer graphics, and we have to extensively exploit the capabilities of modern hardware.

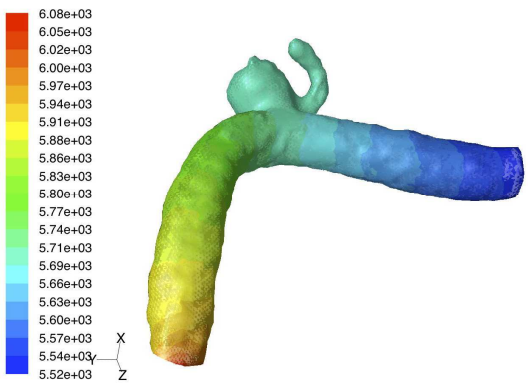


Figure 2: Total pressure of sample data calculated with FLUENT/ABACUS.

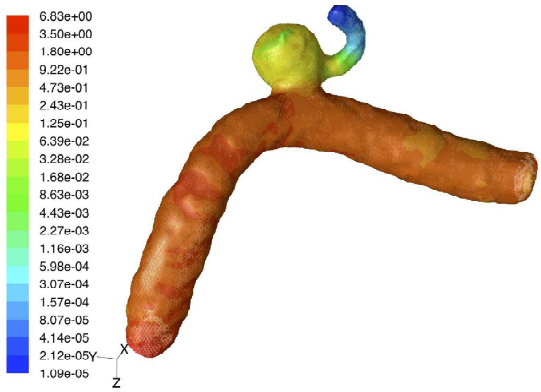


Figure 3: Wall shear stress of sample data calculated with FLUENT/ABACUS.

3. RESULTS AND DISCUSSION

So far, we have calculated the blood flow with our simulation module implemented in C++. Based on geometry data of a cerebral aneurysm measured on a human patient via magnetic resonance imaging (MRI), a 3D mesh was generated and velocity and pressure fields for one pulsatile cycle have been computed.

As a verification basis, on the same geometry velocity, pressure and wall shear stress have been obtained with *Fluent* (CFD) and *Abaqus* (CSD). Moreover, an explicit coupling algorithm providing the Fluid-Structure interaction was implemented by our project partner dTech Steyr GmbH, which will allow further verification of vessel wall displacements due to the blood flow in the future.

3.1.1. Implementation

As a starting point for our implementation, we have mainly studied simple aneurysm model geometries consisting of a cylindrical part representing the blood vessel and a spherical part for the aneurysm. In order to simulate pulsatile inflow, a parabolic velocity profile with periodically varying peak value is used for the boundary condition of the incoming fluid on one end of the cylinder. As outflow condition we set the pressure on the other end equal to zero, while at the vessel walls a no-slip condition is applied. Starting with a fluid at rest, the peak inflow velocity is gradually increased for the first one or two periods, until a stable oscillation is reached. Such a procedure was found to be more efficient than calculating an initial flow configuration with finite inflow velocity with the steady NS equation, especially for higher Reynolds numbers (Re $\geq$ 100) where convergence proved to be difficult. We chose the period of the oscillating inflow velocity with 1s, and used time steps of $\Delta t = 0.01$s. Each time step took about one minute of CPU time on a single core of an Intel Core2 Duo Q6700.

We have also compared the results of our solver with calculations performed with *Fluent* (version 6.3) on the same mesh, using the same boundary conditions and parameters. For the results obtained with the coupled NS equations we observed a good agreement, whereas for the splitting algorithm there were some deviations that are probably due to problems with the implementation of the Neumann boundary conditions for the pressure equation. We are still working on that part of our code.

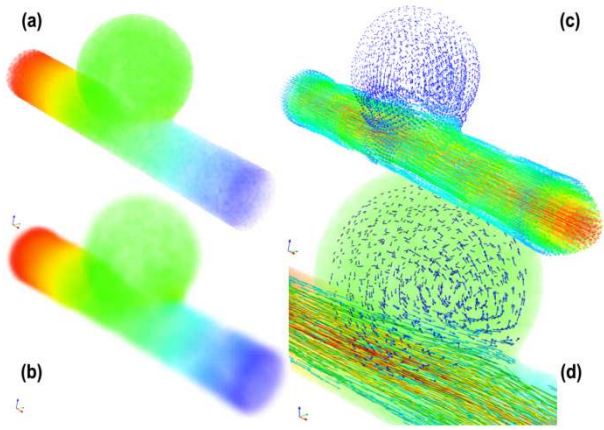


Figure 4: Visualization of pressure [(a), (b)] and velocity field [(c), (d)] for the simplified model geometry, with a peak inflow velocity of 1.5 m/s and Re=100 accomplished with our visualization module.

3.1.2. Benchmarking

For obtaining a verification data set, we have computed the blood flow based on the constant density Navier-Stokes equation in commercial simulation toolkits (*Fluent* and *Abaqus*). The Reynolds number based on the vessel diameter for the intracranial aneurysm presented in figure 5 and figure 6 is about 200. Hence, the blood flow can be considered laminar (e.g. Finol et al. 2003). Nevertheless, at some larger vessel regions the Reynolds number may exceed the critical value of about 2300, which defines the transition between laminar and turbulent flow regimes. Turbulence models based on Reynolds averaging are valid only for fully turbulent flows and hence, not applicable. An alternative consists of using a Large Eddy approach, whereas a huge increase of computational effort has to be accepted. With regard to the project aims we will neglect turbulence considerations.

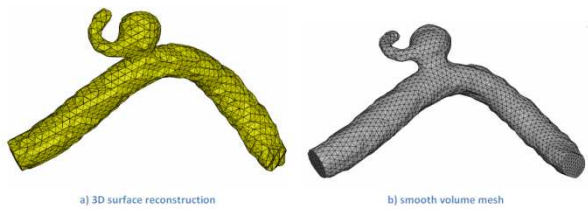


Figure 5: Comparison of 3D surface reconstruction and the resulting flow volume mesh of an intracranial aneurysm

It is common to consider blood as Newtonian fluid (e.g. Finol et al, 2003), Scotti and Finol, 2007)). However, blood is a suspension of red blood cells in

plasma, whereas the viscosity of the blood depends on the volume fraction of the red blood cells in plasma. A possible approach to take into account the inhomogeneity of blood is to use the Herschel-Bulkley fluid model (e.g. Valencia et al. 2006). In this work we use the Newtonian material law because of a more stable behavior and the smaller set of parameters compared to the Herschel-Bulkley model.

To incorporate the coupling effects between the blood flow and the vessel wall, a coupling interface between the CFD software Fluent and the FE code Abaqus has been developed. Here, we assume a local equilibrium between the blood flow and the deformation of the vessel wall. From the CFD simulation we obtain the total pressure and shear forces acting on the vessel wall. By using the coupling interface, these forces can be utilized as boundary conditions for the structural analysis. The resulting deformations are passed back to the CFD simulation and applied to the CFD mesh. Furthermore, it may be necessary to remesh the blood flow volume due to the deformations of the vessel wall, which is achieved by a dynamic mesh algorithm.

To stabilize the fluid structure interaction simulation, we use a subcycling algorithm: i.e. a specific number of CFD iterations (Fluent) are executed between two succeeding structural static finite element iterations (Abaqus).

In figure 6 the displacements and the von Mises stresses caused by the total pressure and shear forces exerted by the blood flow onto the vessel wall for the intracranial aneurysm from Figure 5 are shown. It can clearly be seen that the pressure maximum at the aneurysm causes a peak in the stress distribution.

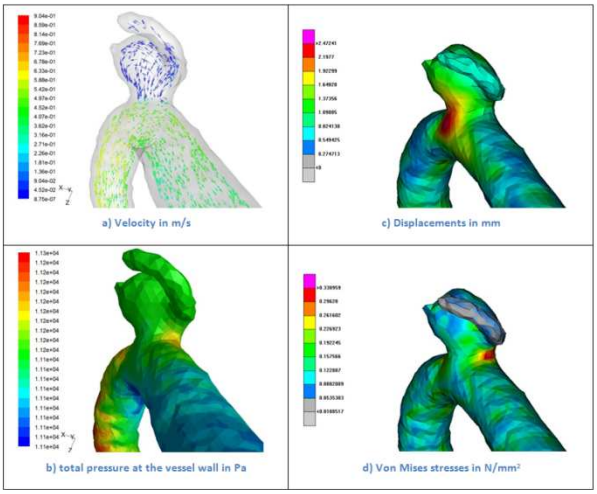


Figure 6: Results of the Fluid Structure Interaction Problem using the example of an intracranial aneurysm.

4. OUTLOOK

As a next step, we are planning to replace the commercial modules by open-source or self-developed libraries.

Moreover, the results obtained in the simulations have to be checked and verified. Forthcoming simulations will be performed on in-vivo measured aneurysm geometries using much larger meshes, and applying as a boundary condition the velocity variation of the incoming blood stream during one pulse cycle, measured on the patient by our medical partners via Doppler Ultrasound.

Besides the comparison to other published results it is planned to compare the calculated flow velocities to the velocities measured by phase contrast MRI or Doppler Ultrasound. Furthermore, we plan to follow the Lagrangian interpolation method introduced by Cheng et al. (2002) to get data about the wall shear stress, the oscillatory wall shear stress and the cyclic strain to compare calculated values with in-vivo measurements. As an outlook, one can also think about simulating stenting, coiling and clipping of aneurysms by a change in the geometry (additional mesh from the stent with appropriate boundary condition to ensure the smoothness and regularity) and changes in the material parameters.

ACKNOWLEDGMENTS

This research project is funded by the Austrian Research Promotion Agency (www.ffg.at) in the frame of the "MODSIM Computational Mathematics" program.

The project board would like to thank Prof. Gundolf Haase (University Graz) and Prof. Heinrich Langer (University Linz) for their valuable input and remarks on our research activities and results.

REFERENCES

Baumgartner S, 2006. Volumenvisualisierung von MR-Daten zur Behandlungsunterstützung bei Schlaganfällen. *Master Thesis*, Institut für Graphische und Parallele Datenverarbeitung, Johannes Kepler University Linz

Cheng CP, Parker D, CAT, 2002. Quantification of wall shear stress in large blood vessels using Lagrangian interpolation functions with cine phase-contrast magnetic resonance imaging. *Ann Biomed Eng*. 30, 1020

Finol EA, Keyhani K and Amon CH, 2003. The Effect of Asymmetry in Abdominal Aortic Aneurysms Under Physiologically Realistic Pulsatile Flow Conditions. *Journal of Biomechanical Engineering*. 125:207-217

Frochte J, Heinrichs W, 2009. A splitting technique of higher order for the Navier-Stokes equations. Journal of Computational and Applied Mathematics. 228:373-390

Frochte J, 2005. Ein Splitting-Algorithmus höherer Ordnung für die Navier-Stokes-Gleichung auf der Basis der Finite-Element-Methode, *Dissertation*, Universität Duisburg-Essen

Haschke H, Heinrichs W, 2001. Splitting techniques with staggered grids for the Navier-Stokes equations in the 2D case. *Journal of Comp. Phys.* 168, 131

Haschke H, 2001. Splitting-Techniken zur spektralen Approximation der Navier-Stokes und Boussinesq-Gleichungen. *PhD thesis*, Universität-GH Essen

Schöberl J, 1997. NETGEN - An advancing front 2D/3D-mesh generator based on abstract rules. *Computing and Visualization in Science.* 1(1), pages 41-52

Scotti CM and Finol EA, 2007. Compliant biomechanics of abdominal aortic aneurysms: A fluid-structure interaction study. *Computers & Structures*. 85:1097-1113

Scotti CM, Shkolnik AD, Muluk SC and Finol EA. 2005. Fluid-structure interaction in abdominal aortic aneurysms: effects of asymmetry and wall thickness. *BioMedical Engineering OnLine*. 4:64

Sieber G, 2001. Numerical Simulation of Fluid-Structure Interaction Using Loose Coupling Methods. *PhD thesis*, TU Darmstadt

Stüben K, 2001. A review of algebraic multigrid. *J. Comput. Appl. Math.* 128, 281

Thumfart S, 2007. Visualisierung und Volumenmessung von Aneurysmen auf Basis von Angiographiedaten. *Master Thesis*, Institut für Graphische und Parallele Datenverarbeitung, Johannes Kepler University Linz

Valencia AA, Guzman AM, Finol EA and Amon CH, 2006. Blood Flow Dynamics in Saccular Aneurysm Models of the Basilar Artery. *Journal of Biomechanical Engineering*. 128:516-526

AUTHORS BIOGRAPHY

Wolfgang Fenz studied Technical Physics at the J. Kepler University of Linz and finished his PhD in 2004. For the following four years he was research assistant at the Institute for Theoretical Physics there, working on different projects dealing with molecular simulation of fluids, liquid state theory and the theory of phase transitions. At the beginning of 2009 he joined the Research Unit Medical-Informatics at the RISC Software GmbH and is now developing the software for blood flow simulation in the project MEDVIS 3D.

Johannes Dirnberger studied Software Engineering at the Polytechnic University of Applied Sciences Hagenberg. After graduation in 2001, he started his scientific career in the research department of the Polytechnic University in Hagenberg. In 2003, he started a research project for assessment of human burn injuries at the Upper Austrian Research GmbH. Since 2008, he works as senior researcher at the RISC Software GmbH and is responsible for project administration and computer visualization in the Research Unit Medical-Informatics.

PUBLIC FUNDING OF MODELLING AND SIMULATION IN AUSTRIA

P. Kerschl[(a)], A. Pogány[(b)]

[(a)]Austrian Research Promotion Agency (FFG), Department Thematic Programmes, Sensengasse 1, 1090 Vienna, Austria, www.ffg.at/tp
[(b)]Federal Ministry of Transport, Innovation and Technology (bmvit), Department III/I 5 – Department for IT, nanotechnology, industrial technologies and space flight, Renngasse 5, 1010 Vienna, Austria, www.bmvit.gv.at

[(a)]peter.kerschl@ffg.at, [(b)]alexander.pogany@bmvit.gv.at

ABSTRACT
The first two calls of the funding initiative ModSim Computational Mathematics showed that there is a high variety of different areas already addressed by Austrian research institutions and companies. It has turned out that the initiative is important to reduce the gap between high level academic modelling and simulation and the commercially relevant use of the tools. The number of high quality proposals is not sufficient to use the full amount of available funding money at once, and it shows us to make smaller calls with a call volume.of about € 2 Mio. .The public intervention for modelling and simulation is justified by the vision to change the behaviour of the company partners in the long run. The aim is to implement modelling and simulation as a standard tool in the every-day business of Austrian companies. Beside the technical innovations, ModSim wants the funded institutions to establish long-term structures for a broader usage of modelling and simulation.

Keywords: computational mathematics, funding, Austria, economy

1. INTRODUCTION

ModSim is an initiative of the Austrian Federal Ministry of Transport, Innovation and Technology (bmvit) in co-operation with the Austrian Research Promotion Agency (FFG). In Austria, the level of academic research in mathematics is rather high compared to other Sciences (FWF 2007). As the Austrian federal ministry for transport, innovation and technology is responsible for promoting the use of innovative technologies by Austrian companies on the long term, it decided to stimulate the knowledge-transfer from mathematics to the companies by exploiting their expertise in computational mathematics. For this purpose a funding initiative was initiated: "ModSim Computational Mathematics - Developing New Applications of Modelling and Simulation for Austrian Business and Research". In ModSim, the term "computational mathematics" is seen as the combination of mathematics, computer sciences and the specific field of application. Following that scheme, it is obvious, that the funded projects should be run by research institutes and industrial partners in close cooperation.

2. SETUP OF THE FUNDING INITIATIVE MODSIM

The funding initiative was planned and prepared by the bmvit. The FFG conducts the overall programme management of ModSim. The funding is based on the directive for research, technology and development ("FTE-Richtlinie", BMVIT, BMWA 2007) in accordance with EU laws. The initiative was introduced to the research community both at research institutions and at companies in the year 2008. In the first two calls three kinds of funding schemes could be chosen by the applicants: 'Stimulation', 'Cooperative RTD' and 'Development of Human Resources'

2.1. Situation in Austria

In Austria the technological field of Modelling and Simulation has a strong presence and strengths, where scientists can gain challenging funding instruments, which function as a screening and amplification tool. The multidisciplinary cooperation of mathematics and medicine, metallurgy , public economics, etc. in the academic sector functions well. Researchers have a strong ability of organisation and leadership. In Industry Modelling and Simulation in Austria is not well developed, as the users needs a high absorption capacity because of a complex process of harmonization and intervention between the way of posing and solving a problem. Concerning funding in Austria Modelling and Simulation needs an expanded framework: scientific quality as a condition, efficiency in structure as a goal.

2.2. General aspects of funding and aims of ModSim

The two aims of the funding initiative ModSim are:

- Intensify the challenging use of computational mathematics in the Austrian business and research
- Setup and development of structures for research and development with the purpose of long-term transfer of knowledge between

science and economy in the area of Computational Mathematics

The proposing institutions have to be located in Austria. Minor contributions of foreign partners are also possible. The proposed applications should cover total costs of 500.000,- € each or more. Depending on the content also smaller projects are worth being funded. The three different funding schemes support the initiative to reach its aims:

- Stimulation projects aim to stimulate the high-level application of Computational Mathematics according to the aims of ModSim. In case that such projects are research projects, they lead up to a functional research-prototype (whereas product development cannot be funded).
- Projects of the type 'Cooperative RTD' are co-operative research and developing projects with structural development at the project partners. These projects lead up to a functional research-prototype (whereas product development cannot be funded).
- In projects of the type 'Development of Human Resources' existing organisational structures for research and the use of Computational Mathematics should be improved or built up.

2.2.1. Evaluation criteria and process

Like in all other FFG-operated funding programmes, four main criteria are relevant for the decision of funding:

1. Quality of proposed activity
2. Relevance of proposed activity to the ModSim-aims
3. Suitability of applicants / partners
4. Economic potential and exploitation

A jury evaluates whether the proposals fulfil these criteria. Each of the criteria has to be fulfilled independently. The jury consists of 4-9 international experts in the field of computational mathematics and/or in one of the application fields. In the ideal case each jury member has knowledge and experience in organisational affairs in the research area.
The targeted composition of the jury is:

- One third of experts from industry,
- One third of experts from academia
- and one third with a broad knowledge and experience in structural / organisational issues (e.g. Deans of Faculty, Heads of Research Centres).

Within the framework conditions set by the bmvit, the jury is free in its decisions and independent in drawing up its recommendation, including any obligatory conditions on the proposers. The jury decides on the basis of the evaluation manual, in which the evaluation process and selection criteria as well as assessment mode are laid down. The basis for the evaluation is the submitted applications.

The funding recommendation formulated by the jury is transferred to the bmvit. The bmvit decides about the funding.

3. RESULTS FROM THE FIRST CALLS

3.1. Applications

For the first two calls the amount of € 7 Mio. was available. In the first call, 13 proposals from several fields of applications where submitted. In the second call 18 proposals were submitted. Table 1 shows the number of applications with respect to the type of funding scheme chosen by the applicants.

Table 1: Overview of the Applications in ModSim

	Coop. RTD	Stim.	Dvlp. of HR	sum
1st Call	8	4	1	13
2nd Call	8	8	2	18
sum	16	12	3	31

3.2. Funded Projects

From the submitted proposals 13 where chosen for funding by the ministry. The funded projects use the sum of 4.4 Mio €. They deal with several application areas of computational mathematics (alphabetic order):

- Automotive
- Building industry
- Economy
- Logistic, Logistic for public health
- Manufacturing technology
- Material science
- Mechatronics
- Medicine
- Meteorology
- Timberwork

The funded projects have total costs between 127000,- € up to 1 Mio. €.

3.3. List of funded projects from the first call:

In the first call, six proposals where chosen to be funded and they already started their work. The titles of the funded projects are:

- M-CFD - Meteodisciplinary Computational Fluid Dynamics
- MIMOSA - Multi Non-Linear Structural Condition Modelling and Assessment

- RoWaFlowSim - Simulations of liquid film flows with free surface on rotating silicon wafers
- MEDVIS 3D - Blood Flow Simulation in Intracranial Aneurysms
- ATHDM-E3 - Austrian Hybrid Dynamic Model E3
- AtoMat - Atomistic Approaches towards Materials Design

The results from the second call were not officially release and therefore not publishable at the time of submission of this paper. You will find the results on the homepage: www.ffg.at/modsim

4. CONCLUSION AND OUTLOOK FOR MODSIM

4.1. Conclusions from the first two calls

Considering the results from the first two calls, the following conclusions can be identified:

- Modelling and Simulation is used in very different application areas in Austria.
- The number of proposals of sufficiently high quality is not high enough to make use of the total money available for funding at once, and it shows us to make smaller calls with a call volume of € 2 Mio.
- The increasing number of applications in the second call shows the ongoing demand for public funding for modelling and simulation.
- The degree of co-operation between research institutions and companies is very high.
- Stimulation projects were more popular in the second call than in the first one.
- Stimulation projects have the highest funding rate considering the number of applications and funded projects.
- Projects for 'Cooperative RTD' are the biggest ones.
- The aimed project size could be reached; nevertheless a big variety of the size of the applications is present which makes it difficult for the jury to compare the projects with each other.

4.2. Third call of ModSim

It is planned to offer a third call for proposal in the ModSim initiative. The results from the first calls will be used to further improve the efficient use of the funds for improvement of the Austrian researchers in research institutions and companies. Information about this will be available on www.ffg.at/modsim

4.3. European funding for Modelling and Simulation

On european level Modelling and Simulations gets funded by top-down calls within the priorities "ICT" (Information and communication Technologies") and NMP ("Nanotechnologies, Material and Production technologies") of the 7th Framework Programme. Austrian participation in ICT is quite successful. With the begin of the 7th Frame work programme 3,5 % of the overall successful participants wihin ICT came from Austrian research institutions, universities and companies. Due to the economic crisis the European Commission has launched a PPP ("Public, Private, Partnership") Initiative for Manufacturing (Factories of the Future"-FoF) as joint call between ICT and NMP, where Modelling and Simulation plays an important part. Austria will build up a national Plattform with the goal to evaluate research topics with are important for Austrian stakeholders, to bring them into the workprogramm of the FoF-initiative, to build up critical mass and in a long term to further more increase participation. of Austrian stakeholders.

ACKNOWLEDGMENTS

The work described in this paper is financed by the bmvit.

REFERENCES

BMVIT, BMWA 2007 : RICHTLINIEN zur Förderung der wirtschaftlich-technischen Forschung und Technologieentwicklung (FTE-Richtlinien) gemäß § 11 Z 1 bis 5 des Forschungs- und Technologieförderungsgesetzes (FTFG) des Bundesministers für Verkehr, Innovation und Technologie vom 19. 11. 2007 (GZ BMVIT-609.986/0011-III/I2/2007) und des Bundesministers für Wirtschaft und Arbeit vom 30. 11. 2007 (GZ BMWA-97.005/0002-C1/9/2007) (in german)

DI Fritz Ohler, DI Brigitte Tiefenthaler, Technopolis Forschungs- und Beratungs- GmbH, 2007, „Modellierung und Simulation – Analyse und Förderung eines emergierenden Forschungsgebietes"

FWF 2007: FWF - Der Wettbewerb der Nationen – oder wie weit die österreichische Forschung von der Weltspitze entfernt ist. Eine Analyse der internationalen Wettbewerbsfähigkeit wissenschaftlicher Forschung Österreichs in den Natur- und Sozialwissenschaften, Wien, 2007 (in german)

AUTHORS BIOGRAPHY

Peter Kerschl is working at the Thematic Programmes of the Austrian Research Promotion Agency (FFG) since 2006. Before he worked at the Institute for Solid State Physics and Materials Science Dresden (IFW Dresden), where he finished his PhD thesis in the area of experimental solid state physics.

Alexander Pogány is working as policy expert for the Federal Ministry for Transport, Innovation and Technology within the area of nano- and industrial technologies. Before he worked as validation expert in quality control for Baxter Bioscience. He holds a Master in Microbiolgy.

MULTI NON-LINEAR STRUCTURAL CONDITION MODELLING AND ASSESSMENT

Wenzel Helmut[(a)], Tanaka Hiroshi[(b)], Schäfer David[(c)], Höllrigl-Binder Michaela[(d)], Allmer Helga[(e)]

[(a) (c) (d) (e)]VCE – Vienna Consulting Engineers
[(b)]University of Ottawa

[(a)]wenzel@vce.at, [(b)]htanaka@uottawa.ca, [(c)]schaefer@vce.at, [(c)]hoellrigl-binder@vce.at, [(e)]allmer@vce.at

ABSTRACT
System identification methodologies have reached a mature state. Starting from mechanical engineering it quickly conquered civil engineering as well as aerospace applications. Damage detection has been the key topic of recent research work in this field. Learning from evolution of dynamic characteristic throughout the damage process, one can validate damage detection methods or set benchmark studies for typical aging processes of a structure. When dynamics gain a guiding role in the behaviour of a structure our conventional approaches fail. The influence of the various non-linearities can not be neglected anymore. This paper is determined to describe a number of these non-linearities and their embedment into the system identification procedure. The results show that changes of natural frequencies are clearly visible, but not a reliable identification approach. Identifying the non-linearities helps to identify damage at an early stage and enables even damage quantification, location and most important remaining lifetime prediction.

Keywords: Non-linear dynamic behaviour, system identification, damage detection, health monitoring

1. INTRODUCTION

Structural Health Monitoring (SHM) is a subject that has received considerable attention in recent years. Incidents such as the collapse of the I-35 Bridge in the U.S. give a clear indication of the importance of SHM. In practice, bridge assessment includes several measures, such as inspection, data interpretation, risk assessment and development of engineering recommendation.

For global structural assessment in particular, vibration monitoring (Wenzel 2005) has been widely used. The authors have been involved at the frontline of development of SHM in civil engineering and particular SHM of bridges (Wenzel 2009). Vibration characteristics captured from vibration monitoring provide global information on structural behaviours such as stiffness, connectivity, boundary conditions, mass distribution and energy dissipation. The basic principle of vibration based structural assessment is that structural performance changes from defects will create changes in the dynamic response that can be detected from the changes in the vibration characteristics. In other words changes in energy distribution, frequencies, mode shapes, vibration intensities and system damping can be used as indicators of the changes of physical properties of structures such as mass distribution, vibration energy contribution, stiffness, connectivity, boundary conditions and energy dissipation. Extensive works have been done on developing methods and algorithms for damage detection using vibration characteristics. Döbling et al (1996) present a comprehensive list of literature in damage detection, divided the detection algorithms into 4 levels of increasing complexity. They are:

- level 1 determination that damage is present in a structure,
- level 2 determination of the geometric location of the damage,
- level 3 quantification of the verity of damage, and
- level 4 prediction of the remaining service life of a structure.

2. MULTI-NON-LINEAR-SYSTEM-IDENTIFICATION

Many algorithms have been developed during the long period this methodology has been identified as being a promising tool. Most of it has been of theoretical or computational nature with finite element simulations as the key element. In parallel laboratory scale experiments were performed for verification. Large field tests became possible after the monitoring equipment reached a mature stage, PCs were introduced and the handling and storage of large quantities of data became feasible.

It quickly showed that the laboratory experiments can not be simply transferred into nature. The simple reason is that the influence from environmental and other sources is big enough to cover the desired information in unintended noise. It further has been found that these influences show a distinct non-linear character. In addition the combination of these non-linearities can almost show random patterns producing random looking results. This has led to a first frustration

phase in the 1980th. With increasing knowledge on these phenomena a new wave of enthusiasm has reached the research community. The underlying MIMOSA project is determined to considerably contribute to this evolution.

The principle of the methodology shall be explained using the equation of motion. In a complex form it describes the dynamic behaviour of a structure.

$$[m]\{\ddot{z}\}+[c]\{\dot{z}\}+[k]\{z\}=F \tag{1}$$

where m represents the mass matrix, c the damping matrix and k the stiffness matrix of the system.

The non-linearities observed in nature can be linked to some of its particular elements.

2.1. Change of Temperature

The first and most basic effect on a structure is the change of temperature. Temperature influences the stiffness of a structure by altering its material properties particular the Young's modulus but also on smaller scale its geometry. Linked to that, a change in boundary conditions might happen. The simplest form would be the linear change of the stiffness matrix over time. The stiffness matrix k in equation (1) has the form

$$k=\begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1n} \\ k_{21} & k_{22} & \cdots & k_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ k_{n1} & k_{n2} & \cdots & k_{nn} \end{bmatrix}. \tag{2}$$

In large structures we observe the fact that non-linearity not only develops over time but also over location. This means that the elements of the stiffness matrix might develop non-linear relationships depending on temperature changes due to radiation from sun on particular local scale. This might add a torsional effect, changing the structures geometry and stress level. Despite the complexity of this phenomenon modern finite element simulations are able to deal with it reasonably.

2.2. Added Mass

The main purpose of bridges is to enable transportation. This brings considerable added masses to the structure. These masses are furthermore moving at various speeds producing a non-linear interface relation due to the vehicle suspensions. At small structures, where the own weight of bridge compared to the added mass is low this effect might dominate the entire behaviour.

At large structures the volume of air captured in the box of a bridge can not be neglected as lesson learned from bridge aerodynamics. These non-linearities are residing in the 1st term of our equation. The vehicle interaction can be alternatively expressed in the right side of the equation.

$$m\ddot{z}+c\dot{z}+kz=F(z,\dot{z}) \tag{3}$$

It is needless to say that a good solution to this problem can only be found if online information on traffic, which has to be classified, becomes available. This is a considerable challenge for the monitoring community.

A typical example of a lively bridge is given below. One graph shows the unprepared result representing all influences from traffic and the other is the cleaned result after simulation.

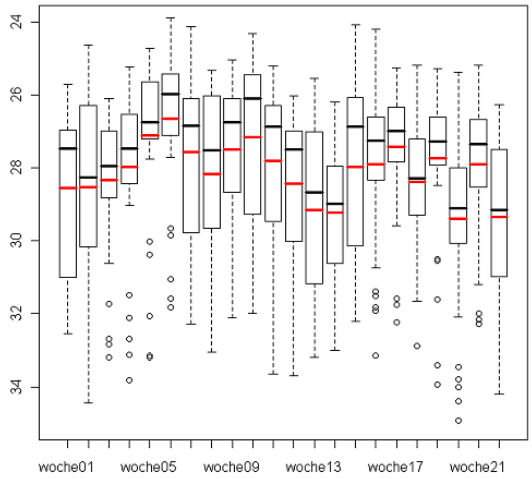


Figure 1: Bridge Performance unprepared

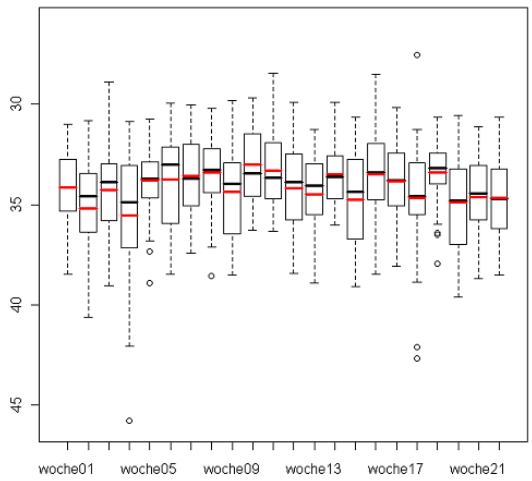


Figure 2: Bridge Performance cleaned

2.3. Material and Structural Aging

Knowledge has been collected by VCE over the thousands of monitoring campaigns performed in the past 12 years. The character of a spectrum changes not only when damage is present, but also due to aging effects. It is no surprise that this is more expressed in concrete structures than in those using any other material. High focus has been put on post-tensioned concrete structures in the past due to their complexity and importance within our transportation networks. This

produced huge amounts of data and resulting knowledge.

The performance of concrete over time (aging) is well defined. Nevertheless the exposure creates considerable uncertainties. The bridges in the alpine regions are particularly exposed to an attack of salt used for de-icing. Carbonation is also highly non-linear and exposure dependent. The dependency on exposure can be handled in the stiffness matrix in case that a respective constitutive law can be formulated properly. How drastic exposure contributes to non-linear behaviour of structures has been demonstrated through the various permanent monitoring campaigns performed. At the Gossensass Bridge on Brennero Autostrade in Italy 2 equal structures are placed side by side. The performance of the one exposed to sunshine from the west is entirely different from the one residing in the shade. A considerable difference on quality has been found in these 40 years old structures. Exposed to the same air temperature the difference between the carriage ways south and north is considerable as shown in the next figure.

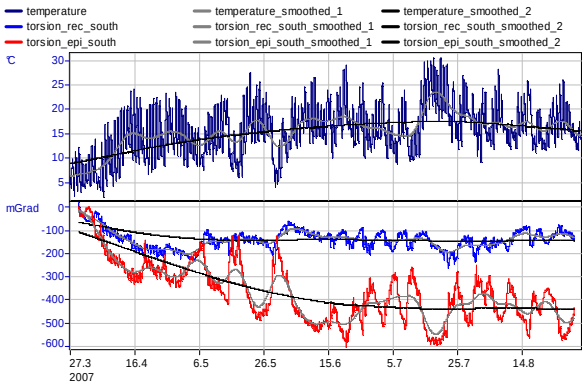


Figure 3: Different response of equal structures due to exposure

2.4. Non-Linear Damping

This most complex phenomenon has turned out to be a promising approach towards structural health monitoring. This phenomenon seems to be much more sensitive to damages than any expected reduction of stiffness. From large scaled damaging tests it has been found that there is a relationship between the amplitude of the fundamental frequencies of a structure and the damage state. Reduced amplitudes normally show a broader peak which implies increased damping (referring to a simple half power band widths approach).

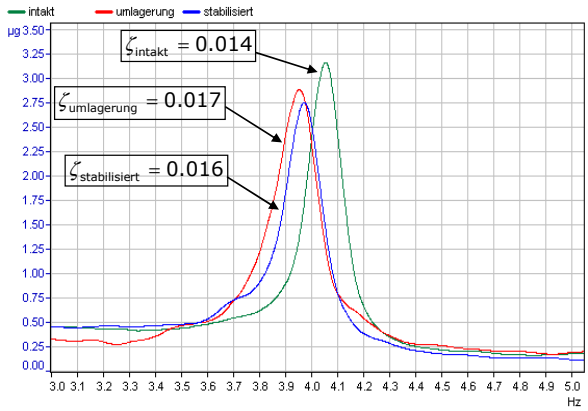


Figure 4: Change in amplitude with damage and related damping value

Targeted laboratory tests (SAFE PIPES Project FP6 STRP-013898) showed a phenomenon which we called energy cascading with the increasing non-linearity of the dynamic system.

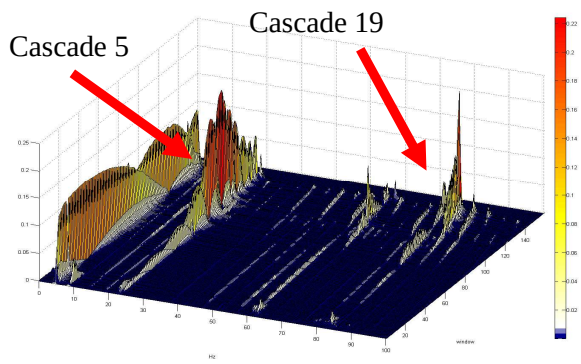


Figure 5: Energy cascading with damage (SAFEPIPES test at MPA Stuttgart)

Cascading energy transfer in a dynamical system could be caused by the development of nonlinear characteristics of the structural response caused by various reasons.

Suppose there is a development of nonlinear mechanism in both damping and stiffness of the structure, whose dynamic behaviour is simply expressed by a SDOF model, associated with the progress of a structural damage. It can be typically represented in the equation of motion by modifying both damping and stiffness terms as follows:

$$m\ddot{z} + c[1+\varepsilon_2(z)]\dot{z} + k[1-\varepsilon_1(z)]z = 0 \qquad (4)$$

where ε_1 and ε_2 are the nonlinear terms introduced corresponding to the development of structural damage and $z(t)$ represents the dynamic response of the structure in general. Eq.(1) can be rewritten as

$$m\ddot{z} + c\dot{z} + kz = k\varepsilon_1(z)z - c\varepsilon_2(z)\dot{z} = F(z,\dot{z}) \qquad (5)$$

where $F(z,\dot{z})$ is generally a nonlinear function of z and/or $\dot{z}$, such as

$$F(z,\dot{z}) = C_1 z^2 + C_2 \dot{z}^3 \tag{6}$$

for example. It implies that if z is given as a vibration with frequency ω, $F(z,\dot{z})$ is generally a function of fluctuations with the frequencies expressed by the multiples of ω. For example, substitution of $z = A\sin\omega t$ to (3) results in

$$F(z,\dot{z}) = \frac{C_1 A^2}{2}(1-\cos 2\omega t) + \\ + \frac{C_2 A^3 \omega^3}{4}(3\cos\omega t + \cos 3\omega t) \tag{7}$$

which in turn will result in the dynamic response including functions of twice, thrice higher frequencies of the original for this case. The same process will be repeated as time allows and, as a result, a part of the system's dynamic energy will be gradually distributed to higher and higher frequency range.

Where would this process end? For the case of damage-caused nonlinearity, the high frequency energy components dissipate as heat or noise and, if not, the destruction or rupture of the structure plays a roll. Even if it does not reach the destruction point, the mechanism of structural response will change largely when damage progressed.

An example from the damage test at the S101 Bridge in Austria is provided below. It shows the results of above described algorithms for the 2 damages introduced during this demonstration within the IRIS Project (FP7 CP-IP 213968-2).

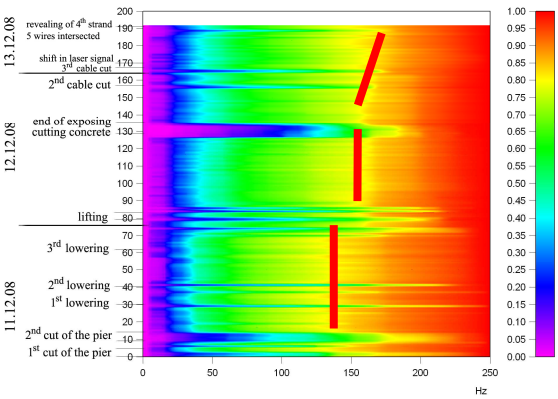


Figure 6: Simulation S101 with damage (IRIS test at A1, Austria)

3. THE MIMOSA PROJECT

MIMOSA is determined to develop necessary models to allow online updating of the multi non linear behaviour of our infrastructure. The product should be a Modeling and Simulation Tool Box able to provide the user with all the necessary information.

The current practise shall be considerably enhanced through expertise in multi non linear modelling solutions with the necessary mathematical formulations in the background.

The Finite Element Model Updating (FEMU) is a model based numerical technique used to minimize the differences between a real structure and a FE-model. The basic idea is to use the recorded structural response to update some selected structural parameters of the numerical model (such as stiffness, mass and internal forces) as well as some boundary conditions (such as translational or rotational springs), until an adequate agreement between numerical and experimental results are achieved. The resulting structure presents a better dynamic agreement with the physical reality. Moreover, the parameter distributions obtained as outcomes can provide useful information about the possible structural damage.

The physical explanations are partly offered in the literature, energy cascading is one of these solutions and shall be explained further.

3.1. Structural Nonlinearity and Energy Transfer

Cascading energy transfer in a dynamical system could be caused by the development of nonlinear characteristics of the structural response caused by various reasons.

A development of nonlinear mechanism in both damping and stiffness of the structure is often associated with the progress of a structural damage. In the description of structural vibration, it can be typically represented by additional nonlinear correction terms introduced to damping and stiffness terms. It implies that if a simple harmonic oscillation of the structure takes place with certain frequency, whatever the reason is, the vibration will soon become mixed with higher frequency components generally expressed by the multiples of the original frequency. This process will be repeated as time allows and, as a result, a part of the system's dynamic energy will be gradually distributed to higher and higher frequency range.

3.2. Nonlinear Damping

Energy cascading can be associated with various types of nonlinear physical phenomena. Another example of it is typically observed in dynamics of turbulent fluid flow. In fact, the case of fully developed turbulence, the energy cascading process is one of the most central issues. Relationship of energy cascading with nonlinearity of dynamical systems is therefore evident in these two different phenomena.

A very interesting aspect of this point is that the detection of energy cascading could be potentially utilized as a tool for the structural health monitoring.

The traditional idea of knowledge-based structural health monitoring is by identifying the reduction of stiffness, which has been proved to be far less sensitive than desired for practical purposes. In contrast to that, by finding the transfer of dynamic energy to higher frequencies through spectral analysis of the ambient vibration survey, it may be possible to detect the damage development in a structure at its earlier stage. Any extent of structural damage can of course change the local structural damping or energy dissipation and stiffness. As a consequence, the global dynamic properties of the structure, i.e., the eigen-frequencies, mode shapes and modal damping would be all somewhat influenced.

It needs to be kept in mind that structural nonlinearity is attributed, however, not only to developing damages. Field experience indicates that the magnitude of modal damping is often amplitude dependent. Increase of damping, when the vibration amplitude is significant, is due to energy consumption at increased friction at bearings, bending action of piers, behaviour of the bridge outfitting and also the structure-vehicle interaction.

Admittedly the present method would also detect the developing structural nonlinearity due to large motion. However, if there is a development of structural damage as its consequence, the nonlinear characteristics will remain with the structure after the large amplitude motion disappeared and should be thus detected.

3.3. Progress in the Proposed Project

The method described above has been experimentally applied to some bridges to investigate how effectively it can be employed for identifying the existence of structural damages. Sample structures chosen for this experimental project include ordinary healthy bridges, bridges with progressive deterioration, and also some bridges that were to be demolished. For the case of the last category, the structure was given artificial damages before their demolition and its influence on their dynamic behaviour was examined. The identification of the energy cascading phenomena was performed by time-limited spectral analysis of acceleration data obtained by a standard BRIMOS ambient vibration survey, which has been established by VCE. An ideal condition for this measurement would be when the structure is excited by micro tremors, which can be regarded as white noise excitation. It would be even better if the measurement was continued for a while under the same conditions. However, the reality is often not under such conditions.

Spectral analysis of the vibration record has clearly indicated that dynamic energy tends to be gradually transferred to higher frequency range when structural damages exist. The most fundamental principle of this measurement has been thus confirmed. This was evidenced particularly clearly when the structure was artificially damaged and measurement was carried out under rather ideal conditions, namely without being disturbed by the on-going traffic loads directly on the structure.

When the vibration was measured with the structures under service conditions, on the other hand, the measured data were much contaminated, as expected, by direct excitation due to traffic loads. The energy transfer due to possible structural nonlinearity, hence, needs to be carefully concluded by somehow subtracting the effects of traffic load excitation. A few different approaches have been tried out but the work is still continuing at this point in time.

4. OUTLOOK

The described non-linearities do not comprise the entire subject in its full complexity. Further non-linearities are to be considered until a closed solution can be offered. These are among others:

- Non-linear changes of the boundary conditions
- Dynamic effects from wind loads
- Earth tide effects on large structures particularly in steep valleys
- Strain from external phenomena, i.e. creeping slopes
- Other influences not yet identified

Out of the simple finite element model a most complex, multi-dimensional simulation is to be developed in order to describe these phenomena properly. A major problem is the combination of the single non-linearities which might lead to extinction resulting in difficulties for identification. Monitoring results provide us the complex system without good ways for isolation of the single phenomena. It will therefore be continuously necessary to search for well expressed non-linearities, trying to establish them in a well controlled laboratory test and subsequently formulate the laws for modelling and simulation. We still see a long way to go until numerical simulation alone can deal with these circumstances.

In the recent years a database containing 1200 structures with more than 4 million measurement files has been erected. It will be tried to apply case based reasoning routines to identify single phenomena out of this extension knowledge base. A breakthrough in identification could be expected.

A further subject of interest will be the online application, where well defined structures equipped with a suitable monitoring system run a numerical simulation online permanently. The final target will be an online risk assessment tool enabling the operators of the structures to see an actual status anytime.

ACKNOWLEDGMENTS

The MIMOSA project is an Austrian National Research Initiative called MOD-SIM initiated by the Austrian Ministry of Infrastructures (BMVIT) and handled by the Austrian Research Promotion Group FFG. It provides the basis for the theoretical works carried out.

The extensive monitoring and testing campaigns in the field and the large laboratories are performed under the European Research Project SAFE PIPES (FP6 STRP-013898) in the 6th framework program and IRIS (CP-IP 213968-2) in the 7th framework program of the European Commission. Without this valuable support from the European Commission and the National Science Foundation this complex work would not have been possible.

REFERENCES

Döbling, S.W., Farrar, C.R., Prime, M.B., 1996. *Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review*. Report Num. LA-13070-MS, Los Alamos National Laboratory.

Wenzel, H., 2009. *Health monitoring of bridges*. Chichester, England: J. Wiley and Sons LTD.

Wenzel, H., Pichler, D., 2005. *Ambient vibration monitoring*. Chichester, England: J. Wiley and Sons LTD.

AUTHORS BIOGRAPHY

Wenzel Helmut: Dr. Wenzel is well known among Europe's leading experts in the field of Ambient Vibration Monitoring. Since 1995 he has pushed the development of the Monitoring system BRIMOS to a level relevant for practical application. This is reflected in various European initiatives aimed to support this development. In this function he serves as a project coordinator and advisor to the Commission in respect of future research and development work to be funded. From the inception of his career, Dr. Wenzel has actively lead theoretical studies into new design methods based on highly advanced computer programs, such as the Finite Element Method (FEM), for oblique angled bridge slabs. Following an early focus on research and development, his attention turned to practical structural design applications. He has written numerous articles on innovation and research in bridge design which have been published in all the major engineering journals. Since 1992 he has taught and lectured at the University of Civil Engineering in Vienna in the field of bridge design and construction technology. Under Dr. Wenzel's leadership, VCE has designed, supervised and managed more than 250 contracts in over 56 countries. Most notable have been the innovations in bridge design and supervision, sophisticated tunneling projects implementing the New Austrian Tunneling Method and large scale industrial projects. Currently 14 R&D projects are under progress.

Tanaka Hiroshi: After seven years of experience at the Boundary Layer Wind Tunnel Laboratory, University of Western Ontario, working on research projects in Wind Engineering and Structural Dynamics, Dr. Tanaka joined the Faculty of Engineering at the University of Ottawa in 1979, where he has been carrying out teaching, research and administrative duties including Chair of Civil Engineering (1988-91, 2001-04) and Director of Ottawa-Carleton Institute for Civil Engineering (1987-88, 93-95, 1998-00). He is a well-recognized expert in the field of Wind Engineering, particularly in Bridge Aerodynamics, and has published over 200 scientific papers, engineering reports and articles. As engineering consultant, he has been involved in wind resistant design and aerodynamic consideration of many significant bridges world-wide, such as the Annacis Island Bridge (Canada), Bronx-Whitestone Bridge (USA) and Storebælt Bridge (Denmark). Meanwhile, he also worked as a guest researcher at National Research Council Canada, Danish Maritime Institute, Vienna Consulting Engineers etc.

Schäfer David: After graduating at the grammar school at AHS Bundesrealgymasium Gmünd, David Schäfer started studying Technical Physics at the Technical University of Vienna. Since he graduated in Technical Physics in 2009 he works as a research engineer at VCE (Vienna Consulting Engineers). He is currently working on bridge monitoring projects as well as geotechnical projects.

Höllrigl-Binder Michaela: After finishing secondary school in Lower Austria, Michaela Höllrigl-Binder started studying Technical Mathematics at Vienna University of Technology in 2002. She concentrated on the field of Mathematics in Science and graduated in 2009 with the diploma thesis "Mathematical models describing physiological processes". Michaela Höllrigl-Binder is currently working as a research engineer at VCE, where she is involved in research activities concerning Bridge Health Monitoring, for example life circle analysis.

Allmer Helga: Having studied structural civil engineering, Helga Allmer worked from 2000-2002 as a research assistant at the Institute of Mechanics at the Vienna Technical University. After that from 2003-2007 she was employed as a technical assistant at the Institute of Leightweight Design and Structural Biomechanics, also Vienna Technical University. Since then she is working at VCE as a research and project engineer on cable dynamics and bridge monitoring projects.

THE ROLE OF THE GREEN QUOTA AND REVENUE RECYCLING SCHEMES IN THE CLIMATE CHANGE OPTIONS: A DYNAMIC GENERAL EQUILIBRIUM ANALYSIS FOR AUSTRIA

Stefan Schmelzer [(a)], Todor Balabanov[(b)]

[(a)]Institut fuer Hoehere Studien, Vienna
[(b)] Institut fuer Hoehere Studien, Vienna

[(a)] schmelzer@ihs.ac.at , [(b)]balabano@ihs.ac.at

ABSTRACT
By simulations with the dynamic equilibrium model TD-BU-E3 DGEM the long term impacts of two alternative policy instruments for responses to climate change were assessed: green quota and double dividend. Electricity demand growth was de-coupled from the economic growth. 3 economic sectors, 5 existing and three new vintage electricity production technologies were considered.

By 2050 the share of renewables in the electricity production could be reaching 0,289 and there are sufficient potential renewable resources. The economic burden is bearable and the welfare is growing.

Checking the double dividend hypothesis (trade-off b/n environmental benefits and gross economic costs): the reduction in the labor tax is increasing consumption; the reduction of consumption tax to a lesser extent so but the reduction in the lump-sum refund to the representative household is detrimental to consumption.

Hence, only for the case of labor tax recycling, we could assume the existence of a strong double dividend.

Keywords: climate change, CO_2 taxation, abatement strategies, general equilibrium models

1. INTRODUCTION

The aim of the paper is to quantitatively assess the macroeconomic and sectoral impacts of future responses to climate change by evaluating policies for adaptation and mitigation aiming at promoting increased market penetration of electricity produced from renewable energy sources in Austria.

The term **adaptation** is to be related to de-coupling of electricity demand from the economic growth by energy and resources conservation in the sense of sustainable development, by changing consumption pattern and habits, etc. – all that are long term measures related to socio economic changes.

For the long-term **mitigation** options for the electric power sector will focus on CO_2 reduction by the mean of a set of the technological options where strong potentials for CO_2 reduction exist.

To grasp synergies in climate policy the adaptation and mitigation options must be analyzed within a consistent, dynamic framework allowing for carrying out of integrated analyses of alternative scenarios for adaptation and mitigation strategies.

Mitigation and adaptation policies should be assessed on their full effects and their quantification calls for the use of the newly developed Top/Down -BU for Bottom-up E3 (energy, environment, economy) dynamic general equilibrium model (TD-BU-E3 DGEM) allowing for systematic trade-off analysis of environmental quality, economic performance and welfare (consumption).

As to policy measures related to mitigation by promotion of renewable energies there had been a shift - as more generally in environmental policy design - from command-and-control policies to market-based instruments such as taxes, subsidies, and tradable quotas. A recent impact assessment by the European Commission, 2008, shows that feed-in tariffs in Austria are the preferred promotion measure. In addition, direct subsidies for renewable energy have been enacted – typically differentiated by the type of green energy, i.e., wind, biomass, solar cells, etc.

A relatively new strand of policy regulation is the use of tradable green quotas where energy supplies are required to produce a certain share of energy services from renewable energy but are flexible to trade these shares between each other in order to exploit potential difference in specific compliance costs.

In this paper, focus on two alternative policy instruments which may be quite relevant to the Austrian strategy for promotion of renewable energy sources: quota obligation systems and Carbon Taxation (double dividend) instruments.

Methodological the focus is set on novel CGE (Computational General Equilibrium) modeling approaches. The methodological objective is to consistently describe the role of specific energy related technologies within a total analytical economic modeling framework. CGE is used as an analytical Top-Down framework that is enhanced by representation of specific technology descriptions.

The paper is structured as follows: Section 2 provides a background to the TD-BU-E3 DGEM and its algebraic representation in the MCP framework, followed by its adjustment to the study's specifics and application to the particular case studies in Section 3 that is dealing with Scenario definition and policy analysis starting with benchmark assumptions, then the description and analysis of the Baseline Scenario followed by the Green quota scenario and respective analysis and ending with the Carbon Taxation (double dividend) Scenario. Section 4 concludes.

2. THE TD-BU-E3 DGEM

Our modeling work was motivated by recent theoretical and practical developments in algorithms for nonlinear complementarity problems and variational inequalities based on the GAMS/MCP modeling format (Rutherford, 2002).

The TD-BU-E3 DGEM where TD stands for Top/Down, BU for Bottom-up, E3 for energy, environment, economy and DGEM for dynamic general equilibrium model.

The TD-BU-E3 DGEM provides a basis for evaluating economic impacts of the chosen energy policies both at macroeconomic and at the sectoral level – indicating the effects of the energy decisions on the economic environment. This approach permits an energy-economy model to combine technological details of an energy system (bottom-up) with a characterization of the market equilibrium (top-down).

TD-BU-E3 DGEM applications include the impacts of scenarios on country´s economic variables, e.g., changes of the main real economic indicators, in the consumption of the households, in the sectoral employment levels, in the energy consumption, of the emission levels, the energy price indices, etc., but TD-BU-E3 DGEM is also used for applied energy and environmental policy analysis, e.g., the impacts of the Green Quotas and the Environmental Tax Reform

2.1. TD-BU-E3 DGEM: algebraic representation in MCP framework

In our formulation of an integrated top-down / bottom-up model we consider a competitive (Arrow-Debreu) economy with n commodities (including economic goods, energy goods and primary factors) indexed by i, m production activities (sectors) indexed by j, and h households (including government) indexed by k. We making use of the MCP framework suggested by Boehringer (2007) formulation of market equilibrium problems as mixed complementarity problems (MCP) thus permitting integration of bottom-up programming models of the energy system into top-down general equilibrium models of the overall economy. The decision variables of the economy can be classified into the following categories:

p denotes a non-negative *n*-vector in prices for all goods and factors,

y is a non-negative *m*-vector for activity levels of constant-returns-to-scale (CRTS) production sectors,
M is a *h*-vector of consumer income levels,

e represents a non-negative *n*-vector of net energy system outputs (including, for example, electricity, oil, coal, and natural gas supplies), and

x denotes a non-negative *n*-vector of energy system inputs (including labor, capital, and materials inputs).

Given the underlying functional forms, we observe that the complementarity conditions only will apply for the energy sector technologies and the shadow prices on the associated capacity constraints; all of the macroeconomic prices and quantities will be non-zero. By use of Shepard's Lemma we can then write the equilibrium as the following mixed complementarity problem:

- **Zero-profit conditions**:

$$\overline{z}_{it} \geq z_{it} \perp \mu_{it} \geq 0 \quad (1)$$

$$-\Pi^{E}_{it} \geq 0 \perp z_{it} \geq 0 \quad (2)$$

$$\prod_{j}^{N} s_j = 0 \quad (3)$$

- **Market clearance conditions:**

$$s_j = \sum_{it} a^{j}_{it} z_{it} + c \frac{\partial \Pi_c}{\partial p_j}$$

$$j = 1 - N$$

$$\overline{L} = \sum_{j=1}^{N} s_j \frac{\partial \Pi_j}{\partial w}$$

$$\overline{K}_j = s_j \frac{\partial \Pi_j}{\partial r_j}$$

$$\sum_{t} z_{it} - \sum_{i't} b_{ii't} z_{i't} = c \frac{\partial \Pi_c}{\partial p^{E}_{i}} + \sum_{j=1}^{N} s_j \frac{\partial \Pi_j}{\partial p^{E}_{i}}$$

$$c = \frac{M}{p_c}$$

(Equations 4 to 8)

- **Income balance**:

$$M = \sum_{j=1}^{N} r_j \overline{K}j + w\overline{L} + \sum_{it} \mu_{it} z_{it} \quad (9)$$

Extending the above setting into an intertemporal model version only requires a few additions (most of the underlying economic relationships hold on a period-by-period basis), regarding capital stock formation and investment, an efficient allocation of capital, i.e.

investment over time, implies two central intertemporal zero profit conditions which relate the cost of a unit of investment, the return to capital, and the purchase price of a unit of capital stock for each time period τ.

Capital evolves through geometric investment and geometric depreciation

Output markets must also account for investment demand.

The consumer allocates lifetime income, i.e., the intertemporal budget, over time in order to maximize utility, solving:

$$\max \sum_{\tau} \left(\frac{1}{1+\rho}\right)^{\tau} u(C_{\tau}) \tag{10}$$

subject to

$$\sum_{\tau} p_{\tau}^{C} C_{\tau} = M \tag{11}$$

With isoelastic lifetime utility the instantaneous utility function is given as:

$$u(c) = \frac{c^{1-\frac{1}{\eta}}}{1-\frac{1}{\eta}} \tag{12}$$

Summary of equilibrium variables in the TD-BU-E3 DGEM:

a. Activity variables

c	Aggregate consumption
S_j	Production of goods in the sectors j= *1- N*
E_i	Aggregated output of energy good i
$z_{i,t}$	Production by technology t for energy good i
E_i^j	Demand for energy good i in the sectors S_j
E_i^c	Final demand for energy good i
L_j	Labor demand of goods in the sectors S_j

b. Price variables

p_c	Price index of final consumption
P_j	Non energy goods from sectors S_j
p_i^E	Energy prices for i = *{OIL, GAS, COL, ELE}*
w	Wage rate
R_j	Returns to non energy capital for S_j
μ_{it}	Energy sector returns

c. Income variable

M	Income of the representative agent

d. Additional variables and parameters for dynamic extension

pK_{τ}	value (purchase price) of one unit of capital stock in period τ ,
K_{τ}	associated dual variable which indicates the activity level of capital stock formation in period τ
I_{τ}	is the associated dual variable which indicates the activity level of aggregate investment in period τ
$u(.)$	denotes the instantaneous utility function of the representative agent,
ρ	is the time preference rate
M	represents lifetime income (from endowments with capital, time, and resources).
η	constant intertemporal elasticity of substitution

For calibration of the TD-BU-E3 DGEM we use the social accounting matrix for 2005 and also the following data:

Intertemporal elasticity of substitution	0,5
Baseline interest rate	5 %/year
Baseline growth rate	0.9 %/year
Depreciation rate	7 %/year

Price of electricity for S_j = { AGR/FOR, PRD/EIS, SRV, ENE (coal, gas, oil, electricity); }, *j =1* - N written as Cobb-Douglass function (the same form is used for all other prices)

$$p_j^E = \left(\frac{P_{ELE}}{\theta_j^{ELE}}\right)^{\theta_j^{ELE}} \left\{ \delta_j \left(\frac{p_{COL}}{\delta_j}\right)^{(1-\sigma_i^E)} + (1-\delta_j)\left[\sum_{ff}\left(\frac{p_{ff}}{(1-\delta_j)\theta_j^{ff}}\right)^{\theta_j^{ff}}\right]^{(1-\sigma_j^E)} \right\}^{1/(1-\sigma_j^E)} \tag{13}$$

θ_j^{ELE}	Cost share of electricity, or oil or gas coal in the composite output Y
δ_j	Cost share of non electricity energy composite in the output Y
σ_j^E	Elasticity of substitution
p_j^E	Price of electricity for sector j

Unit profit functions for S_j.{AGR-FOR, PRD-EIS, SRV, ENE }, j = 1 - N are in turn given by:

$$\Pi_j = p_j - \frac{1}{\phi_j}\left[\gamma_j\left(\frac{p_j^E}{\gamma_j}\right)^{(1-\sigma_j)} + (1-\gamma_j)\left(\frac{r_j}{\theta_j(1-\gamma_j)}\right)^{\theta_j(1-\sigma_j)}\left(\frac{w}{(1-\theta_j)(1-\gamma_j)}\right)^{(1-\theta_j)(1-\sigma_j)}\right]^{1/(1-\sigma_j)} \tag{14}$$

The unit cost of energy inputs to final demand are given by:

$$p_c^E = \left(\sum_i \beta_i \left(\frac{p_i^E}{\beta_i}\right)^{1-\sigma^{EC}}\right)^{1/(1-\sigma^{EC})} \tag{15}$$

And the resulting cost of a unit of final consumption for j=1-N e.g.{AGR, PRD, EIS, ENE, SRV}is:

$$p^c = \left[\alpha \left(\frac{p_c^E}{\alpha} \right)^{1-\sigma^c} + (1-\alpha) \left(\prod_{j=1}^{4} \left(\frac{p_j}{\theta_j^c (1-\alpha)} \right)^{\theta_j^c} \right)^{1-\sigma^c} \right]^{1/(1-\sigma^c)} \quad (16)$$

Where

$$\sum_{j=1}^{N} \theta_j^c = 1 \quad (17)$$

Finally, the unit profit associated with technology *t* for energy good *i = {col, oil, gas, ele}* is:

$$\Pi_{it}^E = p_i^E - \sum_{j=1}^{N} p_j a_{it}^j - \sum_{i'} p_i^E b_{i'it} - \mu_{it} \quad (18)$$

The top-down nesting structure of the production functions is exemplified at the Annex 1.

3. SCENARIO DEFINITION AND POLICY ANALYSIS

3.1. Some technological considerations

In TD-BU-E3 DGEM we have eight different technologies for electricity production, divided into existing and new vintage technologies, and also categorized as renewable (or green) or not renewable.

The existing electricity production technologies are: Gas Power Plants, Oil Power Plants, Coal Power Plants, Hydro Power Plants, and Bio-Wind Power Plants, where the latter accounts for a composite of existing Biomass and Wind electricity production power units. At the Figure 1 the benchmark production shares of the existing technologies for the year 2005 are shown.

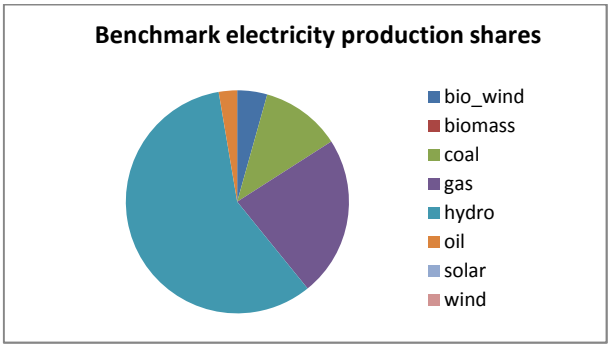


Figure 1: Benchmark electricity production shares

For the future power production we are envisaging the so called new vintage technologies, namely, new wind, new biomass and solar/photovoltaic.

Here the terms new wind and biomass should be understood to be tentative names more the end-of-the pipe technologies that are assumed to be more efficient than the existing but also more costly.

We made assumption that the existing power plants will be functioning in the future and the new technologies will be entering the market after the old have exhausted the limit of their resource allocation. For the existing Bio-Wind technology we have imposed a limit at a level of 2.5 times the value of its benchmark electricity production. Similarly, based on the limiter resource availability, the Hydro Power production was limited to 1.4 times its benchmark production level. According to the trend analysis the production of the coal power plants does not change much and oil power plants are going out of market.

The new renewable technologies have an imposed potential of their maximal contribution to the total electricity production, namely, the new Wind - 7%, new biomass - 15%, and the new solar - 20%.

For the technologies the relative prices per unit of electricity produced have been ranked from the cheapest, hydro power, to the most expensive, new solar which is assumed to be 2.2 more expensive than the hydro. The other technologies are lying in between this range.

The advanced renewables are assumed to be not active at the beginning of the period mainly because they are supposed to be technologically available at a later stage and because they are relatively quite expensive.

3.2. Baseline Scenario

Scenario assumption related to the **adaptation** is the de-coupling of electricity demand from the economic growth. This is assumed to be done by energy and resources conservation in the sense of sustainable development, by changing consumption pattern and habits, etc. – all that are long term measures related to socio economic changes. The growth of total electricity production, shown at Figure 3, is assumed to be 0.7% per year, hence decoupled from the assumed economic growth of 0.9%/year. Just for comparison – till 2008 electricity demand in Austria were growing with 1% per year.

The Scenario assumptions for the main fuel inputs in the power production till the year 2050 are based on energy supply analysis by Kratena and Wrüger (2005) (Figure 2).

The main features of this scenario are:

- doubling the natural gas input for power production,
- hard coal use - almost constant,
- quadrupling the wind and biomass use and
- gradual extinction on fuel oil use in the power plants.

The quadrupling of fuel wood and wind electricity seem to be realizable because the available wind energy potential has been evaluated at 14 - 50 PJ and the fuel wood availability at 30 Mio m3 or 232 PJ (Hantsch and Moidl 2007; Balabanov 2008).

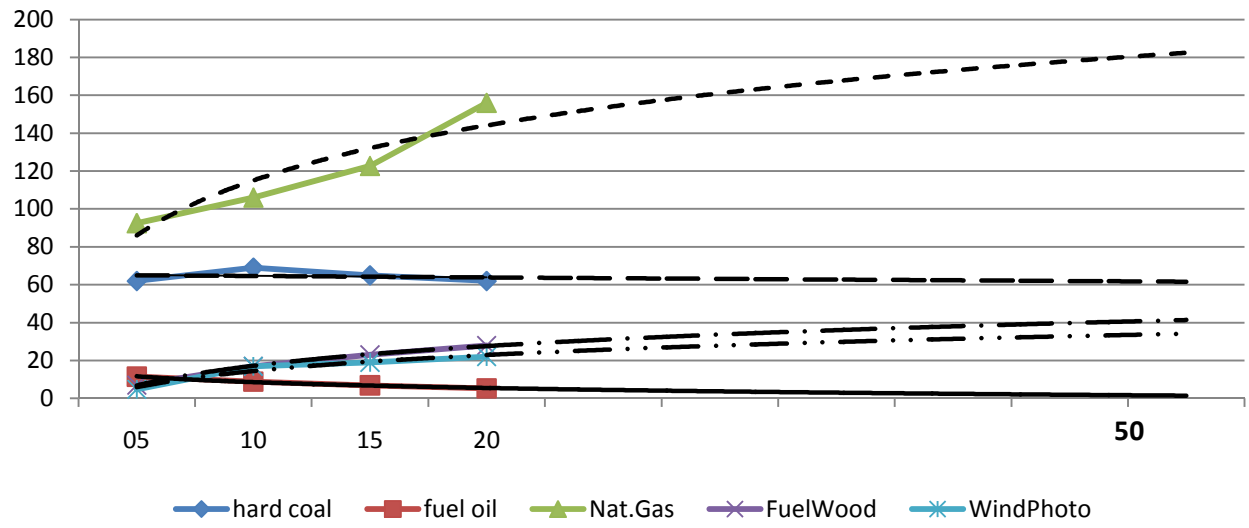


Figure 2: Assumptions for the main fuel inputs till the year 2050 (in PJ)

As said the growth of total electricity production, at Figure 3, is assumed to be decoupled from the economic growth of 0.9%/year so that we are coming to a growth index of 1.64 for electricity production over the 50 year period. In the baseline scenario renewables will increasing their part of the production but at the historical growth rate – reaching approximately 9% by 2050.

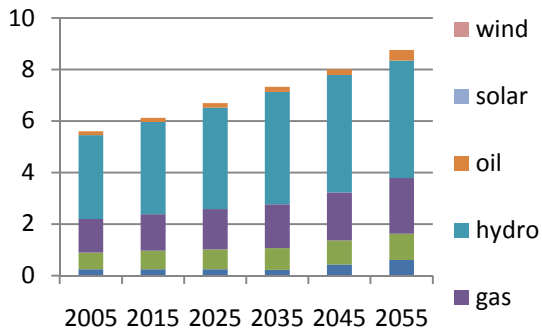


Figure 3: Structure of the power production

3.3. Green quota scenario

A part of the integrated energy and climate change policy guidelines, as adopted by the EU in December 2008 (DG for Energy and Transport 2008), is the obligation by the member states for covering an average of 20% of their total energy needs from renewable sources. Therefore each country agreed to fulfill a different renewable energy quota by 2020.

The target for Austria by 2020 is 34% whilst for the year 2005 it was 23.3% of the total energy use. In so far as this target is recognizing the hydro power as green energy and the *Hydro share for 2020 is projected to be 14.73%* in fulfilling the quota obligation an accelerated growth rate of other renewables would be need in order to reach around 20% by 2020 which is seen by WKOE (2008) as difficult.

In this paper the green quota scenario is attempting to simulate the impacts on the technology mix of the Austrian electricity sector of increasing the share of renewables in the electricity production up to 30% by 2050.

By running the TD-BU-E3 DGEM under the above assumption we have as an output the changes in the main indicators as shown at Figure 4. The growth of the power production indexed with 1.66 is following quite closely the scenario assumption and around 2030 there is a small bump. This is result of the exhaustion of the conventional hydro and bio-wind resources and the slum is due to the significant subsidies needed for the start up of the new wind and biomass technologies.

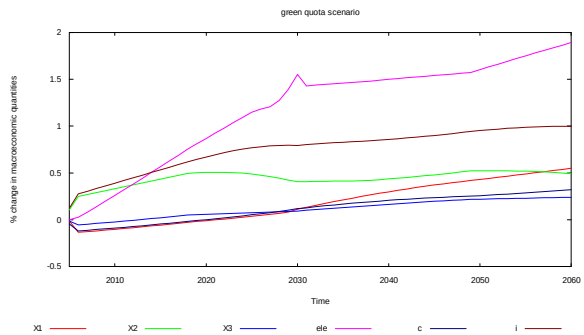


Figure 4: Model output for the main indexes

The accelerated development of the agricultural sector (X1) is a result of the demands of agricultural inputs by the biomass technologies while heavy industry´s production (X2) is slightly declining due the general trend in exporting/downsizing the energy intensive industries.

The growth of investment is following closely the growth of the electricity output and this is due to the

high capital intensity of the power sector. It is quite indicative that the consumption is growing, albeit at a lower rate, despite the significant investment demand.

Here is to be said that by 2030 the share of renewables (without hydro) is reaching 0,184 and by 2050 - 0,289.

To summarize: achieving the quota of close to 30% by 2050 is feasible and there are sufficient quantities of potential renewable resources for that purpose. It also seems that the economic burden is bearable and the welfare is growing.

The next figure shows the electricity production structure by the different technologies in TWh for graphical reasons the dominating Hydro power production is not shown at Figure 5, since it would be depressing the view. The scenario run resulted in steady increase of hydro power production of up to 50 TWh by 2020 when in it reaches its imposed production limit.

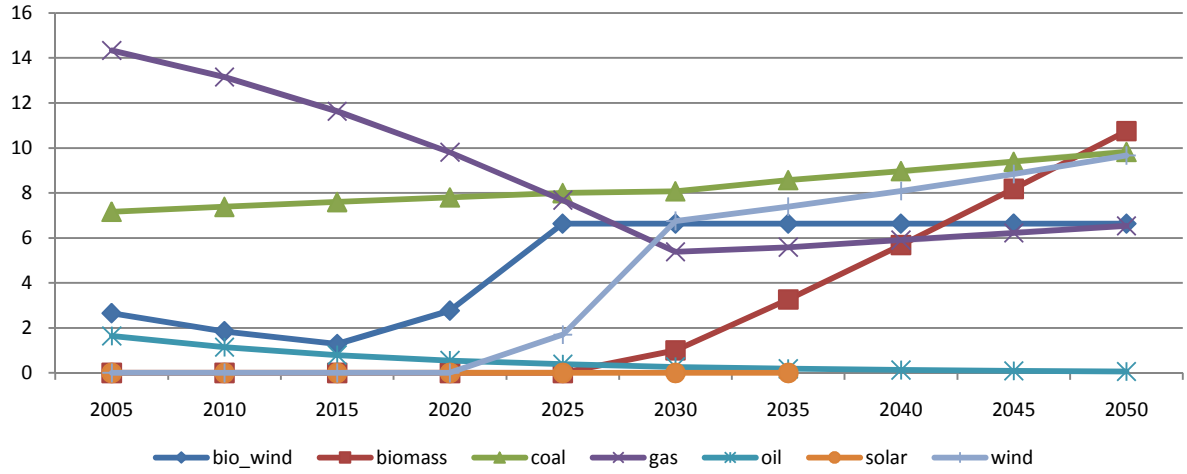


Figure 5: Production (in TWh) of the conventional and renewable energy technologies

Few years later – by 2025 - the bio-wind is also reaching its production limit which results in the output rise by the conventional bio-wind technologies and that is opening the way to entering the market for the new wind and new biomass – the so called backstop technologies.

This start up of the new and expensive technologies result in a jump of the subsidy rate for green technologies, see Figure 6, first in 2025 at the level of 8% from the electricity production cost. When new Vintage reaches its potential, in 2030 there is another jump in subsidy rates reaching to 14%, so that new biomass technologies could start producing electricity.

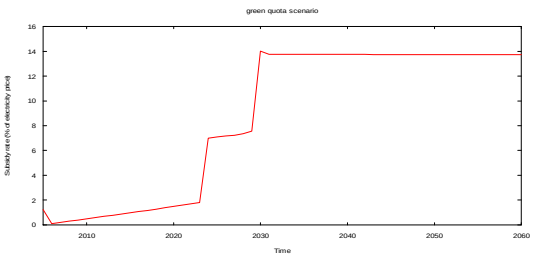


Figure 6: The subsidy rates for the green technologies

As a result of these developments by 2030 the share of renewables in the electricity production (including hydro) is reaching **0,825** or without hydro **0,184** and by 2050 the same share without hydro is **0,289,** while the share (including hydro) remains at 0,825.

3.4. Carbon Taxation (double dividend) Scenario

The **greenhouse gases** are measured in megatons of Carbon dioxide equivalency (MCO_2eq) and there are a number of alternative tax instruments for reducing its emissions.

Carbon dioxide equivalency is a quantity that describes, for a given mixture and amount of greenhouse gas, the amount of CO_2 that would have the same global warming potential (**GWP**), when measured over a specified timescale (generally, 100 years). Carbon dioxide equivalency thus reflects the time-integrated radioactive forcing, rather than the instantaneous value described by **CO_2e**.

For example, the GWP for methane over 100 years is 25 and for nitrous oxide 298. This means that emissions of 1 million metric tons of methane and nitrous oxide respectively are equivalent to emissions of 25 and 298 million metric tons of carbon dioxide.

Over the last decade, several EU Member States have levied some type of carbon tax in order to reduce greenhouse gas emissions from fossil fuel combustion contributing to anthropogenic global warming (OECD 2001).

In this context, the debate on the **double dividend** hypothesis has addressed the question of whether the usual trade-off between environmental benefits and gross economic costs (i.e. the costs disregarding environmental benefits) of emission taxes prevails in economies where distortionary taxes finance public spending.

Emission taxes raise public revenues which can be used to reduce existing tax distortions. Revenue

recycling may then provide prospects for a double dividend from emission taxation (Goulder 1995):

Apart from an improvement in environmental quality (the first dividend), the overall excess burden of the tax system may be reduced by using additional tax revenues for a revenue-neutral cut of existing distortionary taxes (the second dividend).

If – at the margin – the excess burden of the environmental tax is smaller than that of the replaced (decreased) existing tax, public financing becomes more efficient and welfare gains will occur.

The setting of TD-BU-E3 DGEM for simulating Carbon Taxation Scenario differs slightly from the original setting for the Baseline Scenario, e.g., final consumption is being split into public (governmental) and private (household) consumption, where public consumption is estimated at a level of 25% of total consumption.

Therefore a new production activity is defined, indicating a public good (e.g. infrastructure, healthcare, etc.), which is then consumed by the Private households or firms in the economy.

In our dynamic policy simulations, we investigate the economic effects of carbon taxes that are set sufficiently high to reduce carbon emissions by 20% compared to the base year emission level. The figure bellow is showing the rate of decarburization of the produced electricity, namely the reduction of CO2 emissions per TWh of produced electricity.

While keeping public good consumption at the base-year level, the additional carbon tax revenues can be recycled in three different ways:

(i) a reduction in the distortionary labor tax (labeled as "**TL**")
(ii) a cut in the distortionary consumption tax (labeled as "**TC**")
(iii) a lump-sum refund to the representative household (labeled in the Figure as "**LS**")

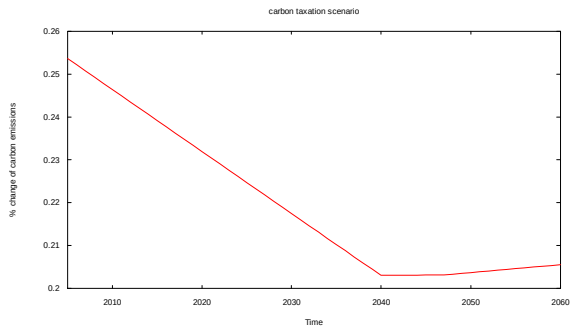


Figure 7: Trajectory of CO2 emissions per unit electricity produced

As seen at the Figure 8 – in line with the undisputed weak double dividend hypothesis (Goulder 1995) - the reduction of the distortionary consumption or labor taxes (TL) is superior in efficiency terms as compared to a lump-sum recycling of carbon tax revenues. In our dynamic simulation, we even obtain a strong double dividend from revenue-neutral cuts in distortonary taxes (TL): Reflecting the larger marginal excess burden of the initial labor tax vis a vis the initial consumption tax, labor tax recycling is distinctly more beneficial than consumption tax recycling. The Figure 8 provides the consumption trajectories for the three different recycling options. In the case of reduction in the distortionary labor tax (**TL**) the consumption levels are increasing over a long period of time. To a lesser extend the same applies to the case of a cut in the distortionary consumption tax (labeled as "TC"): The reduction in the distortionary lump-sum refund to the representative household (labeled as "LS") tends to reducing consumption and respectively the welfare.

Hence, only for the case of labor tax recycling, we could assume the existence of a strong double dividend.

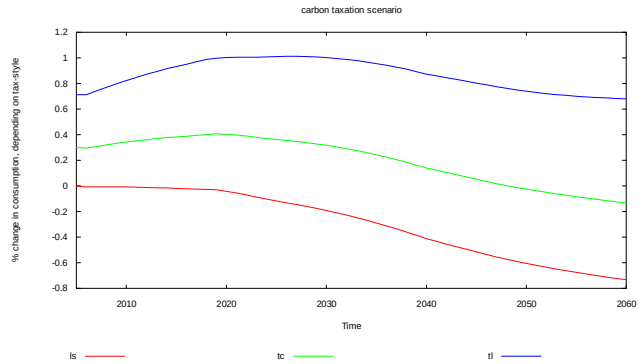


Figure 8: Carbon Taxation Scenarios

Figure 9 shows the associated carbon tax rates, or the marginal abatement cost (MAC), to achieve the target emission reductions. The computed maximum MAC of bellow EUR 100 that correlates very well with other multi country studies for the EU region, e.g. the Marginal Abatement Costs (MAC) levels have been estimated by the EU´s "Impact Assessment of the EU's objectives on climate change and renewable energy for 2020" (EC 2008) to be around € 90/t CO2.

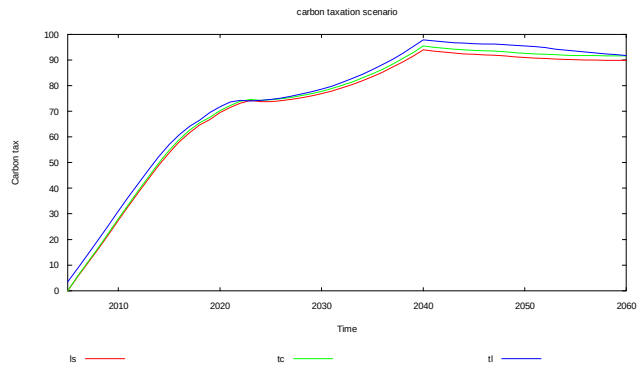


Figure 9: Dynamics of the carbon tax rates/MAC

MAC – as the direct incentive for emission mitigation in production and consumption – increase with the stringency of the emission constraint but hardly differ across recycling variants.

In the dynamic analysis of environmental tax reforms, we impose a linear reduction of carbon emissions compared to baseline emission levels by 20% between 2005 and 2040, holding the percentage reduction vis--vis the Baseline and keeping it constant thereafter.

4. CONCLUSIONS

By adapting and extensively validating the newly developed Top/Down -BU for Bottom-up E3 (energy, environment, economy) dynamic general equilibrium model (TD-BU-E3 DGEM) we assessed the long term impacts on the macroeconomic and sectoral structural components of two alternative policy instruments for responses to climate change and for promotion of renewable energy sources:

Green quota, and

Carbon Taxation (double dividend)

In our baseline Scenario, as a part of the adaptation strategy, we assumed de-coupling of electricity demand growth from the economic growth.

In the model we have introduced 5 existing electricity production technologies, namely: Gas Power Plants, Oil Power Plants, Coal Power Plants, Hydro Power Plants, and Bio-Wind Power Plants (a composite of existing Biomass and Wind electricity production power units).

The new vintage technologies, namely, new wind, new biomass and solar/photovoltaic – are tentative names and should be better seen as the end-of-the pipe technologies that are assumed to be more efficient than the existing but also more costly.

The model runs for the Green quota scenario have shown that as a result of the inversing demands of agricultural inputs by the biomass technologies there is accelerated development of the agricultural sector while heavy industry´s production is slightly declining due the general trend in exporting/downsizing the energy intensive industries.

The growth of investment is following closely the growth of the electricity output and this is due to the high capital intensity of the power sector. It is quite indicative that the consumption is growing, albeit at a lower rate, despite the significant investment demand.

Here is to be said that by 2030 the share of renewables in the electricity production (without hydro) is reaching 0,184 and **by 2050 - 0,289** and the renewables share (including hydro) is 0,825.

To summarize: achieving the quota of close to 30% by 2050 is feasible and there are sufficient quantities of potential renewable resources available for electricity production. It also seems that the economic burden is bearable and the welfare is growing.

The **double dividend** hypothesis has addressed the question of whether the usual trade-off between environmental benefits and gross economic costs (i.e. the costs disregarding environmental benefits) of emission taxes prevails in economies where distortionary taxes finance public spending.

Emission taxes raise public revenues which can be used to reduce existing tax distortions. Revenue recycling may then provide prospects for a double dividend from emission taxation.

While keeping public good consumption at the base-year level, the additional carbon tax revenues can be recycled in three different ways:

(i) a reduction in the distortionary labor tax
(ii) a cut in the distortionary consumption tax
(iii) a lump-sum refund to the representative household

The results of the simulations are showing that the reduction in the distortionary labor tax is leading to increases over a long period of time of the consumption levels. To a lesser extend the consumption increases in the case of a cut in the distortionary consumption tax. From the other side the reduction in the distortionary lump-sum refund to the representative household tends to reducing consumption and respectively the welfare.

Hence, only for the case of labor tax recycling, we could assume the existence of a strong double dividend.

ACKNOWLEDGMENTS

The Authors are grateful for the generous support to their work provided by the Innovation Fund of the Austrian Research Promotion Agency; www.ffg.at/modsim

REFERENCES

Balabanov Todor, 2008, The enhanced use of wood-biomass: Macroeconomic, sectoral and environmental impacts, WP 227, IHS, Vienna, Austria

BMLFUW, 2009, Erneuerbare Energie 2020 - Potenziale & Verwendung in Österreich ,Page 37, http://www.energiestrategie.at/images/stories/pdf/02_bmlfuw_09_erneuerbare2020.pdf

Böhringer Christoph, Hoffmann Tim, Rutherford Thomas F., 2006, Alternative Strategies for Promoting Renewable Energy in EU Electricity Markets, Centre for European Economic Research (ZEW), Mannheim, Germany, http://www.mpsge.org/qtool/electricity.pdf

Böhringer Christoph, Rutherford Thomas F., 2003,. Computable General Equilibrium Analysis: Opening a Black Box, ZEW Discussion Paper No. 03-56

Böhringer Christoph, Rutherford Thomas F., 2007, Combining Top-Down and Bottom-up in Energy Policy Analysis: A Decomposition Approach, ZEW Discussion Paper No. 06-007, Mannheim, Germany, ftp://ftp.zew.de/pub/zew-docs/dp/dp06007.pdf

DG for Energy and Transport , 2008, Austria – Renewable Energy Fact Sheet (23 January 2008), EUC, Brussels, http://www.energy.eu/renewables/factsheets/2008_res_sheet_austria_en.pdf

EC, 2008, “Impact Assessment of the Package of Implementation measures for the EU's objectives on climate change and renewable energy for 2020”, http://ec.europa.eu/environment/climat/climate_action.htm

Goulder, L. H., 1995, “Environmental taxation and the double dividend: A readers guide,” International Tax and Public Finance, 1995, 2, 157–183.

Hantsch Stefan, Moidl Stefan ,2007 , Das realisierbare Windkraftpotenzial in Österreich, bis 2020, St.Pölten

Kurt Kratena, Michael Wüger ,2005, Energieszenarien für Österreich bis 2020, WIFO, Vienna

Markusen James R., 2002, General-Equilibrium Modeling using GAMS and MPS/GE: Some Basics, Boulder, Colorado

Rosenthal Richard E., 2007, GAMS | A User's Guide, Tutorial, GAMS Development Corporation, Washington, DC, USA,

Rutherford Thomas F., 2002, Applied General Equilibrium Modeling with MPSGE as a GAMS Subsystem: An Overview of the Modeling Framework and Syntax,

Rutherford Thomas F., 2002, Mixed Complementarity Programming with GAMS, Lecture Notes, Colorado School of Mines

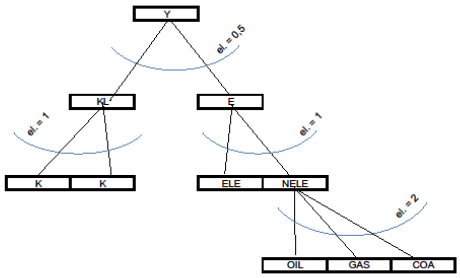
Rutherford Thomas F., 2008, Calibrated CES Utility Functions: A Worked Example, Department of Management, Technology and Economics, ETH Zuerich March 24, 2008

WKOE (Wirtschaftskammer Österreich), 2008, EU-Klima- und Energiepaket(Vorschlag für eine Richtlinie zur Förderung der Nutzung von Energie aus erneuerbaren Quellen) Page No 3 http://wko.at/up/enet/stellung/WKO-Position_RES.pdf

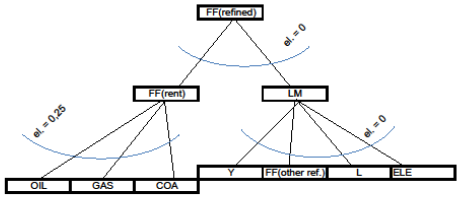
ANNEX.1: TD-BU-E3 DGEM: NESTING STRUCTURE OF THE PRODUCTION FUNCTIONS

TOP DOWN NESTING STRUCTURE.

STRUCTURE FOR SECTORAL PRODUCTION:



STRUCTURE FOR FOSSIL FUEL PRODUCTION



COMPARATIVE STUDY WITH AND WITHOUT SENSOR OF SPEED OF A COMMAND BY DTC OF A VARIABLE RELUCTANCE MOTOR WITH SMOOTH STATOR

M. Birame[(a)], L. Mokrani[(b)], M.kadjoudj[(c)],A. Naamane[(d)], N.K.Msirdi[(d)]

[(a)] Amar Telidji University , Laboratory of valorization of renewable energies and environment agrissifs, Department of Electrical Engineering, Laghouat 03000, ALGERIA
[(b)] Amar Telidji University, Laboratory Materials Semiconductor and Dielectric, Department of Electrical Engineering, Laghouat 03000, ALGERIA
[(c)] Hadj Lakhdar University, Laboratory of Electrical engineering. Department of Electrical engineering, Street Chahid Med El-Hadi Boukhlouf, Batna 05000, ALGERIA
[(d)] Paul Czanne University , LSIS - Laboratory of the System and Information sciences,marseille, FRANCE

[(a)]bir_moh@yahoo.fr, [(b)]Mokrani lakhdar@hotmail.com, [(c)] m.kadjoudj@lycos.com,
[(d)] aziz.naamane@lsis.org, [(d)] nacer .msirdi@lsis .org

ABSTRACT
This paper describes a comparative study with and without sensor of speed by simulation of an control by the Direct Torque Control of switched reluctance motor (SRM) with a smooth stator and massive rotor. Because of the design and reluctant effect between the stator and rotor, the torque of is pulsating. The great difficulty is to control these undulations which are undesirable and generally vermin with the good performance of the device.

Keywords: DTC, switched reluctance motor (SRM), modeling, with sensor and without sensor of speed,

1. INTRODUCTION

In recent years, especially the need for variable speed in servo applications has increased the interest to theSRM. The major reasons of these interests for SRM are robustness, high efficiency, low cost, high speed, simple structure, easy to maintain, high torque in low speed, simple power converter circuits with reduced number of switches, excellent controllability and smaller dimension of the motor in comparison to the other motors.

Although it has been known for a long time, SRM has not been very often used because of acoustic noise that comes from torque dip and detection of rotor position [KIM]. SRM has been started to be used with the development of power electronic devices. If the problems of the SRM can be solved,

The method of control of order DTC is based on the algorithmic principles of the vectorial orders applied to the electric actuators of the asynchronous, synchronous motor type and other motorizations with alternative source. Here, decoupling between flow and the couple are obtained by preparing a logical table of the orders of commutation of the inverter (sa,sb,sc) [Kioyyur].

The calculation of the flow quantities (ϕsq, ϕsd,) and torque is obtained from the acquisition and processing, through the Concordia transform, only electrical voltage (uo) and currents (isa , isb). Controls dynamic flux quantities (ϕref) and torque (Celm) are made from two simple hysteresis comparators and a calculation of the sectoral location of the flux [Ameur].

.

2. PRINCIPLE OF OPERATION OF A MRV

The operation of variable reluctance machines is based on the principle of magnetic attraction. It can be described from a single basic structure similar to that presented in Figure 1 [KIM].

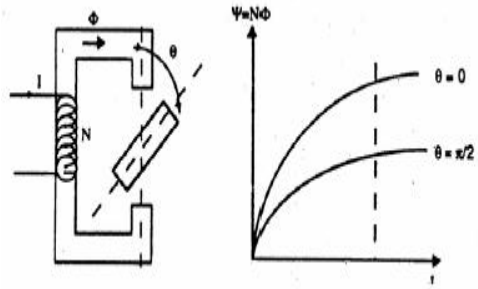


Figure1: Elementary structure of a réluctant system

Consider the simple magnetic circuit of Figure 1, consisting of a fixed ferromagnetic cylinder and a mobile ferromagnetic part of which the axis is indicated by the angle θ,

The transition from an I in N turns of the coil, creating a magnetomotive force fmm defined by [KIM]:

$$F = NI$$

The Flow through the N turns of the coil depends on the strength of different portions of the circuit, this resistance is called reluctance R, it is the sum of the reluctances of the cylinder head and the moving part, as that of the air-gap which is generally dominant.

The reluctance can be represented by another parameter known as perméance, and defined by [KIM] :

$$P = \frac{1}{R}$$

By considering a flux variation Ψ compared to current I, we can write [KIM]:

$$\Psi = L\,I = N^2 I\,P$$

L: is the inductance of the coil clean..
For a current I in rolling up, the variation of the air-gap defines two limiting values of flow (see Figure 1). For θ=0 and θπ , flow is maximum, and it becomes minimal for θ = ± π/2.

This structure has two remarkable rotor positions. A position of opposition, for which the magnetic circuit presents a maximum reluctance.

A position of conjunction where the magnetic circuit has a minimal reluctance [Kioyyur].

If the rotor is in an intermediate position between the opposition and the conjunction, the injection of a current in the exciting winding modifies the state of the system which tends to minimize its energy and thus to obtain a minimal reluctance. Thus the rotor turns of an angular step to take the position of conjunction.

3. MODELING OF THE SRM

The model of the SRM with smooth stator is [Kioyyur], is made up mainly in the electric equations of the machine given by:

$$\begin{aligned} v_d &= R_s i_{sd} + \frac{d\phi_d}{dt} - p\frac{d\theta}{dt}\phi_q \\ v_q &= R_s i_{sq} + \frac{d\phi_q}{dt} + p\frac{d\theta}{dt}\phi_d \end{aligned} \qquad (1)$$

Where flows following the axis d and the axis q are given by:

$$\begin{bmatrix} \phi_d \\ \phi_q \end{bmatrix} = \begin{bmatrix} L_d & 0 \\ 0 & L_q \end{bmatrix} \begin{bmatrix} I_{sd} \\ I_{sq} \end{bmatrix} \qquad (2)$$

The matrix form of (1) is:

$$\begin{bmatrix} v_d \\ v_q \end{bmatrix} = R_s \begin{bmatrix} i_{ds} \\ i_{qs} \end{bmatrix} + \begin{bmatrix} L_d & 0 \\ 0 & L_q \end{bmatrix} \frac{d}{dt} \begin{bmatrix} i_{ds} \\ i_{qs} \end{bmatrix} + p\Omega \begin{bmatrix} 0 & -L_q \\ L_d & 0 \end{bmatrix} \begin{bmatrix} i_{ds} \\ i_{qs} \end{bmatrix} \qquad (3)$$

The torque of a SRM can be expressed by:
:

$$C_e = p(\phi_d i_{sq} - \phi_q i_{sd}) = p\left((L_d - L_q) I_{sd} \quad I_{sq}\right) \qquad (4)$$

On the other hand, the mechanical equation of the machine is:

$$J\dot{\Omega} = C_e - C_r - f_r \Omega \qquad (5)$$

4. DESCRIPTION OF THE DIRECT TORAUE CONTROL

The method of Direct Torque Control (DTC) was introduced in 1985 by Takahashi and Depenbrock especially for asynchronous machines [KIM]. Then, several studies have developed more precise knowledge of this command. This technique of command was also applied on the variable reluctance machines. The direct control of a couple of MRV is based on identifying "direct" in the command sequence used to switch a voltage inverter. This choice is usually based on the use of hysteresis regulators whose function is to control the system state, namely here by the amplitude of stator flux and electromagnetic torque. This type of strategy is therefore classified in the category of orders in amplitude, In opposed to the laws of contracts in duration more traditional and time-based adjustment of the average value of voltage vector pulse width modulation (PWM) [Takahashi],[Zhong].

The method of estimate consists to measure the currents and the stator tensions of the machine, and to employ them in the law of direct order of couple (DTC). This method uses the figuration of order flow and torque for determining the sequence of power to impose the three-phase PWM inverter vector [Takahashi] and [Zhong].

The classical DTC proposed by Takahashi, is based on the following algorithm:

1. To divide the time domain during Te reduced period of duration (of the order of tens of µs);
2. For each shot clock, measuring the line currents and voltages in phase of SRM;
3. Reconstruct the components of stator flux vector;
4. Estimate the electromagnetic torque of SRM, through the estimate of the vector of flow stator and the measurement of the currents of lines;
5. To introduce the difference Δ□s between the flow of reference □s* and the flow estimated □s^ into a comparator at hysteresis at two levels, which generates at its output value +1 to increase flow and 0 to reduce it.

5. APPLICATION OF THE DTC TO THE SRM

The reference mark related to the stator makes it possible to estimate flow and the couple, which allows the knowledge of the amplitude and the position of stator flow.

$$\begin{cases} \phi_{s\alpha} = \int_0^t (v_{s\alpha} - R_s i_{s\alpha})dt \\ \phi_{s\beta} = \int_0^t (v_{s\beta} - R_s i_{s\beta})dt \end{cases} \qquad (6)$$

The DTC is deduced while being based on the two approximations described by the formulas (7) et (8) [Takahashi] :

$$\bar{\phi}_s(k+1) \approx \bar{\phi}_s(k) + \bar{V}_s T_E \quad \rightarrow \Delta\bar{\phi}_s \approx \bar{V}_s T_E \tag{7}$$

It was mor

$$\begin{cases} \hat{\phi}_s = \sqrt{\hat{\phi}_{s\alpha}^2 + \hat{\phi}_{s\beta}^2} \\ \angle \hat{\phi}_s = arctg \dfrac{\hat{\phi}_{s\beta}}{\hat{\phi}_{s\alpha}} \end{cases} \tag{8}$$

According to the formula (6), we can neglect the voltage drop private to the resistance of the stator. With this assumption, the variation of the flux is proportional to the applied stator voltage.. A voltage applied in the same direction as the flux vector increases the module to it and vice versa.

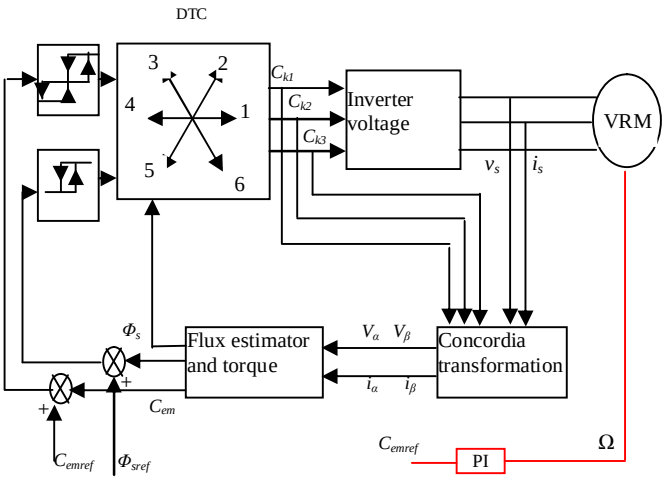


Figure 2: Diagram of an order From a DTC of a SRM Powered by a voltage inverter three-phase PWM sensorless From speed.

The DTC is based on the direct determination of the sequence of command used to switch a voltage inverter.

This choice is based on the use of hysteresis comparators whose function is to control the system state, namely the amplitude of stator flux and electromagnetic torque [Ameur], [Carlos].

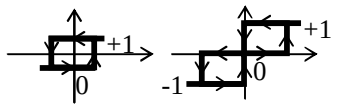


Figure 3: Hysteresis comparators used to control the flux and torque.

The equation (7) implies that the tip of the stator flux vector moves along a straight line whose direction is given by the vector of applied voltage, as shown in Figure (3).

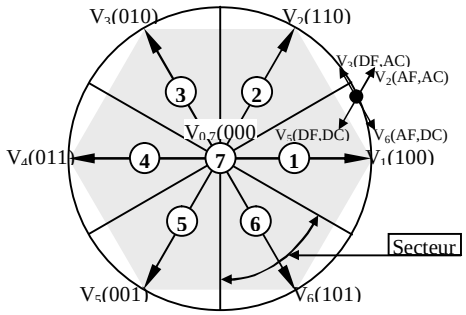


Figure 4: Evolution Of the extreme of ϕ_s

Table (I) presents the sequence for each position, knowing that the error between the reference flux and the estimated flux is introduced into a comparator with hysteresis at two levels, which delivers 1 if the error is positive and 0 if it is negative. Similarly, the error between the reference torque and estimated torque is introduced into a comparator with hysteresis at three levels, which delivers 1 if positive, 0 if zero, and -1 if it is negative [Ameur] and [Carlos].

Table 1: Table de localisation des vecteurs voltage

$\Delta\phi_s$	ΔC_e	S_1	S_2	S_3	S_4	S_5	S_6
1	1	V_2	V_3	V_4	V_5	V_6	V_1
	0	V_0	V_7	V_0	V_7	V_0	V_7
	-1	V_6	V_1	V_2	V_3	V_4	V_5
0	1	V_3	V_4	V_5	V_6	V_1	V_2
	0	V_7	V_0	V_7	V_0	V_7	V_0
	-1	V_5	V_6	V_1	V_2	V_3	V_4

where, Si=1,…, 6 are the sectors of localization of the stator vector of flow

6. SIMULATION RESULTS

Presented in the following simulation results of an control by DTC of SRM, supplied by a three-phase inverter voltage vector to PWM. Without and with the regulation the speed which is done by a PI regulator.

For the same parameters of the regulator, we simulated a no load starting,step load change and we reversed the direction of speed.

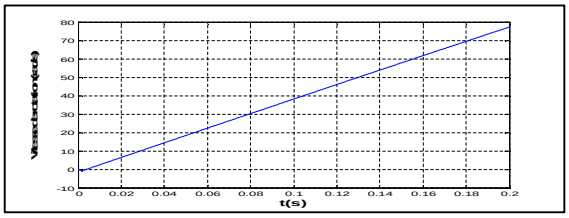


Figure.5: Starting without loop of regulation speed

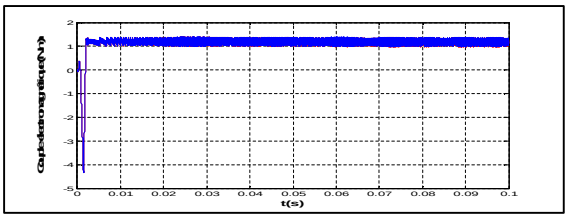


Figure.6: Electromagnetic torque (estimated and real)

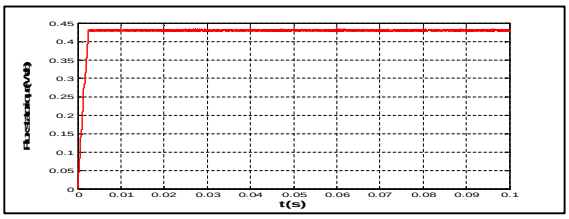


Figure.7: Stator flux (estimated and real)

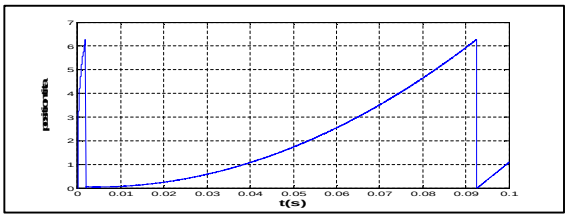


Figure.8 Evolution of the positon of the stator vector of flow (estimated and real)

The results of simulation presented by the Figures (5, 6, 7,8), for the starting without loop of regulation speed with a torque of reference of 1.2Nm, we find that the speed does not answer on the other increases and exceeds the value of 50 rad/s over against the stator flux (estimated and real) follows its reference of 0.43Web.

The results of simulation presented by the Figures (9, 10, 11 and 12), for the (void) starting with inversion of the direction of rotation at the moment 0,25s of 50 rad/s to -50 rad/s, We note that satisfies the speed without overshoot and during the reverse rotation with a short response time because the machine is void and that the inertia is low. With starting, the electromagnetic torque reaches its maximum value limited (5Nm) and stabilizes at a value of almost zero in the regime established.

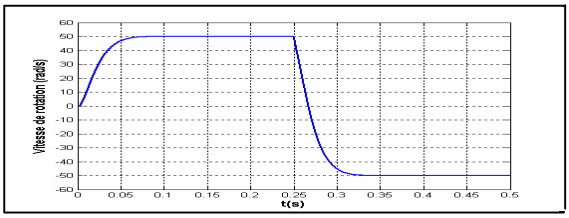


Figure.9: starting with inversion of the direction of rotation

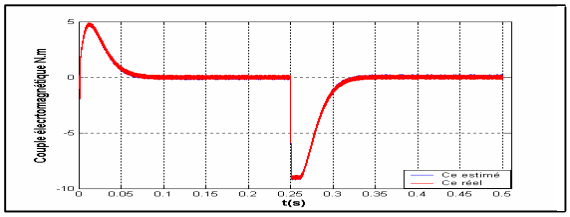


Figure.10: Electromagnetic torque (estimated and real)

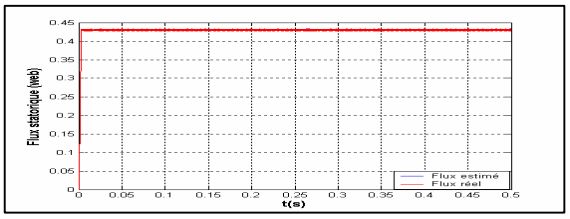


Figure.11: Stator flux (estimated and real)

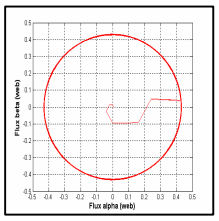


Figure.12: flux in the (α, β).

The following figures (13, 14, 15 and 16) show the introduction of a load at t = 2s. The machine is loaded by a resistant torque equal to (2 Nm). The electromagnetic torque reaches its maximum value limited (5 Nm) and stabilizes at a value of (2 Nm).

Note that the electromagnetic torque follows a decent record of the couple with relatively large undulations, while the flux follows well the reference speed and responds with a small overshoot at the introduction of resistant torque with a short response time. Figure (d) illustrates the flow in the (α, β).

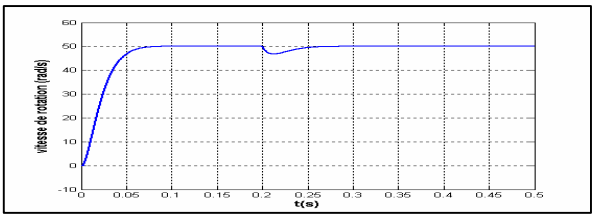


Figure.13: starting with introduction of a load

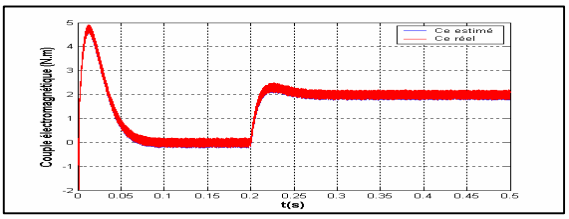


Figure.14 : Electromagnetic torque (estimated and real).

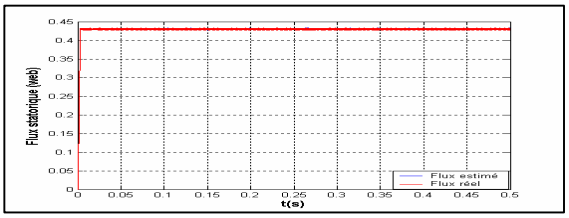


Figure.15: Stator flux (estimated and real)

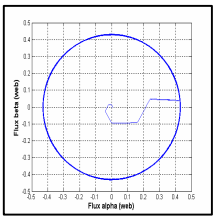


Figure.16 : flux in the (α, β).

7. CONCLUSION

In this article, we made a comparative study by simulation of the DTC applied to the reluctance variable machine with a smooth stator and massive rotor without and with the loop of the speed of the machine, and we showed that the machine meets soon start to empty and change of direction of rotation and the introduction of a resistant torque

Table 2: VRM parametrs [Zhong]

Symbole	Valeurs (S.I)
f	50
P_n	1000
p	2
r_s	1.0
L_d	0.072
L_q	0.028
J	0.003

REFERENCES

Ameur, A. Mokhtari, B. Mokrani, L. Boukhtache, S. 2005 Commande sans capteur de vitesse par DTC d'une machine synchrone à aimants permanents dotée par un observateur à mode glissants. *SISEM'05*, May 22-23, Annaba, Algérie.

Carlos, C, 2000. Modélisation Contrôle Vectoriel et DTC, *Commande des Moteurs Asynchrones 1, Edition Hermes Science Europe* .

Kioyyur, B. 2005. Modélisation et contrôle vectoriel de la machine à reluctance variable comparaison entre deux modèles avec et sans amortisseurs, *PCSE'05*, May 09-11, O.E.bouaghi, Algérie.

KIM, M.H, 2002. A Position Sensorless Motion Control System Of Reluctance Synchronous Motor With Direct Torque Control. *Conférence IEEE*, pp. 1148-1153.

Takahashi, I., Ohmori, Y. 1989. High-performance Direct Torque Control of an Induction Motor, *IEEE Transactions on Industry Applications*, Vol. 25, pp. 257-264, March/April.

Zhong, L. Rahman, M. F. Hu, W.Y 1999 A Direct Torque Controller for Permanent Magnet Synchronous Motor Drives. *IEEE Transactions on Energy Conversion*, Vol. 14, pp. 637-642.

AUTHORS BIOGRAPHY

Birame M'hamed Teaching Assistant 2005 Department of Electrical Engineering, Laghouat, Algeria. Member of , Laboratory of valorization of renewable energies and environment agrissifs. He specializes in artificial intelligence applications to power system and Electrical Machines.

Mokrani Lakhder MCC degree from Department of Electrical Engineering, Laghouat, Algeria. Member of Laboratory Materials Semiconductor and Dielectric. He specializes in electrical energy systems, design optimization of electromagnetic devices, and artificial intelligence applications to energy systems as well as electromagnetic field computations in nonlinear systems.

Kadjoudj Mohamed MCC degree from Department of Electrical engineering of Hadj Lakhdar University Batna Algeria., Member of Laboratory of Electrical engineering.

Naamane Aziz MCC degree of Paul Czanne University member of LSIS - Laboratory of the System and Information sciences, Marseille, FRANCE

Msirdi Nacer kouider is a Professor of Paul Czanne University member of LSIS - Laboratory of the System and Information sciences, Marseille, FRANCE

CONTROL OF A PH NEUTRALIZTION PROCESS VIA THE BIG BANG BIG CRUNCH OPTIMIZATION

Tufan Kumbasar[(a)], İbrahim Eksin[(b)], Müjde Güzelkaya[(c)], Engin Yeşil[(d)]

Istanbul Technical University, Faculty of Electrical and Electronics Engineering,
Control Engineering Department, Maslak, TR-34469, Istanbul, Turkey

[(a)]kumbasart@itu.edu.tr, [(b)] eksin@itu.edu.tr, [(c)] guzelkaya@itu.edu.tr, [(d)] yesileng@itu.edu.tr,

ABSTRACT

The control of pH is common in the chemical process and biotechnological industries. The main complexity consists of the nonlinearity reflected in the S-shaped gain curve of the system. However, a model-based controller will inherit all the nonlinearities of pH system. Therefore using the inverse model as a controller in an open loop fashion will produce perfect control if there does not exist any disturbance or parameter variations. In this paper, a new model inversion technique that is based on an evolutionary search algorithm called Big Bang Big Crunch optimization is introduced. Moreover, a hybrid control scheme is proposed where a parallel PI controller, which will be only activated in a model mismatch, is implemented. The beneficial sides of the hybrid control approach based on the proposed model inversion technique are illustrated in a simulation study. This control scheme is applied to control the pH variable in a neutralization process.

Keywords: Inverse model control, Big Bang- Big Crunch optimization, pH neutralization process

1. INTRODUCTION

The control of the neutralization process has been studied for several years but still remains a challenging problem. The control of pH is not just a control problem, but also a chemical equilibrium problem (Wright and Kravaris 1991). It is challenging problem since the pH process inherits nonlinearity, high sensitivity at and near the neutralization point (Pishvaie and Shahrokhi 2006). Control of pH process plays an important task in chemical plants like biological, wastewater treatment and electrochemistry (Fuente et al. 2006). However, modeling of a complex pH process is a difficult. It has been verified that high nonlinear behavior of pH system is caused by titration curve of process stream. Therefore it can be concluded that, the most nonlinear term of process is described by the relationship which describes the neutrality condition.

Several control techniques have been proposed lately. They range from linear controller to nonlinear controllers through fuzzy controllers. Most of the successful linear controllers are based on using multiple linear models (Nystrom et al. 1998). In this technique, the controller design is based linear quadratic (LQ) technique, and then the controllers are combined with two gain scheduling methods. Beside the fact that linear controllers are easy to implement, the best performances are given by nonlinear controllers since the neutralization process inherits highly nonlinear terms. In Loh et al. (2001), an adaptive control strategy is proposed where the control and parameter estimation laws are derived based on a reference model. Moreover, Henson et al. (1994) proposed and implementable version of the indirect adaptive nonlinear control strategy. And Babuska et al. (2002) have proposed a fuzzy self-tuning PI controller for a pH control in a fermentation system, where PI parameters are tuned on-line. Another fuzzy control methodology is proposed by Fuente et al. (2006), in this strategy the pH process is divided in several sub fuzzy regions, so that the operating the process region is determined and according to this information a fuzzy PI controller is designed.

Recently, a new evolutionary computation algorithm named as Big Bang Big Crunch (BB-BC) is presented by Erol and Eksin (2006). The leading advantage of BB-BC is the high convergence speed and the low computation time. The working principle of this method can be explained as the transformation of a convergent solution to a chaotic state and then back to a single tentative solution point. This evolutionary search algorithm was first used as an online adaptation of the fuzzy model by updating the consequent parameters of the model (Kumbasar et al 2008a); moreover this optimization method is used as an inverse fuzzy model control structure (Kumbasar et al 2008b).

In this paper, a model based inversion technique is introduced. It is known fact that inverting the system is an effective way to control nonlinear systems. Therefore a model based open loop control structure is proposed in which the inverse model is used as the controller. In this structure, the output of the inverse model which would be the control signal for the system is generated via an optimization problem. The optimization problem can be defined as to decrease the error between the

process model output and the reference signal. Since the BB-BC optimization algorithm has a high convergence speed and low computational time, the optimal inverse process model control signal is generated within each sampling time. As the process model output converges to the set point; the process output will converge to the set point; unless there does not exist any disturbance or parameter variation in the system. In the case of disturbances of parameter perturbations, a hybrid control structure is proposed in which a PI controller is implemented in order to avoid steady state errors.

In the next chapter, brief information about the BB-BC optimization technique is given. In the third chapter, the BB-BC optimization based inverse model controller is introduced. Moreover the hybrid control structure is presented. Later, the highly nonlinear pH model described. Finally, the performance of this control structure is illustrated.

2. BIG BANG–BIG CRUNCH OPTIMIZATION

The Big Bang-Big Crunch (BB-BC) optimization method is built on two main steps: The first step is the Big Bang phase where candidate solutions are randomly distributed over the search space and the next step is the Big Crunch where a contraction procedure calculates a center of mass for the population. The initial Big Bang population is randomly generated over the entire search space just like the other evolutionary search algorithms. All subsequent Big Bang phases are randomly distributed about the center of mass or the best fit individual in a similar fashion. In (Erol and Eksin 2006), the working principle of this evolutionary method is explained as to transform a convergent solution to a chaotic state which is a new set of solutions. The procedure of the BB-BC optimization is given in the table below:

Table 1: BB-BC Optimization Algorithm

Step 1 *(Big Bang Phase)*
An initial generation of N candidates is generated randomly in the search space.
Step 2
The cost function values of all the candidate solutions are computed.
Step 3 *(Big Crunch Phase)*
The center of mass is calculated. Either the best fit individual or the center of mass is chosen as the point of Big Bang Phase.
Step 4
New candidates are calculated around the new point calculated in Step 3 by adding or subtracting a random number whose value decreases as the iterations elapse.
Step 5
Return to Step 2 until stopping criteria has been met.

After the Big Bang, a contraction procedure is applied during the Big Crunch. In this phase, the contraction operator takes the current positions of each candidate solution in the population and its associated cost function value and computes a center of mass. The center of mass can be computed as:

$$x_c = \frac{\sum_{i=1}^{N} \frac{1}{f^i} x_i}{\sum_{i=1}^{N} \frac{1}{f^i}} \tag{1}$$

where x_c = position of the center of mass; x_i = position of candidate; f^i =cost function value of candidate *i*; and *N* = population size. Instead of the position of the center of mass, the best fit individual can also be chosen as the starting point in the Big Bang phase.

The new generation for the next iteration Big Bang phase is normally distributed around x_c

$$x_i^{new} = x_c + \sigma \tag{2}$$

where x_i^{new} = the new candidate solution *i*; and σ standard deviation of a standard normal distribution. The standard deviation decreases as the iterations elapse according to the following formula

$$\sigma = \frac{r\alpha(x_{\max} - x_{\min})}{k} \tag{3}$$

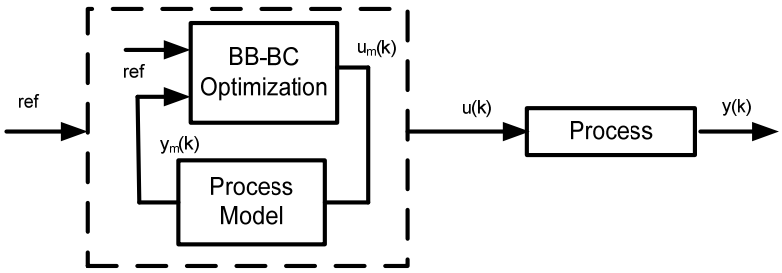
where *r* is random number; α is a parameter limiting the size of the search space, $x_{\max}$ and $x_{\min}$ are the upper and lower limits; and *k* is the number of the iterations. Therefore, the new point is generated as follows:

$$x_i^{new} = x_c + \frac{r\alpha(x_{\max} - x_{\min})}{k} \tag{4}$$

Since normally distributed numbers can be exceeding ±1, it is necessary to limit population to the prescribed search space boundaries. This narrowing down restricts the candidate solutions into the search space boundaries (Erol and Eksin 2006).

3. CONTROLLER STRUCTURE

In the first part of this section the proposed BB-BC based Inverse Model controller will be introduced. This controller structure assumes that there is neither disturbance nor model mismatch. It is common known fact it is almost impossible to perfectly model industrial process. Since this proposed control structure is based on the obtained model, the controller will not be able to reject parameter perturbations or disturbances. Therefore, in the second part of this section, a hybrid control structure will be proposed where a parallel PI controller is proposed to force the system output to the desired set point. The PI controller will only be activated if there is model mismatch.



BB-BC Optimization based Inverse Controller

Figure 1: BB-BC based Inverse Open Loop Control Scheme

3.1. BB-BC based Inverse Model Controller

In this proposed control structure, the control signal generation is handled as an optimization problem. The problem can be defined as to decrease the error between the process model output and the reference signal.

Assuming that the process model matches the process perfectly, as the process model output converges to the set point; the process output will converge to the set point, too. Since an online implementation of this evolutionary algorithm is feasible, at each sampling the optimal control signal can be generated. Then the optimal control signal is then applied to the process. The scheme is illustrated in Figure 1.

The cost function J_1 which is minimized at every sampling time is chosen as:

$$J_1 = \left(ref - y_m\right)^2 \tag{5}$$

Since BB-BC is a stochastic evolutionary algorithm, the performance of the controller will vary for each trial.

3.2. Hybrid Controller Structure

As it has been mentioned before, the BBBC inverse control structure will not be able to reject disturbances or parameter perturbations. Therefore a hybrid control structure is proposed. This structure is similar to the well known Internal Model Scheme; however the error is only used to activate the PI controller. The PI controller will be activated in the case of a model mismatch or disturbance. So it will be guarantied that the system output will converge to the set point. The scheme is illustrated in Figure 2.

4. PH PROCESS MODELING

4.1. Process chemistry

The pH process is characterized by the presence of acid/base reactions. In this study the reaction of strong acid (HCl) and strong base (NaOH) reaction will be considered. It is a known fact that this type of reactions occur quickly. In the reaction tank the decomposition of HCl, NaOH and H_2O is represented with the following equations (McAvoy et al. 1972).

$$\begin{aligned} &H_2O \leftrightarrow H^+ + OH^- \\ &HCl \rightarrow H^+ + Cl^- \\ &NaOH \rightarrow Na^+ + OH^- \end{aligned} \tag{6}$$

The electro neutrality principle indicates that the concentration of positive and negative ions in a solution has to be neutral in the equilibrium. So, the following must be satisfied;

$$[H^+] + [Na^+] = [OH^-] + [Cl^-] \tag{7}$$

Beside this equation, the equilibrium of the pure water must also be satisfied which is given as:

$$k_w = [H^+][OH^-] = 10^{-14}\, mol^2 / L^2 \tag{8}$$

where k_w is the dissociation constant of the water.

Defining,

$$\begin{aligned} w_a &= [Cl^-] \\ w_b &= [Na^+] \end{aligned} \tag{9}$$

Replacing in the equation (7),

$$w_b + [H^+] = w_a + \frac{k_w}{[H^+]} \tag{10}$$

After rearranging the equation, the following second-order polynomial is found

$$[H^+]^2 + [H^+](w_b - w_a) - k_w = 0 \tag{11}$$

So the morality of H^+ is found as:

$$[H^+] = \frac{(w_a - w_b) + \sqrt{(w_a - w_b)^2 + 4k_w}}{2} \tag{12}$$

The morality of H^+ is directly related with pH value which indicates the acidic value of the process.

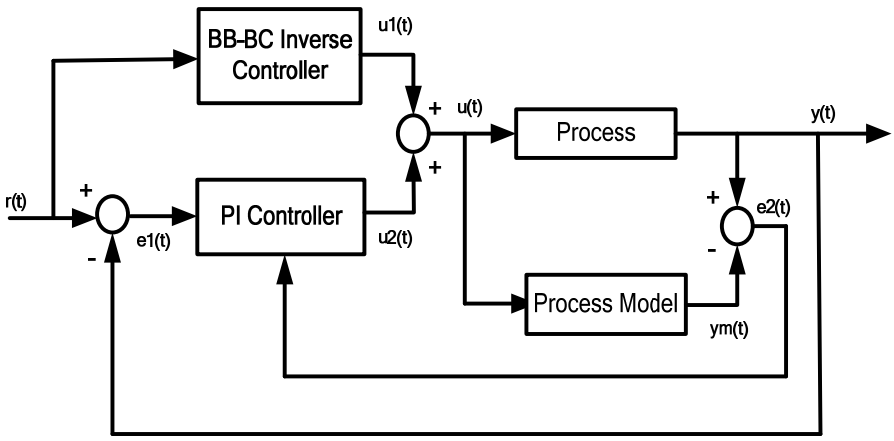


Figure 2: Proposed Hybrid Control Scheme

The relationship is defined as:

$$pH = -\log\left([H^+]\right) \tag{13}$$

Replacing the solution from equation (12), we obtain

$$pH = -\log\left(\frac{(w_a - w_b) + \sqrt{(w_a - w_b)^2 + 4k_w}}{2}\right) \tag{14}$$

Via this equation the pH value of the mixture of strong acid and base can be calculated. It can be clearly seen from Figure 3 that there is static nonlinear relationship between pH and $[H^+]$.

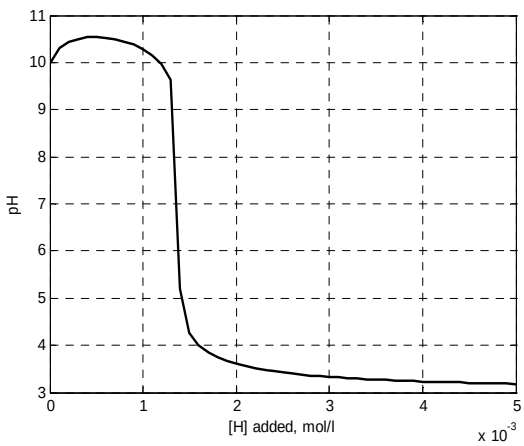


Figure 3: pH Variation

4.2. Process dynamics

The dynamic model of the pH process is proposed by (McAvoy et al. 1972) and is shown in Figure 4. Assumption of perfect mixing is general in the modelling of pH processes. Balances in the reactor can be given by

$$\left.\begin{aligned} V\frac{dw_a}{dt} &= F_a C_a - (F_a + F_b) w_a \\ V\frac{dw_b}{dt} &= F_b C_b - (F_a + F_b) w_b \end{aligned}\right\} \tag{15}$$

where V is the volume of the mixture tank, C_a and C_b are the acidic and basic concentration, F_a and F_b are the acidic and basic flow rates respectively.

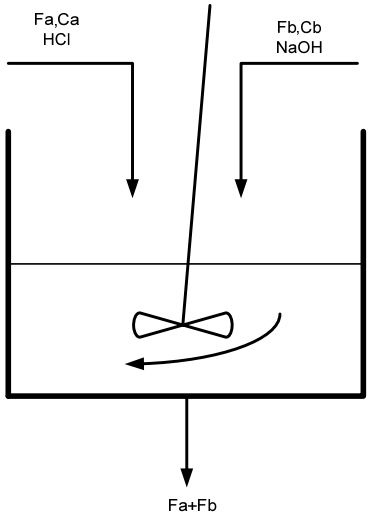


Figure 4: Schematic of CSTR

5. SIMULATION STUDY

In this section a simulation study will be performed in order to show the effectiveness of the proposed control structure. As it has been explained briefly in the previous section, the pH neutralization process inherits not only a static nonlinear gain but also the dynamics of CSTR are also nonlinear. Simulations are performed on MATLAB®/Simulink toolbox.

The description of the pH process used in this simulation study is given in Table 2.

The initial condition of the CSTR is assumed to be that the morality of H^+ and OH^- are equal, therefore the initial pH value is 7.

Table 2: Description of the pH process

Symbols	Description	Value
V	Volume of the CSTR	0.8L
F_a	Flow rate of the influent stream	1 l/h
F_b	Flow rate of the titrating stream	0-2.1 l/h
C_a	Concentration of the influent stream	0.001 mol/l
C_b	Concentration of the titrating stream	0.001 mol/l

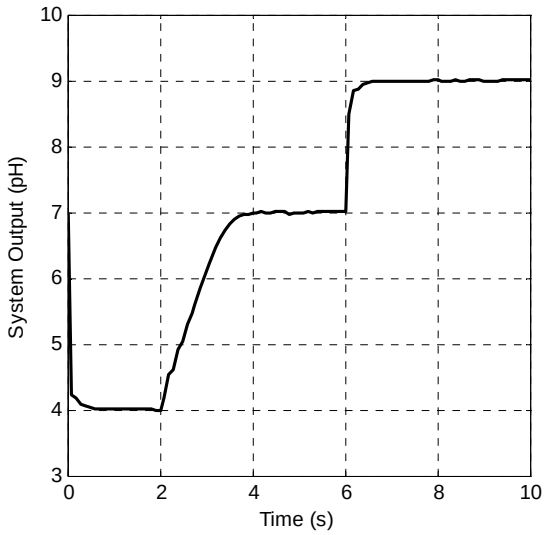


Figure 5: System Output for Varying Reference Signal

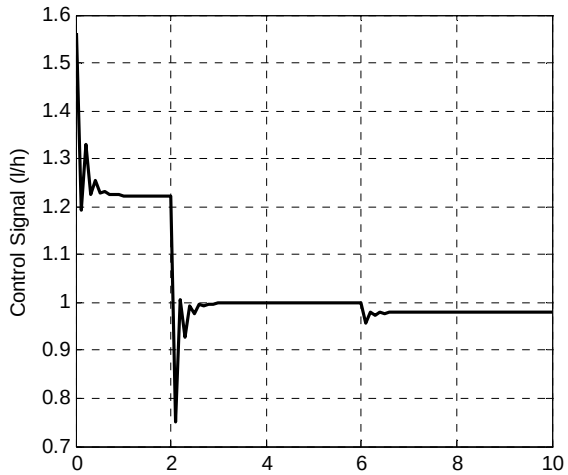


Figure 6: Control Signal for Varying Reference Signal

In the first part, since the dominant nonlinearity is related to the morality of H^+, the controller has been tested under varying pH reference values. It can be seen from Figure 5 that the controller provides satisfactory performances for different pH reference signals. Especially, it is observed that, the process output has been the set point value 7 pH successfully. In this case, since there is nor parameter perturbation nor disturbance, the PI controller is not activated. The process is forced the set value, only by BB-BC inverse controller. The control sign is also presented in Figure 6.

In the second part, at first the neutralization process has been forced to the set point reference signal 7 pH. Then in order to examine the disturbance case, the flow rate of the influent stream has been increased %10 percent in the 5th second. As soon as the disturbance affects the process, the PI controller is activated and disturbance has been compensated in a short period of time. The system output and the control signal are presented in Figure 7 and Figure 8, respectively.

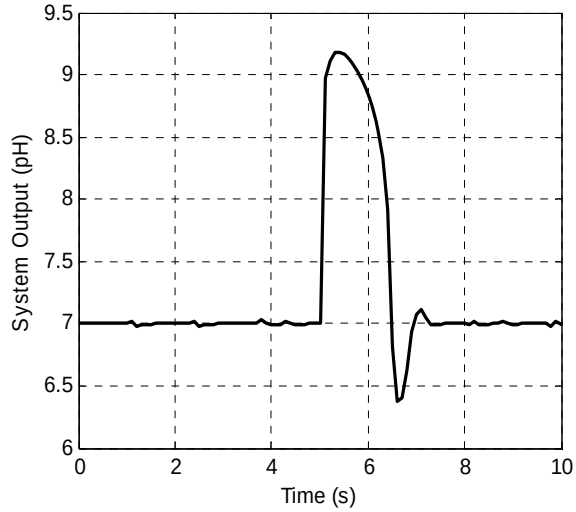


Figure 7: System Output Under Disturbance

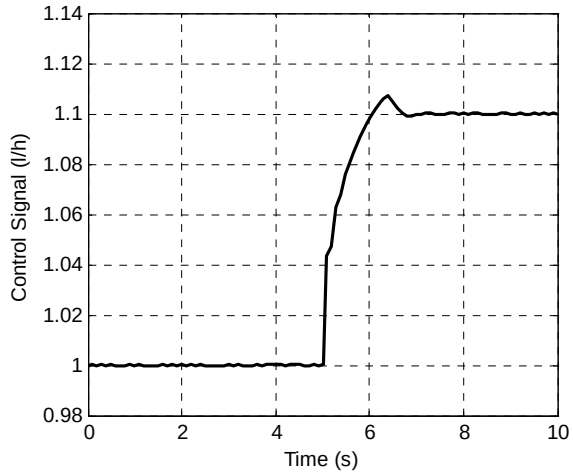


Figure 8: Control Signal Under Disturbance

6. CONCLUSIONS

In this study, a new iterative inversion technique based on the BB-BC optimization has been presented. In this new technique, the inverse model control signal generation is handled as an optimization problem. Since the BB-BC optimization algorithm has a high convergence speed and low computational time, the optimal inverse model control signal can be generated

within each sampling time in an on line fashion. In order to show performance of the new approach a simulation study has been performed. In this study, the pH process which has high sensitivity at and near the neutralization point has been controlled. It has been shown that pH process could have been controlled perfectly in an open loop control scheme. Since the proposed methodology is based on a stochastic global search algorithm, the generated inverse model control signals may vary at each trial and this is the only arguable point of this inversion technique. Moreover, a hybrid control structure is proposed where a parallel PI controller is implemented to the BBBC inverse model controller. In the case of various disturbances and parameter perturbations, PI controller will get activated so that the convergence of the system output to the reference signal is guaranteed. It has been demonstrated via simulation examples that the proposed controller scheme really provides quite satisfactory results.

ACKNOWLEDGMENTS
This research is supported by the project (108E047) of Scientific and Technological Research Council of Turkey. All of these supports are appreciated.

REFERENCES

Babuska, R., Oosterhoff, J., Oudshoorn, A., Bruijn, P.M., 2002. Fuzzy self-tuning PI control of pH in fermentation. *Engineering Applications of artificial Intelligence* 15, 3–15.

Erol, O. K. and Eksin I., A new optimization method: Big Bang–Big Crunch *Advances in Engineering Software 37* (2006) 106–111

Fuente, M.J., Robles, C., Casado O., Syafiie, S., Tadeo, F.,2006, Fuzzy control of a neutralization process, *Engineering Applications of Artificial Intelligence*, 19 905–914

Henson ,M. A., Seborg, D. E., 1994, Adaptive non linear control of a pH neutralization Process, *IEEE Transactions On Control Systems Technology*, 2(3),169–184.

Kumbasar T., Eksin İ., Güzelkaya M., Yesil E., 2008. Big Bang Big Crunch Optimization Method Based Fuzzy Model Inversion, *MICAI 2008: Advances in Artificial Intelligence - Lecture Notes in Computer Science*, 5317, 732-740.

Kumbasar, T., Yesil, E., Eksin, İ., Güzelkaya, M., 2008. Inverse Fuzzy Model Control with Online Adaptation via Big Bang-Big Crunch Optimization, *The 3rd International, Symposium on Communications, Control and Signal Processing, ISCCSP*, Malta.

Loh, A, Sankar De, D., Krishnaswamy, P. R., 2001, pH and Level Controller for a pH Neutralization *Process, Ind. Eng. Chem. Res*., 40, 3579-3584

McAvoy, T.J., Hsu, E., Lowenthal, S., 1972, Dynamics of pH in controlled stirred tank reactor, *Ind. Eng. Chem. Process Dev.*11 68–70.

Nystrom, R.H., Sandstrom, K.V., Gustafsson, T.K., Toivonen, T.H.,1998. Multimodel robust control applied to a pH neutralization process. *Computer Chemical Engineering* 22, s467–s474.

Pishvaie, M. R., Shahrokhi, M, 2006, Control of pH processes using fuzzy modeling of titration curve, *Fuzzy Sets and Systems* 157 2983 – 3006.

Wright, R.A. , Kravaris, C. ,1991, Nonlinear control of pH processes using the strong acid equivalent, *Indust. Eng. Chem. Res.* 30 (7) 1561–1572.

Yoon, S.S., Yoon, T.W., Yang, D.R., Kang, T.S., 2002. Indirect adaptive nonlinear control of a pH process. *Computers and Chemical Engineering* 26, 1223–1230.

TUFAN KUMBASAR received B.Sc. degree from the Department of Control Engineering of the Faculty of Electrical and Electronics Engineering of Istanbul Technical University, Istanbul, Turkey, in 2006 and M.S. degree from the Institute of Science and Technology of the same university in 2008. He is currently a Ph.D. student and is working as a research assistant in the Department of Control Engineering of Istanbul Technical University. His research interests are intelligent control, optimization, and process control.

IBRAHIM EKSIN received the B.Sc., M.S. and Ph.D. degrees from the Department of Electrical Engineering of Bogazici University, Istanbul, Turkey, in 1976, 1979 and 1983, respectively. He has worked in Honeywell from 1983 to 1985. He then joined the Department of Control Engineering of Istanbul Technical University and is currently working as a professor in the same division. He is the General Secretariat of NMO of IFAC. His research interests are intelligent control, optimal control and evolutionary computation.

MUJDE GUZELKAYA received the B.Sc. degree from the Department of Electronics and Communication Engineering of the Faculty of Electrical Engineering of Istanbul Technical University, Istanbul, Turkey, in 1982 and M.S. and Ph.D. degrees from the Institute of Science and Technology of the same university in 1984 and 1990, respectively. She is currently a professor in Department of Control Engineering of Istanbul Technical University. Her research interests are intelligent control, multivariable control and stochastic optimal control.

ENGIN YESIL received B.Sc. degree from the Department of Electrical Engineering of the Faculty of Electrical and Electronics Engineering of Istanbul Technical University, Istanbul, Turkey, in 1998 and M.S. and Ph.D. degrees from the Institute of Science and Technology of the same university in 2000 and 2008, respectively. He is working as a research assistant in the Control Engineering Department of Istanbul Technical University. His research interests are intelligent control, process control.

FUZZY BLENDING HYBRID STRUCTURE WITH FUZZY AND CONVENTIONAL PID CONTROLLERS

Cihan Bulut [(a)], Ibrahim Eksin[(b)], Mujde Guzelkaya[(c)], Tufan Kumbasar[(d)], Engin Yesil[(e)]

Istanbul Technical University, Faculty of Electrical and Electronics Engineering,
Control Engineering Department, Maslak, TR-34469, Istanbul, Turkey

[(a)]bulutc@itu.edu.tr, [(b)]eksin@itu.edu.tr, [(c)]guzelkaya@itu.edu.tr, [(d)]kumbasart@itu.edu.tr, [(e)]yesileng@itu.edu.tr

ABSTRACT

The aim of this study is to develop a fuzzy blending hybrid controller (FBHC) which mixes the control outputs of a conventional PID and a fuzzy PID controller. The idea behind this design methodology is to combine the beneficial sides of both controllers in its own structure. The fuzzy and the conventional controllers are put into parallel form within this the blending mechanism and generally the advantages of conventional controller in steady-state characteristics and the fuzzy controller in transient characteristics are exploited. In this paper, a new hybrid controller scheme with a blending mechanism that uses simple fuzzy rule base instead of complicated algorithms has been presented. Moreover, the proposed blending mechanism is independent of the type of controllers used in hybridization. Thus, this feature provides the designer an opportunity to use other control strategies within the same mechanism for different processes.

Keywords: fuzzy logic, fuzzy PID + conventional PID, hybrid control, fuzzy blending

1. INTRODUCTION

Despite a lot of research and different effective solutions, conventional PID controllers (proportional-integral-derivative) are the most popular controllers used in industry due to their simplicity and cost affectivity. According to the different sources, the use of conventional PID controllers in industry is in between 90% and 99% (Reznik et. al. 2000). When the system to be controlled is linear the performance of PID controllers is superb, but if the system is a nonlinear or certain uncertainties exist within system, PID controllers cannot achieve a good performance (Er and Sun 2001).

On the other hand, fuzzy controllers are another type of controller and they are widely and increasingly been used by control engineers for too many systems with nonlinearity and uncertainty over the past two decades (Sugeno 1985; Driankov 1996; deSilva 1995). The main advantage of this method is that there is usually no need for a model in designing the fuzzy controller (Passino and Yurkovich 2001). However, defining fuzzy rules and designing the membership functions may unfortunately be time consuming.

These drawbacks and advantages remind a hybrid structure which involves both a linguistic part and a numeric part in its topology. FBHC integrates the advantages of both conventional PID controller and fuzzy controller. This idea naturally interested many engineers. Various hybrid controller designs have been arisen in literature (Xiaoyin and Belmin 1993; Kwok et. al. 1990; Brehm and Rattan 1993).

A self optimal regulating factor is added to the control rule of the fuzzy controller in order to have not only quick dynamic response, but also high steady-state accuracy of a PID (Liang and Qu 1993). An interesting approach with the parallel connection of those two controllers is to use both control outputs in some combination (Li 1998). In that method, a fuzzy P and an ID controller were used in the hybrid controller structure.

In this paper, a new approach toward designing a hybrid controller using a fuzzy blending mechanism has been presented. This is a way to design effective combinations of conventional PID controllers and intelligent methodologies for the industry. The proposed controller FBHC compares the controller outputs within a fuzzy mechanism and that mechanism produces a blending factor. Then, the controller outputs are mixed up appropriately using this factor. Therefore, FBHC can be considered as a mechanism which tries to determine and use the controller output that gives the best system response more effectively. The leading advantage of this blending mechanism is the fact that it is independent from the nature of the controllers used. FBHC structure can easily be applied to both linear and nonlinear systems. In this study, the results of FBHC are compared with the results obtained using both the pure conventional and fuzzy PID controllers and it has been proven that the proposed hybrid controller outperforms the pure forms of both controllers both in transient and steady-state even under disturbances.

Performance comparison between the proposed hybrid controller and the pure of the controllers involved has been carried out by two simulation examples that confirm the superiority of the hybrid controller. The structure of the hybrid controller and the pure of the controllers are presented in section 2, and blending mechanism is described in section 3. The simulation examples are given in section 4. The last part is the conclusion stage.

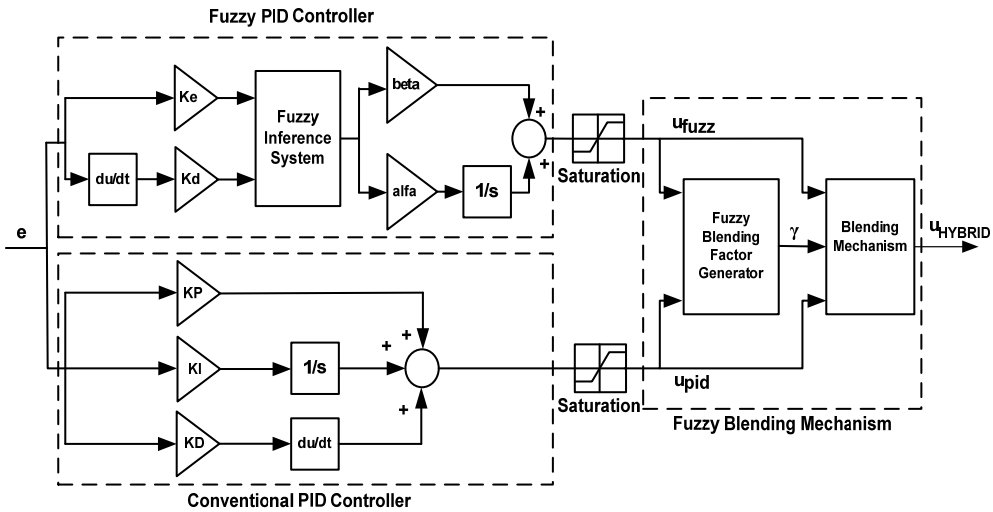


Figure 1: Structure of Fuzzy Blending Mechanism

2. HYBRID FUZZY PID + CONVENTIONAL PID CONTROLLER SCHEME

The proposed structure of the FBHC with a Fuzzy PID and a conventional PID is shown in Figure 1. The first part of this hybrid structure is a Fuzzy PID controller and the inputs of this controller are the system error (e) and the rate of the change of the system error ($\dot{e}$). These inputs are defined as %50 percent overlapped triangular membership functions in the range of [-1, 1], while the output (u) is defined with singleton membership functions as shown in Figure 2. The rule base of the fuzzy controller is composed of 49 rules as given in Table 1.

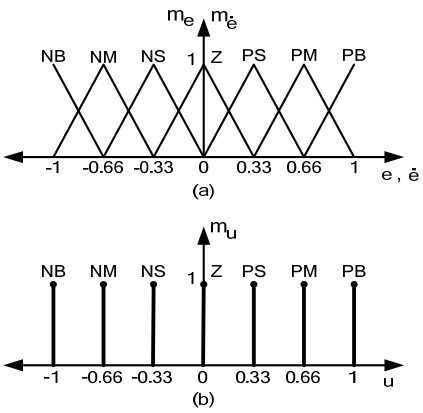


Figure 2: (a) Error, and Derivative of Error, (b) Control Signal Membership Functions of the Fuzzy PID Controller

Table 1: Fuzzy PID Controller Rule Base

$\Delta e / e$	**NL**	**NM**	**NS**	**Z**	**PS**	**PM**	**PL**
PL	Z	PS	PM	PL	PL	PL	PL
PM	NS	Z	PS	PM	PL	PL	PL
PS	NM	NS	Z	PS	PM	PL	PL
Z	NL	NM	NS	Z	PS	PM	PL
NS	NL	NL	NM	NS	Z	PS	PM
NM	NL	NL	NL	NM	NS	Z	PS
NL	NL	NL	NL	NL	NM	NS	Z

The second part of this hybrid control structure is a conventional PID controller and its transfer function can be given as:

$$G_{PID}(s) = K_p + \frac{K_I}{s} + K_D s \tag{1}$$

where K_P is the proportional gain, K_I the integral gain, K_D the derivative gain.

Either one of the two controllers might have been chosen with aggressive response; that is, small rise time and high overshoot and the other one having a smooth system time response; that is, high settling time and low or no overshoot. Then, this hybrid mechanism will provide a system response exploiting the beneficial sides of both controllers.

3. FUZZY BLENDING MECHANISM (FBM)

FBM is a structure where the control signals of the two control signals are mixed. Different blending algorithms can be suggested but here a method based on fuzzy logic is proposed.

In literature, there also exist other ideas in calculating the blending factor γ; for instance, a simple function depending on the system error e can be used (Erenoglu et. al. 2006). However, the algorithm could have become more complex in order to cover all the system situations. Here, as an alternative, it has shown that a simple fuzzy rule base can be used instead of complicated algorithms. The blending rules used in this study are not complicated ones and they are defined to control the process over a wide range of operating points.

The outputs of the Fuzzy PID controller and the conventional PID controller are multiplied by either the output blending factor γ or (1-γ). The key point of the blending mechanism is to get a reasonable tradeoff between the pure forms of the two controllers. FBM can be given in two main parts; namely, Fuzzy Blending Factor Generator (FBFG) and Blending Mechanism (BM).

3.1. Fuzzy Blending Factor Generator (FBFG)

FBFG is the part of the FBM where the blending factor γ is produced.

3.1.1. Membership Functions

Triangular-shaped functions shown in Figure 3 are chosen as the membership functions due to the resulting simplicity. The fuzzy members for the input are defined as Very Large (VL), Large (L), Medium (M), Small (S) and Very Small (VS).

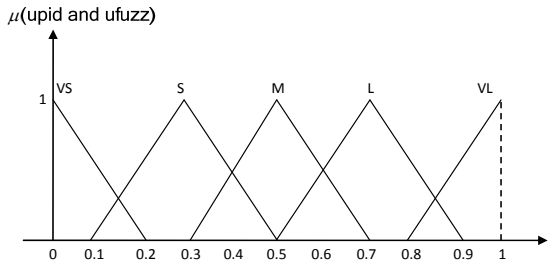


Figure 3: Membership Functions of Control Inputs

The output variable γ is also triangular membership functions are defined as illustrated in Figure 4 the linguistic labels for the memberships are given as Very Very Large (VVL), Very Large (VL), Large (L), Medium (M), Small (S), Very Small (VS) and Very Very Small (VVS).

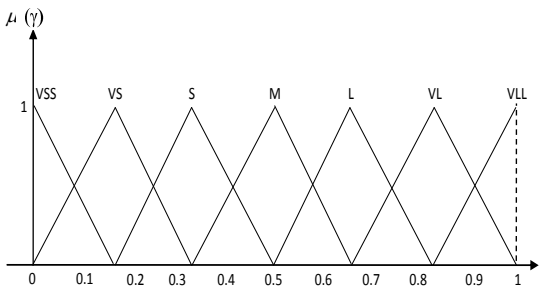


Figure 4: Membership Function of Blending Factor

3.1.2. Rule Base

FBFG rule base is composed of 25 rules as given in Table 2. The number of rules could have been increased; but it has been avoided for simplicity. The rules are between the two controller outputs: Fuzzy PID and Conventional PID and three or more controller outputs might have been used.

It is obvious from the rules that when a controller output is aggressive compared to the other one, the aggressive output multiplied with a higher blending factor and this controller output activates more than the other controller in the hybrid control output. In other words, the controller output which gives the faster response must be multiplied by a greater value of γ in the transient phase of system response. The main idea of this structure is to multiply the dominant control signal with higher fulfillment degree to produce faster system response.

Table 2: FBFG Rule Base

u_{pid} \ u_{fuzz}	**VL**	**L**	**M**	**S**	**VS**
VL	M	L	VL	VLL	VLL
L	S	M	L	VL	VLL
M	VS	S	M	L	VL
S	VSS	VS	S	M	L
VS	VSS	VSS	VS	S	M

The rule base is designed on this basic idea. When Fuzzy PID output (u_{fuzz}) is Very Large (VL) and Conventional PID output (u_{pid}) is Very Small (VS), the bigger blending factor must affect the control output so a Very Very Large (VVL) factor must be multiplied with u_{fuzz}. This means that a Very Very Small factor (1-γ) will be multiplied with u_{pid}.

On the other hand when u_{pid} is VL and u_{fuzz} is VS, the blending factor will be VSS, so the conventional PID ratio will be close to 1 in the hybrid control output. On the other hand, the fuzzy PID fulfillment in the hybrid output will be close to 0.

When u_{fuzz} is Medium (M) and u_{pid} is Medium (M), blending factor is Medium (M) for an equal blending. The surface of the FBFG is illustrated in Figure 5.

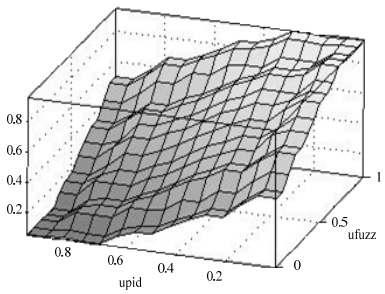


Figure 5: The Surface of FBFG

3.1.3. Scaling Factor

As it can be seen from that Figure 1, after both of the controller outputs, saturation blocks are used. This saturation blocks gives the mechanism the ability to work with the same rule base in different systems and controllers. After the saturation blocks, the controller outputs are mapped in [0,1] region.

By using that kind of saturations and mapping procedure, there is no need to search for the scaling factors. Moreover, saturation is more physical than a searching algorithm.

3.2. Blending Mechanism (BM)

BM has three inputs: Conventional PID controller output u_{pid}, Fuzzy PID controller u_{fuzz} and FBFG output γ. Here the new controller output is calculated as;

$$u_{HYBRID} = u_{fuzz}.\ \gamma\ +\ u_{pid}.(1-\gamma) \qquad (2)$$

where u_{HYBRID} is the new control output.

It is so apparent from the rule base that when a controller output is larger from the other that controller output is multiplied a bigger blending factor so it is

activated more than the other controller part. FBHC tries to catch the bigger one of the control efforts of the two controllers. Behind this lays the idea of that the higher control effort should produce faster system response.

4. SIMULATION STUDIES

In order to show the benefit of the proposed control structure two simulation examples are presented. For each example, the transient response for the reference changes, the input ($d_1(t)$) and the output disturbance ($d_2(t)$) rejection performance of the proposed hybrid control (FBHC) is compared with a fuzzy PID and conventional PID controllers. The control scheme used for the simulations is presented in Figure 6.

Simulations are performed on MATLAB®/Simulink toolbox to illustrate the efficiency of the FBHC.

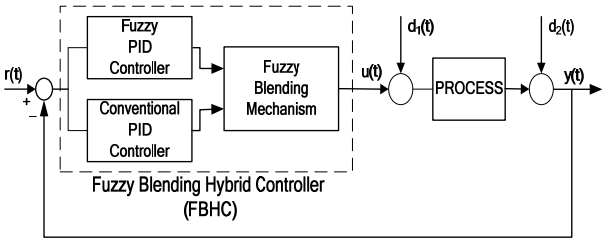


Figure 6: Control Scheme of the Proposed Hybrid Controller

4.1. Linear System

Most of the systems in industry can be modeled as second-order with time-delay systems, therefore the following system has been considered.

$$G(s) = \frac{1}{s^2 + 3s + 1} e^{-0.2s} \tag{3}$$

The controller parameters of the conventional PID controller are designed so as to give a system response with no overshoot and the parameters are as follows: KP=0.92, KI=0.73, KD=0.11. On the other hand, the fuzzy PID parameters have been designed so as to provide a fast rising time as follows: Ke=1.89, Kd=1.35, alfa=1.1, beta=0.1.

A unit step reference is applied in the beginning, and then at 18th second, the reference is changed from one to 1.5. In addition, an output ($d_2(t)$) and an input ($d_1(t)$) disturbance with amplitude of 0.2 units are applied to the system at 30th second and at 50th, respectively.

The corresponding system responses and controller outputs are given in Figure 7 and Figure 8, respectively. It can be seen from Figure 7 that the proposed controller provides satisfactory performances for different reference signals and disturbances. The system response of the FBHC is always between responses of other controllers as it is expected from the blending mechanism.

The FBHC control signal is between the Fuzzy PID and Conventional PID control signals, since it is blended with different values of γ. When γ is close to 1, FBHC controller performs like Fuzzy PID controller. On the other hand, when the value of γ becomes 0 the output of the proposed controller is equal to the Conventional PID.

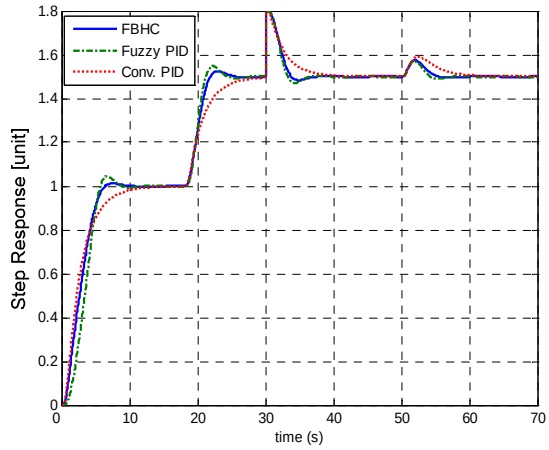


Figure 7: The System Output For Varying Reference Values And Under Disturbances.

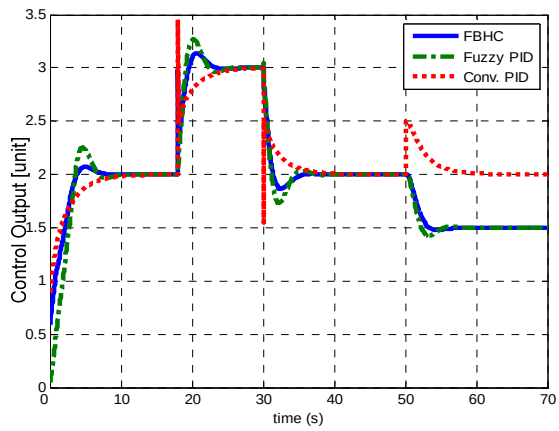


Figure 8: The Control Signal For Varying Reference Values And Under Disturbances.

4.2. Nonlinear System

In the nonlinear simulation study, the proposed controller will be used for a nonlinear spherical tank process as presented in Figure 9. The Simulation results confirm the better performance of the proposed hybrid controller.

A spherical tank system is a nonlinear level control system (Agrawal and Lakshminarayan 2003). The parameters of system are given in Table 3.

The differential equation can be gives as;

$$Q_i(t-d) - Q_o = (\pi - \pi(R-y)^2)\frac{dy}{dt} \tag{4}$$

where R is the radius of tank, Q_i is the inlet flow rate (volumetric), and Q_o is the outlet flow rate (volumetric). d is the delay from input Q_i to the controlled output y. With Bernoulli equation;

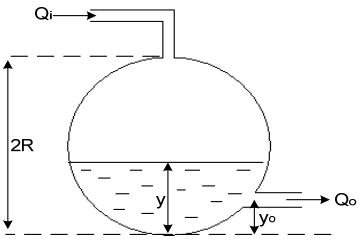


Figure 9: Spherical Tank System

$$Q_o = \sqrt{2g(y - y_o)} \quad (5)$$

where y_o is the height of output pipe and g represents the gravitation constant.

Table 3: Spherical tank parameters

Radius of tank [m]	R = 1m
Delay from Q_i to y	d = 0 s
Gravity acceleration [m/s^2]	g = 9.81
Height of the output pipe [m]	y_o=0.1m
Inlet volumetric flow rate [m^3/s]	$Q_i(t)$
Outlet volumetric flow rate [m^3/s]	$Q_o(t)$
Height of liquid level [m]	y

The controller parameters of the conventional PID controller are designed so as to give a system response with no overshoot and the parameters are as follows: K_P=0.72, KI=2.51, KD=0.051. On the other hand, the Fuzzy PID controller parameters have been designed so as to provide a fast rising time as follows: Ke=4.62, Kd=1.61, alfa=6.021, beta=1.523.

It is assumed that the initial value of the tank height is 0.1m. Since the nonlinearity is directly related to the level of the water, the controller has been tested for different reference values as 0.4 m and 1 m. After the process output is converged to the set point, at 25th second an input disturbance with a value of 0.1 m, and an output disturbance with a value of 0.1 m at 35th second are applied in order to examine the disturbance rejection of the control structure.

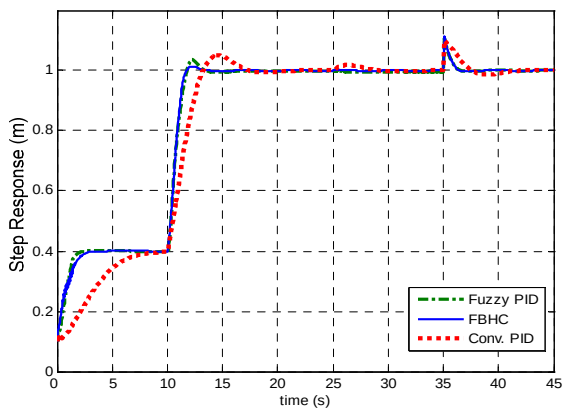


Figure 10: The System Output For Varying Reference Values and Under Disturbances.

The corresponding system responses and controller outputs are given in Figure 10 and Figure 11, respectively. It can be seen from Figure 10 that the proposed controller provides satisfactory performances for different reference signals and disturbances.

The FBHC control signal is between the Fuzzy PID and Conventional PID control signal, since it is blended with different values of γ. When γ is close to 1, FBHC control output is like Fuzzy PID output.

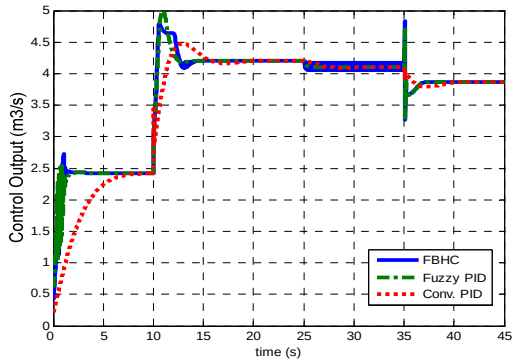


Figure 11: The Control Signal For Varying Reference Values And Under Disturbances.

5. CONCLUSIONS

A new method toward designing a Fuzzy PID and Conventional PID type hybrid controller using a fuzzy blending mechanism has been presented in this paper. Conventional PID controllers are very popular controllers in industry due to their simplicity and cost affectivity. With this hybridization methodology, intelligent technology can be inserted more easily into the real-life industrial projects. Here, a fuzzy blending mechanism, which hybridized two well known controllers, is designed so that it produces a remedy for most of the system situations such as reference changes, different types of disturbances. The main idea of this design is to use the dominant control signal with higher fulfillment degree to produce faster system response for rising time and then blending with the other control output in order not to have a major overshoot.

The proposed controller has been applied to both linear and nonlinear systems. Performance comparison between the presented controller and the controllers involved has been carried out by a system simulation results all confirm the advantage of the presented controller.

The fuzzy blending mechanism is independent of the controllers that have been hybridized; so for the future work, other control methods can be used to have a different type of hybrid controllers for different processes.

REFERENCES

Agrawal, P. Lakshminarayan, S., 2003. Tuning proportional-integral-derivative controllers using achievable performance indices, *Ind. End. Chem. Res*, 42, pp. 5576–5582.

Brehm, T., Rattan, K.S., 1993. Hybrid fuzzy logic PID controller. *Proceeding of the IEEE National Aerospace and Electronics Conference*, 2, pp. 807-813.

C.W. deSilva, 1995. Intelligent Control and Fuzzy Logic Applications. *Boca Raton, FL: CRC Press*.

Driankov, D., Hellendoom, Reinfrank, M., 1996. An Introduction to Fuzzy Control, *2nd ed. New York: Springer-Verlag.*

Er, M.J., Sun, Y.L., 2001. Hybrid Fuzzy Proportional–Integral Plus Conventional Derivative Control of Linear and Nonlinear Systems. *IEEE Transactions on Industrial Electronics*, 48 (6), pp. 1109-1117.

Erenoglu, I., Eksin, I., Guzelkaya, M., Yesil, E., 2006. An Intelligent Hybrid Fuzzy PID Controller. ECMS'06, 20th European Conference on Modelling and Simulation, Bonn, Germany.

Ketata, R., Geest, D.D., Titli, A., 1995. Fuzzy controller: design, evaluation, parallel and hierarchical combination with a PID controller. *Fuzzy Sets and Systems,* 71, pp. 113-129.

Kwok, D.P., Tam, P., Li, C.K., Wang, P., 1990. Linguistic PID controllers. *Proceedings of the 11th World Congress, Tallin, Estonia*, 7, pp. 192-197.

Lee, J., 1993. A method for improving performance of PI-type fuzzy logic Linguistic PID controllers. *IEEE Transactions on Fuzzy Systems*, 1, pp. 298-301.

Li,W., 1998. Design of a hybrid fuzzy logic proportional plus conventional integral-derivative controller. *IEEE Transactions on Fuzzy Systems*, 6 (4), pp. 449-463.

Li, W., 1997. A method for design of a hybrid neuro-fuzzy control system based on behavior modelling. *IEEE Tranactions on Fuzzy Systems,* 5, pp. 128-137.

Liang, X., Qu, B., 1993, Fuzzy- PID controller. *Proceedings of the TENCON`93, 1993 IEEE Region 10 Conference*, 4, pp. 296-299.

Matsunaga, N., Kawaji, S., 1991. Fuzzy hybrid control for DC servomotor. *Trans. Inst. Electrical Eng. Japan D*, 111, pp. 195-200.

Parnichkun, M., Ngaecharoenkul, C., 2001. Kinematics control of pneumatic systems by hybrid fuzzy PID. *Mechatronics*, 11, pp. 1001-1023.

Passino, K. M., Yurkovich S., 2001. Fuzzy Control. *Tsinghua University Press, ADDSION-WESLEY. Beijing, China.*

Reznik, L., Ghanayem, O., Bourmistrov, A., 2000, PID plus fuzzy controller structure as a design base for industrial applications. *Engineering Applications of Artificial Intelligence*, 13 , pp. 419-430.

Sugeno, M., 1985, An introductory survey of fuzzy control. *Inform. Sci.*, 36, pp. 59–83.

White, A.S. and Sinclair, R., 2004, Quantitative validation techniques a database. (I) Simple examples, *Simulation Modeling Practice and Theory*, 12, pp. 451-473.

Xiaoyin, L., Belmin, Q., 1993, Fuzzy-PID controller. *IEE TENCON'93, Beijing*, pp. 296-299.

Zhong, H.S., Shaohua, T., Zhuang, W.P., 1993. Fuzzy self-tuning of PID controllers. *Fuzzy Sets and Systems*, 56, pp. 36-46.

CIHAN BULUT received B.Sc. degree from the Department of Control Engineering of the Faculty of Electrical and Electronics Engineering of Istanbul Technical University, Istanbul, Turkey, in 2006 and M.S. degree from the Institute of Science and Technology of the same university in 2009. He is currently working as an automation engineer. His research interests are intelligent control and process control and industrial automation.

IBRAHIM EKSIN received the B.Sc., M.S. and Ph.D. degrees from the Department of Electrical Engineering of Bogazici University, Istanbul, Turkey, in 1976, 1979 and 1983, respectively. He has worked in Honeywell from 1983 to 1985. He then joined the Department of Control Engineering of Istanbul Technical University and is currently working as a professor in the same division. He is the General Secretariat of NMO of IFAC. His research interests are intelligent control, optimal control and evolutionary computation.

MUJDE GUZELKAYA received the B.Sc. degree from the Department of Electronics and Communication Engineering of the Faculty of Electrical Engineering of Istanbul Technical University, Istanbul, Turkey, in 1982 and M.S. and Ph.D. degrees from the Institute of Science and Technology of the same university in 1984 and 1990, respectively. She is currently a professor in Department of Control Engineering of Istanbul Technical University. Her research interests are intelligent control, multivariable control and stochastic optimal control.

TUFAN KUMBASAR received B.Sc. degree from the Department of Control Engineering of the Faculty of Electrical and Electronics Engineering of Istanbul Technical University, Istanbul, Turkey, in 2006 and M.S. degree from the Institute of Science and Technology of the same university in 2008. He is currently a Ph.D. student and is working as a research assistant in the Department of Control Engineering of Istanbul Technical University. His research interests are intelligent control, optimization, and process control.

ENGIN YESIL received B.Sc. degree from the Department of Electrical Engineering of the Faculty of Electrical and Electronics Engineering of Istanbul Technical University, Istanbul, Turkey, in 1998 and M.S. and Ph.D. degrees from the Institute of Science and Technology of the same university in 2000 and 2008, respectively. He is working as a research assistant in the Control Engineering Department of Istanbul Technical University. His research interests are intelligent control, process control.

MODELING A SPATIAL COMMUNICATION ACTIVITY IN WIRELESS SENSORS NETWORK

Jan Nikodem [(a)], Ryszard Klempous[(a)], Maciej Nikodem[(a)], Marek Woda[(a)], Zenon Chaczko[(b)]

[(a)] Wroclaw University of Technology, Poland
[(b)] University of Technology Sydney, NSW, Australia

[(a)]{jan.nikodem, ryszard.klempous, maciej.nikodem, marek.woda}@pwr.wroc.pl, [(b)]zenon.chaczko@uts.edu.au

ABSTRACT
This paper presents a novel method of modeling spatial communication activity in wireless sensor network (WSN). We define native aspects of communication in WSN. Focusing on local/global activity dilemma, cooperation, interference, network topology, and optimization aspects. A neighborhood abstraction is defined and we involve three binary relations: subordination, tolerance and collision to describe the cooperation in WSN. Using digital terrain model tools we model communication activity aspects as surfaces, stretched over WSN network. A network topology features are modeled using bare drainage surface. It is a component of a topographic map, which gives a direction towards the base station, determined by a slope of the modeled surface. Modeling node's instant energy level, we construct another surface represents node's instant level of consumed energy. Finally, we construct a drainage surface spread over each node neighborhood as superposition of bare drainage surface, energy consumed and relational surfaces.

Keywords: wireless sensors network, spatial communication, relations in complex system

1. INTRODUCTION

Regular node's measure parameters of the environment they reside. Their basic task is to measure, collect and to send a data to the base station (BS). Wireless Sensor Networks have been studied for a long time and there are plenty of publications in this subject focusing on different aspects of network operation (Vaidya 2005; Cohn 1997; Braginsky 2002). Multiplicity of issues and topics leads to restrictions and assumptions that aim at simplifying the analysis and focus on a particular case. Unfortunately, taking assumptions usually cause some aspects to be omitted. This may not be desirable especially when these aspects are important for some reasons. That is why in our paper we first focus on native aspects of WSN and communication activities. Native aspects are the most important ones and cannot be omitted in modeling process, especially if we try to get reasonable simulation results. We define five native aspects of WSN and communication in WSN:

1. principle task of the WSN is to measure, collect and send data from nodes to the BS (one or many).
2. any WSN is created to achieve some globally defined aims. From this point of view, we may treat the WSN as one device performing tasks. However, WSN is a collection of spatially spread nodes, which take actions based on local information they have. Moreover, software that runs nodes is also implemented and executed on every node independently, having no information about the whole network, but rather some neighborhood of the node. It has to be ensured that local actions taken by each node cause the whole WSN to perform the globally defined aim.
3. cooperation and interference means that nodes influence each other through cooperation and disruption. Since disruptions arise from WSN properties and are unavoidable, thus one can only try to minimize its influence through proper cooperation between nodes. Cooperation is even more important in multi-hop networks where nodes cannot fulfill commissioned tasks on their own. In such situations, cooperation between nodes is crucial and is the only one way to achieve global aims. Positive aspect of cooperation and interference is the possibility to model both aspects of communication.
4. concerning network topology we assume that the topology remains unchanged throughout the whole lifespan of the network. Based on such assumption we can adjust topology of the WSN only once, during the deployment of the network.
5. optimization problem is focused on a maximization of WSN lifespan. Lifespan can be defined in many different ways (e.g. until the first nodes dies) but taking into account the principle task of the WSN we may assume that network dies when it cannot collect and send data from nodes to the BS.

2. NOVEL APPROACH BASED ON SETS AND RELATIONS

2.1. Motivation

Basic problem of our work is how to model a behavior of data flow (generated in WSN nodes) which traverses a network towards base station (BS). Even considering a simple model of such transmission, we came up against many problems. We consider a sensors network composed of nodes, which all reside in the communication range of the base station. Sensor measures parameters of its surrounding environment and transmits this data to the BS. This is a typical way people used to think about the WSN simplifying its operation to point-to-point communication.

It is usually assumed that network is a set of independent homogenous nodes and such simplification of communication activity model is unacceptable due to number of different transmission aspects. Practically, separation of two transmissions: from node *A* to the BS (A→BS) and (B→BS) is inadmissible. These two transmissions use the same radio communication channel , causing collisions, arbitration and priorities important and native issues that have to be solved. Assumption of point-to-point transmission omits vital aspects of WSN communication activity, so it is unacceptable. In fact, in order to model WSN communication activity it is necessary to consider set-to-set (set of sensors to set of base stations) transmission.

A multi-hop WSN networks with limited radio communication range and restricted energy are also widely studied in the literature (Fang 2005; Veyseh 2005). In order to send data from a node to the BS in such networks, it is necessary to use relayed transmission. It causes even more challenging problems because collisions could occur for any element of the routing path. An abundance of routing path elements yields collision, arbitration or priorities problems.

Abundance of interferences forces again point-to-point approach of a transmission. We determine routing path between data source and base station and next we model a multi-hop transmission. Such approach settles and simplifies our theoretical consideration, but a process of path determining causes new problems. If we determine routing path rarely it causes abundant load of routing path nodes (unbalanced distribution of energy consumption). If we determine path too often, we waste energy and communication channel resources more than necessary.

This is a reason why a number of papers focus on optimization of routing path selection in WSN. Developing flat (Burmester 2007), data-centric, hierarchical (Manjeshwar 2001; Sung-Min 2007) or location based routing protocols as well as developing reactive or proactive scenarios. We are working on problems that result from accepted assumption but are not native WSN problems.

Communication activity in WSN should be considered as set-to-set (set of sensors to set of base stations) transmission. Hence, we postulate a set theory as a tool for modeling this type of WSN activity. Such decision is very well justified. Already published works take advantages of functions, which are defined using the language of a set theory. Therefore using a set theory in our approach allows integrating novel approaches with solutions proposed so far.

Functions are nothing but restricted relations and relations can be viewed as a multivalued functions. Restricting relations into function for modeling communication activities in WSN leads to many problems and difficulties. Hence, we postulate relational approach as more general one. However at any time and whenever it is necessary, it is possible to reduce prepared model to traditional (functional) conditions.

2.2. Relations

As mentioned above, the novel approach proposed in this paper is based on such abstract fields of mathematics like theory of relations and sets. To describe communication activities in WSN we involve three binary relations, which are defined on a set of actions (*Act*). These relations (represented as a set of ordered pairs <*x*,*y*>; where $x,y \in Act$), namely: subordination, tolerance and collision (Jaron 1978, Nikodem 2008) are defined as follows:

Subordination $$\pi := \{< x, y >; x, y \in Act \mid x \pi y\}, \qquad (1)$$

where $x \pi y$ – means that the action *x* is subordinated to action *y*. In other words *y* dominates over *x*.

Tolerance $$\vartheta := \{< x, y >; x, y \in Act \mid x \vartheta y\}, \qquad (2)$$

where $x \vartheta y$ – states that actions *x* and *y* tolerate each other.

Collision $$\chi := \{< x, y >; x, y \in Act \mid x \chi y\}, \qquad (3)$$

where expression $x \chi y$ – means that actions *x* and *y* are in collision one to another.

Basic properties of π, ϑ and χ relations were discussed in (Jaron 1978). Here we outline only some of them:

$$\pi \cup \vartheta \cup \chi \subset Act \times Act \neq \emptyset, \qquad (4)$$

and

$$\iota \cup (\pi \cdot \pi) \subset \pi, \qquad (5)$$

where ι is the identity on the set *Act*. Eq. (4) states that all three relations are binary on non-empty set of actions *Act*. Eq. (5) states that subordination is reflexive ($\iota \subset \pi$) and transitive ($\pi \cdot \pi \subset \pi$).

Further

$$\pi \cup \vartheta^{-1} \cup (\vartheta \cdot \pi) \subset \vartheta, \quad (6)$$

where ϑ^{-1} is inverse of ϑ, means that:

- subordination implies tolerance - if π holds for some $x,y \in$ Act, then ϑ also holds for these,
- tolerance is symmetrical - if $x \vartheta y \Rightarrow y \vartheta x$,
- subordinated action tolerates all actions tolerated by the dominant - if $(x \pi y \wedge y \vartheta z) \Rightarrow x \vartheta z$.

For collision relation we have that

$$\chi^{-1} \cup (\pi \circ \chi) \subset \chi \subset \vartheta', \quad (7)$$

where ϑ' is the complement of ϑ. Eq. (7) states that collision is symmetric ($\chi^{-1} \subset \chi$) and disjoint to tolerance ($\chi \subset \vartheta'$). Moreover all subordinated actions must be in collision with action being in collision with its dominant (($\pi \circ \chi) \subset \chi$).

2.3. Neighborhood abstraction

A neighborhood abstraction, is defined (Nikodem, Klempous and Chaczko 2008) by a set of criteria for choosing neighbors and set of common resources to be shared, is very useful in almost all algorithms of WSN routing protocols. Realizing distributed operation in which nodes communicate only with other nodes within vicinity; sensor network takes advantage of some concept of a neighborhood. Each node selects some set of important neighbors within the network community and its activity is restricted to this set of nodes. Routing trees, graphs as well as ranges and clusters are specific types of neighborhoods.

Now, let us define *Map(X, Y)* as a set of mapping functions from *X* onto *Y* (surjection). Where Sub*(X)* is defined as a family of all *X* subsets. We define the neighborhood *N* as follows:

$$\mathcal{N} \in Map(Nodes, Sub(Nodes)). \quad (8)$$

Thus, *N(k)* is the neighborhood of node *k* defined as:

$$\mathcal{N}(k)\big|_{k \in Nodes} = \{y \in Nodes \,\big|\, y R_{\mathcal{N}} k\}, \quad (9)$$

In the paper (Nikodem 2009), the native neighborhood was advised as the most suitable form of the local range. Therefore, we define an indexed family of sets $\{N_i \,|\, i \in I\}$, where *I* denotes the index set and N_i has the following properties:

$$(\forall i \in I)(N_i \neq \emptyset) \wedge \bigcup N_i = Nodes, \quad (10)$$

$$(\forall i, j \in I \,\big|\, i \neq j)(N_i \cap N_j \neq \emptyset). \quad (11)$$

It means that native neighborhoods do not divide a set of WSN nodes into mutually exclusive subsets. Using a neighborhood abstraction we can try to decompose globally defined activities to locally performed identical task ascribed to each node of the network. It will not be an easy task to cast all global dependencies from network area to the neighborhood one. It will be even more difficult because neighborhood conditions for the network nodes might be, and usually are, quite dissimilar.

3. BASIC CONCEPT

A local/global activity dilemma is a starting point of our consideration of modeling communication activity in WSN. We split all-important aspects of communication activity into two classes. First class is composed of invariable aspects, second class relates to_aspects with local/global or $local_1$/$local_2$ sensibility.

The network topology and node's energy states constitute the first (invariable aspects) class. In contrast, cooperation and interference have been taking into account as second (relative aspects) class.

3.1. Digital terrain model and drained surface

Using digital terrain model (DTM) tools, we model communication activity aspects as surfaces, stretched over WSN network. According to this methodology, a result is obtained as a superposition of a few digital surface models (DSM). Each component (i.e. digitally modeled surface) describes some aspect-related additive properties.

When modeling data flow from network area towards base station we do this similar to rainwater surface flow. Data produced in WSN nodes flow like raindrops which streaming down in a direction determined by a slope of the modeled surface. During this process, drops merge with another (data aggregation), carve terrain or build it like lava tears (energy consumption). A resulted flow has been finally conditioned by the local neighborhood conditions and environmental stimulus (cooperation and interference).

We model natural network topology features using digital surface model (DSM). It is a component of a topographic map (bare drainage surface), which gives a basic reference frame that ensures messages are send towards the BS. In a real WSN network nodes usually have no information about their Euclidean distance from the BS. Therefore in the paper (Nikodem, Klempous, Nikodem, Chaczko and Woda 2009) we propose a measure of *dis(k)* (distance between BS and node *k*) based on maximal node's energy (E_{max}) and the amount of hops (*h*) required to send data from node *k* to the BS

$$dis^h(k) = 0.95 * h * E_{\max}. \quad (12)$$

Now, a bare drainage surface can be defined as follows:

$$D(x, y, z) = \{z = dis^h(k) \,\big|\, k(x, y) \in N\}. \quad (13)$$

The data required for representation (13) is collected and processed by the whole WSN area during

the self-organization process. This enables to determine nodes that are one hop closer to the BS immediately. It is vital that the message from a node k traverses in a direction determined by a slope of the modeled surface (13). Surface D created in such a way corresponds to spatial localization of WSN and is invariant in time since we assume that the network is not mobile. For these reasons, digital surface model (13) is used in our algorithm as bare drainage surface.

For the purpose of modeling node's instant energy level, we construct another surface

$$L_{en}(x,y,z) = \{z = E_{con} \mid k(x,y) \in N\}. \qquad (14)$$

In that case we assign nodes the energy E_{con} they spent (their consumed energy). Therefore, if more energy is used by a node, then greater value of the coordinate z will have a surface above this node.

Modeled surface L_{en} represents node's instant level of consumed energy. The communication activity during a network lifetime declines these levels, so a surface (14) also tends to fall off. Because of that surface L_{en} is recalculated all the time during the simulation process.

3.2. Cooperation and interference relational model

In case of cooperation and interference (Vakil 2006), the problem is more sophisticated than considered above. For a modeling: cooperation and interference purpose it is not reasonable to construct surface draped over WSN network. Aspect of cooperation and interference relates not only on global/local dilemma but first of all varies from one neighborhood to another. This is the reason that attempts to model cooperation and interference based on modeling of global surface stretched over WSN have been failed.

We focused our attention on two aspects when modeling cooperation and interference in WSN. First is a cooperation interpreted as a method of achieving globally defined strategy through tactics i.e. activities performed locally by each node. This aspect can be ensured if relational attempt is used.

Global strategy determined by intensity quotients of π, ϑ, χ relations is determined by base stations that adopt the strategy to the actual state of the WSN and situation. Later on, this strategy is send to nodes that setup their tactics in order to achieve the strategy. Using tactics each node performs operations within its neighborhood interacting and sending data to its neighbors. Additionally node's measure parameters of the environment they reside and align their operation accordingly.

Because both: node's neighborhood and environment differs for each node therefore for a given global strategy each node chooses some of its neighbors, he will cooperate according to subordination, tolerance and collision relations. Since these relations may differ for each node, we are not able to represent tactics as one common surface spread over the whole WSN. On the other hand, cooperation can be modeled individually for each neighbor as follows:

$$L_{rel}^{N(k)} = \{z = f(n,\pi,\vartheta,\chi) \mid n(x,y) \in N(k)\}. \qquad (15)$$

Second aspect concerns how to bind together different tactics that implement the same global strategy. Due to different local conditions and interactions with environment, tactics performed by each node may vary. If there are w nodes in the neighborhood of node *k*, then there may be up to *w+1* different tactics neighboring nodes may take in order to achieve the strategy.

DTM approach requires a construction of a surface that represents cooperation and interference. Surface that represents the global strategy is simply a plane since there is only one strategy for the whole network. On the other hand, surface that represents tactics of all nodes such that local interactions with the environment are also considered is difficult to draw. It is so, since surfaces representing the tactics of each node are different and span over the neighborhood area rather than the whole WSN. Therefore, for a network consisting of *n* nodes we get *n* surfaces that overlap. Since surfaces may differ, therefore it is difficult to draw one common surface that represents tactics of all nodes. On the other hand, tactics are implemented and performed by nodes and in this perspective; the interpretation of the model is easier.

Based on DTM each node may construct a drainage surface spread over its neighborhood. When constructing this function nodes can use bare drainage surface (13) and information about energy consumed (14), restricted both to its neighborhood. It can also use a relational surface (15) and superposition of these surfaces (13)-(15) constitutes drainage surface over the node's neighborhood.

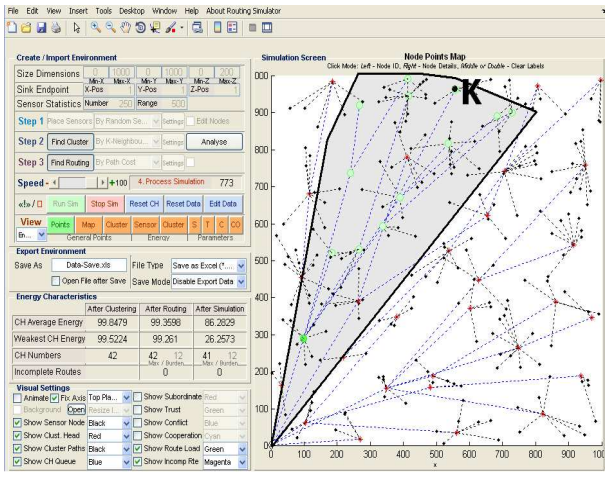


Figure 1: Modeling a spatial communication in WSN

4. CONCLUSIONS

Proposed application of sets theory and relations, allows solving the compliance dilemma posed against global

net of requirements along with their distributed local implementation within nodes of the network. At the same time, it proved that the data transmission treated

There is a demo situation presented on the fig.1, as set-to-set relation makes available new, feasible features to WSN modeling. It shows how a node *K* constructs to a routing path set which is simply an area (marked with grey color) through which data from *K* node is transmitted to the BS. Fig 1 compares proposed routing path set approach with traditional cluster and routing paths solutions.

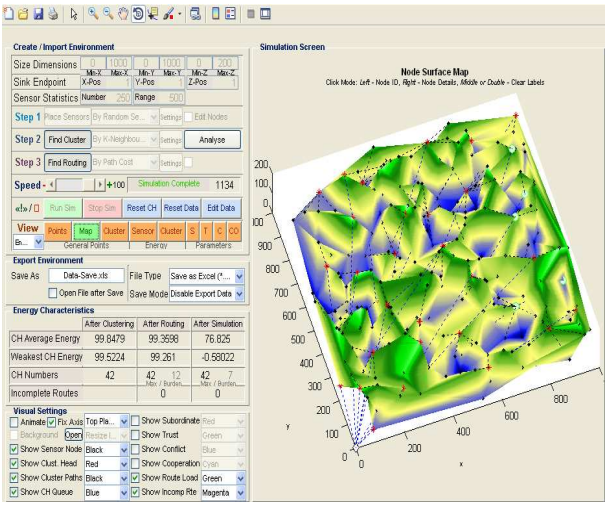


Figure 2: Modeling a drained surface in WSN

Fig. 2 presents a drainage surface generated for a certain network. Base modeled surface, was not so diverse at the beginning of simulation, a distinct surface slope clearly indicated drainage direction towards the BS. However, after a number of transmissions one may notice erosion of the drainage surface. One may also observe unevenness of energy use in particular nodes (cluster heads). These aforementioned inequalities can be leveled by a different tactic definition, from one, which is being currently used.

REFERENCES

Braginsky, D., Estrin, D., 2002. Rumor Routing Algorithm for Sensor Networks, *Proc. of the 1-st Workshop on Sensor Networks and Applications*, Atlanta, GA.

Burmester, M., Le T.V., Yasinsac, A., 2007. Adaptive gossip protocols: Managing security and redundancy in dense ad hoc networks, *Ad Hoc Networks*, Volume 5, Issue 3, pp.313-32.

Cohn, A.G., Bennett, B., Gooday, J.M., Gotts, N.M., 1997. Representing and Reasoning with Qalitative Spatial Relations about Regions. In: Cohn, A.G., Bennett B., Gooday J.M., Gotts N.M, eds. Spatial and Temporal Reasoning, Dordrecht, Kulwer, 97-134

Fang O., Gao J., Guibas L.J., de Silva V., Zhang L., 2005. GLIDER: gradient landmark-based distributed routing for sensor networks, *24th Annual Joint Conference of the IEEE Computer and Communications Societies*. INFOCOM 2005, 13-17 March 2005, vol.1, pp. 339-350.

Jaroń, J., 1978. Systemic Prolegomena to Theoretical Cybernetics, *Scient. Papers of Inst. of Techn. Cybernetics*, Wroclaw Techn. Univ., no. 45, Wroclaw.

Manjeshwar, A., Agrawal, D.P., 2001. TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks, *15th International Parallel and Distributed Processing Symposium* (IPDPS'01), *Parallel and Distributed Processing Symposium, International*, vol. 3, pp. 30189a, Workshops.

.Nikodem, J., 2008. Autonomy and Cooperation as Factors of Dependability in Wireless Sensor Network, *Proceedings of the Conference in Dependability of Computer Systems*, DepCoS - RELCOMEX, pp.406-413. Szklarska Poreba, Poland.

Nikodem, J., Klempous R., Chaczko, Z., 2008. Modelling of immune functions in a wireless sensors network.W: *20th European Modeling and Simulation Symposium*. EMSS 2008, Campora S. Giovanni, Italy.

Nikodem, J., 2009. Relational Approach Towards Feasibility Performance for Routing Algorithms in Wireless Sensor Network, *12 International Conference on Computer Aided Systems Theory and Technology* EUROCAST'2009,, *in: Lecture Notes in Computer Science* (in printing), Springer-Verlag Berlin Heidelberg.

Nikodem, J., Klempous, R., Nikodem, M., Chaczko, Z., Woda, M., 2009. Directed Communication in Wireless Sensor Network Based on Digital Terrain Model, *2nd International Symposium on Logistics and Industrial Informatics* (in printing), September 2009, Linz, Austria.

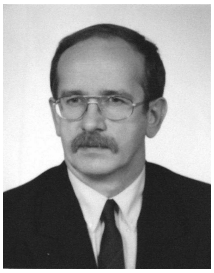
Sung-Min, J., Young-Ju, H., Tai-Myoung, C, 2007. The Concentric Clustering Scheme for Efficient Energy Consumption in the PEGASIS, *9th International Conference on Advanced Communication Technology*, Vol.1, page(s): 260-265

Vaidya, D., Peng, J., Yang, L., Rozenblit, J.W., 2005. A Framework for Sensor Management in Wireless and Heterogeneous Sensor Network, *12th IEEE International Conference on the Engineering of Computer-Based Systems*, pp.155-162. 4-7 April 2005, Greenbelt, USA.

Vakil, S., Liang, B., 2006. Balancing Cooperation and Interference in Wireless Sensor Networks, *IEEE Communications Society on Sensor and Ad Hoc Communications and Networks*, SECON '06, vol.1, pp.198-206.

Veyseh, M., Wei, B., Mi,r N.F., 2005. An Information Management Protocol to Control Routing and Clustering in Sensor Networks, *Journal of Computing and Information Technology* - CIT 13 (1), pp.53-68.

AUTHORS BIOGRAPHY



Jan Nikodem received the B.Sc. in electrical engineering, M.Sc. in artificial intelligence in 1979 and Ph.D. degree in computer science in 1982 from Wroclaw University of Technology (WUT), Poland. Since 1986, he has been an Assistant Professor in the Institute of Technical Cybernetics, WUT. Since 2005 in the Institute of Computer Engineering, Automatics and Robotics (ICEAR). His current research are focused on the area of complex and distributed systems, cybernetics, wireless sensor networks and digital data transmission.

Ryszard Klempous holds a M.Sc. in Automation (1971) and Ph.D. in Computer Science (1980) from Wroclaw University of Technology (WUT). Since 1980 he has been an Assistant Professor in the Institute of Computer Engineering, Automatics and Robotics, WUT. Senior member of IEEE and NYAS, has already published over 90 papers in Optimization Methods and Algorithms, Simulation and Data Processing and Transmission.

Maciej Nikodem graduated a M.Sc. in Computer Science in 2003 and a M.Sc. in Control and Robotics in 2005 from the Wroclaw University of Technology, Wroclaw in Poland. In 2008 he completed Ph.D. studies in Computer Science at Faculty of Electronics, Wroclaw University of Technology. For last 5 years Maciej Nikodem has worked on Countermeasures to Fault Analysis, Boundary Scan Security as well as security aspects of Wireless Sensor Networks. Maciej Nikodem is a Assistant Professor in the Institute of Computer Science, Control and Robotics, Faculty of Electronics at WUT.

Zenon Chaczko completed a B.Sc. in Cybernetics and Informatics in 1980 and a M.Sc. in Economics in 1981 at the University of Economics, Wroclaw in Poland., as well as completed MEng in Control Engineering at the NSWIT 1986, Australia. For over 20 years Mr Chaczko has worked on Sonar and Radar Systems, Simulators, Systems Architecture, Telecommunication network management systems, large distributed Real-Time system architectures, network protocols and system software middleware. Mr Chaczko is a Senior Lecturer in the Information and Communication Group within the Faculty of Engineering at UTS.

Marek Woda is an Assistant Professor in the Institute of Computer Engineering, Control and Robotics at Wroclaw University of Technology. In 2001 he graduated at WUT. In 2007, he got PhD degree in Computer Science from the Faculty of Electronics WUT. His research interests focus on multi-agents systems, e-learning, Internet technologies. He participated in international projects sponsored by European Union (e.g. PL96-1046 INCO-COPERNICUS project “Multimedia Education: An Experiment in Delivering CBL Material”, VI FP Integrated Project FP6-IST 26600 DESEREC “Dependability and Security by Enhanced Reconfigurability”). He is the author of about 30 scientific articles and conference papers.

ROBUST TUNING PSS FOR MULTIMACHINE POWER SYSTEMS USING MULTIOBJECTIVE HYBRIDATION TECHNIQUE

A. Choucha, A. Hellal, S. Arif and A. Hadjadj

Université Amar Telidji – Laghouat

(a.choucha, a.hellel, a.arif, a.hadjadj)@mail.lagh-univ.dz

ABSTRACT-
Power system stabilizers (PSS) are usually used in power system plants to damp out power system oscillations. In this paper, a simultaneous coordinated tuning procedure of multiple power system stabilizers (PSSs) in multimachine power systems based multiobjective functions Hybridation technique using genetic algorithms (GAs) and the requit simulé (RS). To validate the effectiveness of this tuning approach in enhancing the stability of power systems, modal analysis and nonlinear simulations have been carried out on a multimachine power system**.**

Keywords:Genetic algorithm, Requit simulé, modal analysis, power system stabilizer, simultaneous tuning, small signal stability.

1. INTRODUCTION

In a power system, the rotor generators angular stability depends on two electromagnetic torque components; the synchronizing torque and the damping torque. A lack of synchronizing or damping torque leads respectively to an aperiodic or an oscillatory instability (Y.L. Abdel-Magid and M.A. Abido, 2003)

The excitation system with its voltage regulator has an important impact on the two torque components, and thus, on the stability. During a disturbance, the increase of the excitation generator voltage has the effect of increasing the internal machine voltage and the synchronization power. So, the use of rapid excitation systems has long been considered as an effective method to increase the synchronizing torque and transient stability limits (Y.L. Abdel-Magid, M.A. Abido, , S. Baiyat and A.H. Mantawy, 1999).

The static excitation systems, where the excitation current is provided by a controlled rectifier, seem to be the appropriate way to achieve both rapid performances and high gain; the limits of the transient stability system will be increased.

Unfortunately, this benefit may be outweighed by the excitation system effect on the damping oscillations by reducing the damping torque (F.P. DeMello and C. Concordia, 1969, A.L.B. Do Bomfim, G.N. Taranto and D.M. Falcao, 2000).

To compensate this effect and improve the damping system, additional signals provided by a Power System Stabilizer (PSS) are generally injected into the excitation system.

The PSS parameters fixed under certain operating conditions are set to the predetermined values. It is important to know that the generator parameters change with the load variation: the machine dynamic behaviour varies with the operating points.

The PSSs must be correctly adjusted and coordinated in such a way that the stability of the global system can be guaranteed for a wide range operating points.

In PSSs tuning, adjustment sequences and location are critical parameters for stabilizing optimal performance. A PSS can be adjusted to improve the damping mode. However, it can produce undesirable effects for other modes. Moreover, the various investments in of PSSs make oscillation behavior different according to the operating points.

In the literature, several approaches using genetic algorithms (GA/RS) have been proposed for coordinated tuning of multiple power system stabilizers (K. Hongesombut & all 2005, M.A. Abido *2000,* K. Hongesombut and Y. Mitani 2004, Y.Y. Hsu and C.L. Chen 1987, P. Kundur, 1994, E.V. Larsen and D.A. Swann, 1981). The main advantage of the AG compared to other optimization techniques is to be a global population optimization method where the solution is independent of the problems complexity.

In many searches, the location of PSSs is chosen before selected tuning methods. The participation factors (PF) method has been extensively used to identify the PSSs possible locations (K. Hongesombut & all 2005, M.A. Abido *2000,*Panda and N. Prasad Padhy, 2007, F. Rashidi and M. Rashidi, 2000).

Generally, when having too much and/or poorly positioned PSSs may cause dysfunctioning of the system. Thus, to reduce these undesirable effects, it is necessary to locate and select an appropriate number of PSSs.

The objective of this work is to ensure an optimum tuning of PSSs with a better location, while reducing their numbers. Hence, we have developed an hybridation method using AG/RS simult anneling program tackling a multi-objective function, based on the real part poles and the damping factor.

The multi-machines power system studied consists of 16 generators and 68 nodes; it represents the New England/New York interconnected network system.

2. PROBLEM FORMULATION

A widely used conventional lead-lag PSS is considered in this study. Its transfer function, given in (1), consists of an amplification block with a control gain K_{PSS}, a washout block with a time constant T_w and two lead-lag blocks for phase compensation with time constants T_1, T_2, T_3 and T_4.

$$V_{PSS}(s) = K_{PSS} \cdot \frac{sT_w}{1+sT_w} \cdot [\frac{(1+sT_1)}{(1+sT_2)} \cdot \frac{(1+sT_3)}{(1+sT_4)}] \cdot \Delta\omega(s) \qquad (1)$$

Where, the PSS output signal, V_{PSS}, is a voltage added in the system excitation input. The generator speed deviation $\Delta\omega$ is almost used as an input signal of PSS.

In small signal stability studies, the linearized system model around an equilibrium point is usually applied (P. Kundur, 1994).

The analysis of the system eigenvalues, resulting from model linearization, provides important information about the system stability. It indicates the presence of unstable or poorly damped modes and identifies the characteristics of these modes and their origins.

To complete the analysis of system stability and confirm the results obtained from the eigenvalue analysis, nonlinear time domain simulations should be applied on the system (Y.Y. Hsu and C.L. Chen 1987).

The criteria used for examining the results are based on eigenvalues analyses: the real part (σ) of an eigenvalue (λ), given in (2), and the damping factor (ζ), given in (3) (P. Kundur, 1994).In PSS tuning, All system eigenvalues must be placed in the D stability region in the s-plan; this region is determined by the following criteria: $\sigma_{cr} = -1$, $\zeta_{cr} = 10\%$ (Y.L. Abdel-Magid and M.A. Abido, 2003).

$$\lambda = \sigma \pm j\omega \qquad (2)$$

$$\zeta = \frac{-\sigma}{\sqrt{\sigma^2 + \omega^2}} \qquad (3)$$

The multiojective function is formed to optimize a composite set of two eigenvalue-base objective functions, comprising the real part eigenvalue (σ) and the damping factor (ζ) of the dominate electromechanical modes.

We propose in this study a GA program that allows to choose the best PSS location and to reduce their number simultaneously, and we use GA/RS for coordinated tuning of PSS parameters.

In order to evaluate the approach performance and consequently the damping performance, the study will be divided in three applications. In the first one (**Case A**) the PSS locations and their number are predetermined by Participation Factor method and the PSSs are then tuned by GA programme. In second application (**Case B**), we have used the same PSS number as the previous application but we have used the GA for finding PSS locations and tuning their parameters simultaneously. This application can be represented graphically by Figure 1. In the last application (**Case C**), PSS parameters, are optimized simultaneously by GA/RS programme..

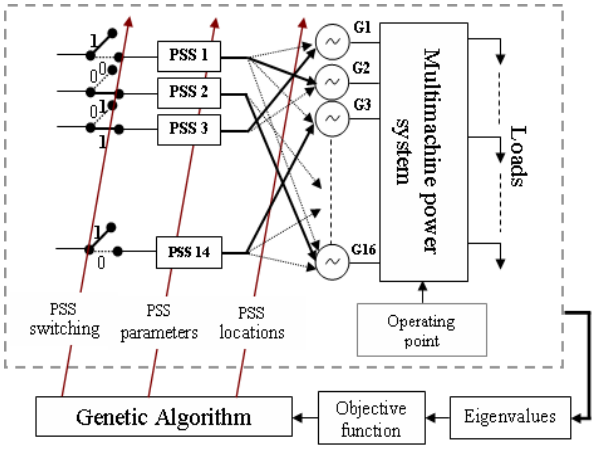


Figure 1. Graphic representation of case C.

Each PSS is connected in series with a switch that can take the values of 0 or 1. The variables representing these switches are added to the other PSS variables of the AG programme. The switch can connect (if its value is 1) or disconnect (if its value is 0) the PSS of the associated generator which can permit finally to reduce the PSS number. During the optimization, when the multiobjective function does not vary significantly because of little influence of some PSSs on the system damping, the AG will not take them into consideration and will eliminate them from the list of PSSs to be installed.

3. APPLICATION RESULTS AND DISCUSSION

The applications were made on New England/New York multimachine power system (16 generators and 68 buses). Details of the system data can be found in (G. Rogers, 2000).

To test the PSS performance tuning, eigenvalues analysis of linear model will be done and nonlinear simulation analysis will be also carried out. The nonlinear time domain simulations were performed for a three phase-fault, with duration of 100 ms on the line 59 # 23. The simulation results represent the speed variations of five generators, the most affected by this defect; they are G.56, G.58, G.59, G.61 and G.68.

The open-loop electromechanical modes of the system at the studied operating point are shown on Figure 2. The mode repartition shows obviously that the system is strongly unstable.

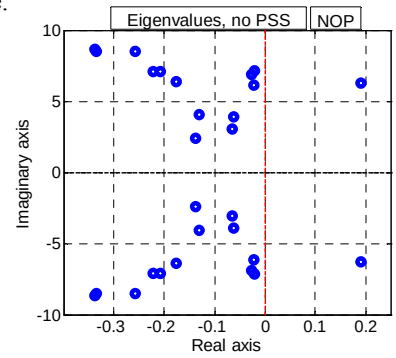


Figure 2. Electromechanical modes of the open-loop system.

3.1 Cases A

Participation factor method is used in this application to find the suitable generators for installing PSSs. 14 generators are found necessary by this method. The coordinated synthesis of parameters for these 14 PSSs is optimized using genetic algorithms. Figure3 demonstrates the electromechanical modes of the closed-loop system using these optimized 14 PSSs. The analyse of system eigenvalues gives a minimum damping factor ζ_{min} = 16.19%and a maximum eigenvalue real part $\sigma_{max} = -0.991$.

Hence, all electromechanical modes were generally shifted in the *D* stability region.

Analyse of generator speed variations, shown on Figure 4, demonstrates clearly that the system becomes stable and the oscillations are damped in less than 10 seconds.

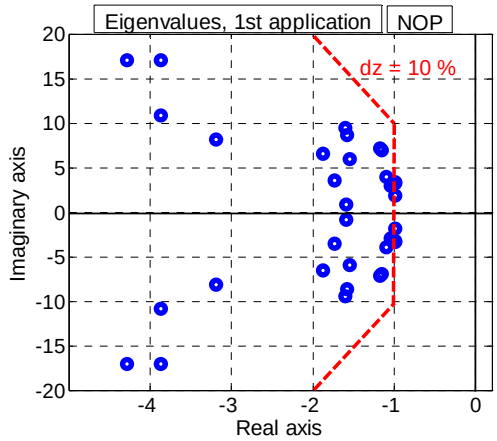


Figure 3. Electromechanical modes, cases A.

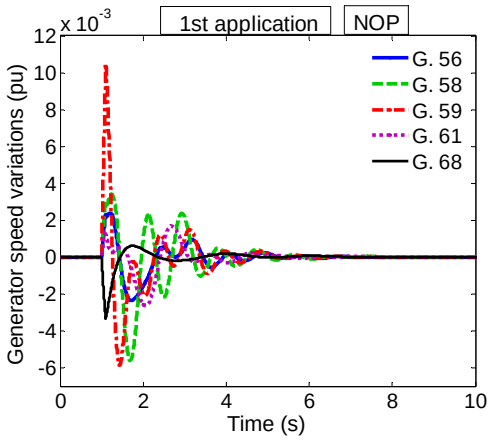


Figure 4. Generator speed variations, cases A.

3.2 Cases B

Considering a fixed number of PSSs equal to 14, these PSS locations are considered as variables to be optimized simultaneously with PSS parameters using a GA programme.

Figure 5 shows the s-plan system mode repartition. The modes are clearly more shifted in the *D* stability region. The minimum damping factor and maximum eigenvalue real part are respectively ζ_{min} = 17.28% and $\sigma_{max} = -1.338$.

Thus well improvements are won on the minimum factor damping (of 6.7%) and on the maximum real part eigenvalue (of 35%) compared to the first application.

Figure 6 demonstrates the non-linear time domain simulation of generator speed variations. The evaluation of system damping performance shows that the settling times of generator speed variation response are improved very well compared to the first application; for example, oscillations of generators (61 and 68) are damped two times faster than their equivalents in the first application.

Thus, the system stability can be improved very well when PSS locations are chosen by GA.

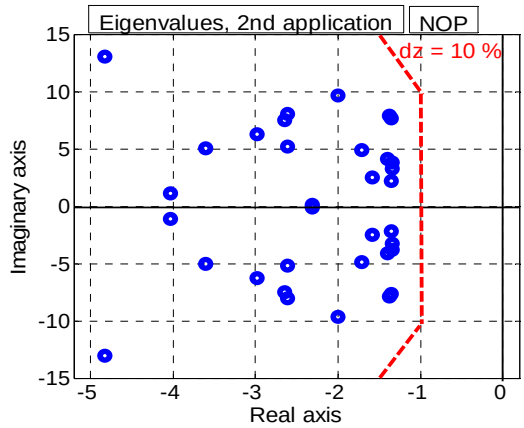


Figure 5. Electromechanical modes, cases B.

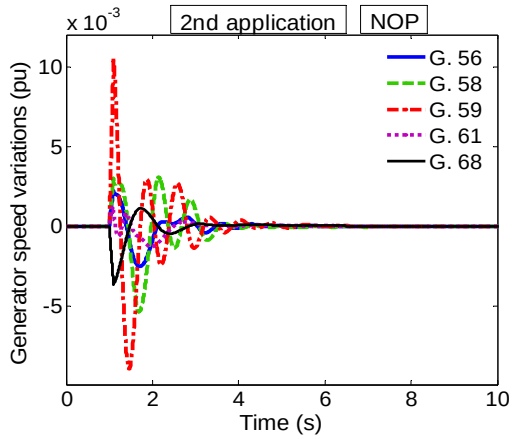


Figure 6. Generator speed variations, cases B.

3.3 Cases C

Considering in this application 12 PSSs number and to be optimized simultaneously with their parameters, we find that are quite sufficient to get a well stable system.

Figure 7 shows the system electromechanical modes, with ζ_{min} = 10.42% and $\sigma_{max} = -0.912$. All the modes are then shifted in the *D* region. We notice that we get almost the same performance as compared with case A, using only 12 PSSs.

Figure 8 gives the simulation results of generator speed variations. The overall stability of the system is thus ensured and the oscillations are damped in less than 10 seconds.
This confirms that it is necessary to choose the PSS optimal locations simultaneously with a coordinated tuning while using only the necessary PSS number.

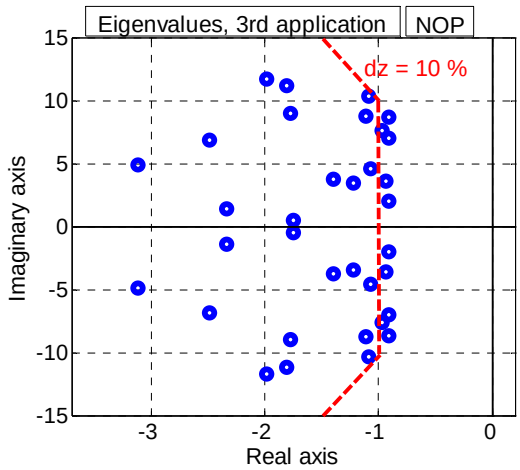


Figure 7. Electromechanical modes, cases C.

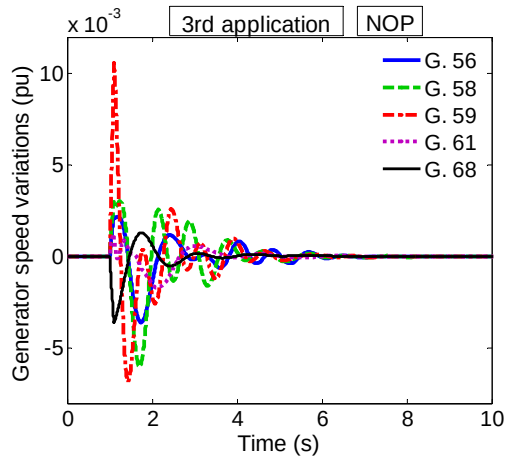


Figure 8. Generator speed variations, cases C.

4. CONCLUSION

In this study, a simultaneously tune multiple power system stabilizers (PSSs) based on multiobjective functions by using genetic algorithms (GA/RS) is presented. Eigenvalues analysis and nonlinear time domain simulations are done to verify the effectiveness of this technique in enhancing the small signal stability of multimachine power systems. The analysis of results showed that the dominant electromechanical modes are well shifted in the D-stability region. To avoid using the participation factor method to find the possible PSS locations, the GA is proposed to search the best PSS locations. The results demonstrate a well improvement of close-loop plant performance. The GA is also proposed to reduce the needed number of PSSs to realise good damping system. Using hybridation method in the nonlinear time domain simulations which are applied for large disturbance and for many operating conditions has also confirmed the obtained results and proved that the power system oscillations are well damped.

5. REFERENCES

[1] Y.L. Abdel-Magid and M.A. Abido, 2003 "Optimal Multiobjective Design of Robust Power System Stabilizers Using Genetic Algorithms," *IEEE Trans. Power System*, Vol. 18. n°. 3, pp. 1125-1 132,

[2] Y.L. Abdel-Magid, M.A. Abido, , S. Baiyat and A.H. Mantawy, 1999 "Simultaneous Stabilization of Mulfimachine Power Systems Via Genetic Algorithms," *IEEE Trans. on Power Systems*, Vol.14, n°.4, pp.1428-1439,

[3] F.P. DeMello and C. Concordia, 1969 "Concepts of synchronous machines stability as affected by excitation control," IEEE Trans. *On Power Apparatus and Systems*, Vol. PAS-88, pp. 316-329,

[4] A.L.B. Do Bomfim, G.N. Taranto and D.M. Falcao, 2000 "Simultaneous Tuning of Power System Damping Controllers Using Genetic Algorithms," *IEEE Trans. Power Sys.*, Vo1.15, pp.163-169,

[5] K. Hongesombut, S. Dechanupaprittha, Y. Mitani and I. Ngamroo, 2005 "Robust power system stabilizer tuning based on multiobjective design using hierarchical and parallel micro genetic algorithm," 15th *Power Systems Computation Conference (PSCC)*, Liege, Belgium,.

[6] M.A. Abido *2000* "Simulated annealing based approach to PSS and FACTS based stabilizer tuning," *Electrical Power and Energy Systems 22 247* – 258

[7] K. Hongesombut and Y. Mitani 2004 "Implementation of Advanced Genetic Algorithm to Modern Power System Stabilization Control," *IEEE PES, Power Sys. Conference & Exposition*, Vol. 2, pp.1050-1055

[8] Y.Y. Hsu and C.L. Chen 1987, "Identification of Optimum Location for Stabilizer Applications Using Participation Factors," *IEE Proc.*, vol. 134, No. 3, pp. 238-244

[9] P. Kundur, 1994 *Power System Stability and Control*, The EPRI Power System Engineering Series, McGraw-Hill. Inc..

[10] E.V. Larsen and D.A. Swann, 1981. "Applying power system stabilizers," *IEEE Trans.*, Vol. PAS-100, no. 6, pp. 3017-3046,

[11] Panda and N. Prasad Padhy, 2007 "Power system with PSSs and FACTS Controller: Modelling, Simulation and Simultaneous Tuning Employing Genetic Algorithm," *International Journal of Electrical, Computer and Systems Engineering*, Vol.1, No. 1, pp.9-18,.

[12] F. Rashidi and M. Rashidi, 2000 "Tuning of Power System Stabilizers via Genetic Algorithm for Stabilization of Power Systems," *Innovations in Applied Artificial Intelligence Ed.* Springer Berlin 2004, pp. 1210-1219.

[13] G. Rogers, 2000 *Power System Oscillations*, Boston, Kluwer Academic Publishers..

AUTHOR BIOGRAPHY

Abdelghani CHOUCHA was born on April 28, 1963 in Laghouat. He obtained his Electrical Engineering degree in 1987 from USTO of Oran (Algeria) and his Magister diploma in Electrical Power System in 2000 at Tiaret University (Algeria).

He is currently Researcher Teacher at the Electrical Engineering Department, Amar Telidji University of Laghouat. His research interests are mainly stability and security in power systems using new techniques.

VEHICLE DYNAMICS CONTROL USING BOND GRAPH WITH SLIDING MODE OBSERVERS

A.Naamane, N.K. Msirdi

LSIS UMR CNRS 6168
Avenue Escadrille Normandie Niemen
13397 Marseille cedex 20

Email: aziz.naamane@lsis.org; n.msirdi@lsis.org

ABSTRACT

This paper deals with the design of a sliding mode observer using Bond graph for the control of the dynamics of a vehicle. The control objective is to define and develop effective procedures making it possible to observe dynamics correctly in a robust way and to detect certain situations sufficiently early. This method is shown to be robust with respect to perturbation and parametric uncertainties. Experimental results illustrate the efficiency of the proposed approach.

Keywords: Bond graphs, sliding mode observers, control, diagnosis workstation design, work measurement, ergonomics, decision support system

1. INTRODUCTION

During these last decades important studies (Hocine,2003),(Rill and,Zbiri 2005), were carried out to integrate driver assistance system with more and more control with a great number of methods of observation, and detection of critical situations. The goal consists to define and develop effective procedures making it possible to observe dynamics correctly in a robust way and to detect certain situations sufficiently early. The vehicles are a complex Dynamics System with unknown inputs (like contact forces, road profile, external perturbations…). Their behaviour is affected by several factors that may depend or not on its structure. The external influence depends mainly on the contact between the pneumatic tyre and the road and the aerodynamic forces introduced by the wind flowing around it. Tire forces affect the vehicle dynamic, performance and behavior properties.

Robust observer with unknown inputs is shown efficient for estimation of road profile and for estimation of the contact forces. Different dynamic controls on the vehicle like tracking, braking and cornering, reduce the friction coefficient. The traction and braking control reduces the wheel slip, and this can be done by the use of sliding mode approach for observation and control. For vehicles and road safety analysis, it is necessary to take into account the contact force characteristics. However, the friction coefficient and different force (like traction force) cannot be directly measured. They are complex to precisely represent by some deterministic model equations. Usually some experimentally fitted and approximated model are used to deduce their values.

In this work we develop a method to observe tire forces. The proposed estimation procedures have to be robust, and can then be used to improve the security detecting some critical driving situation. This estimation can be used in several vehicles control system such as Anti-lock Brake System (ABS), traction control system (TCS), diagnosis systems, etc... An observer is then proposed to estimate the forces and friction coefficient. The estimations are produced using only the angular wheel position and longitudinal velocity as measurement and they are the input to the specially designed robust observer based on the Second Order Sliding Modes (SOSM). (Imsland and al 2006) The method of estimation is verified through simulation with as contact model a "Pacejka Model" (Magic Formula). (Pacejka & al. 1997,2000),

2. OBSERVERS

The readers can find in (M'Sirdi & al. 2003,2007), (Rabhi 2004,2005) the different models of the whole vehicle. In this work, we focus on the use of the sliding mode technique with Bond Graph, to show, that this combination can replace advantageously the classical approaches. We use Second Order Sliding Modes to develop a second order differentiator in order to obtain the tire road friction estimation. (A. Levant,2003), (Fridman2004)

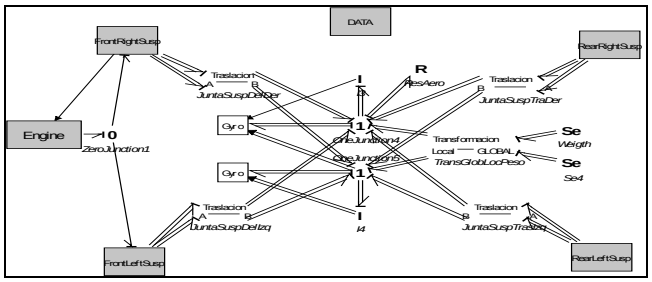


Fig 1: Complete Bond Graph Vehicle

2.1. High Order Sliding Mode Observer (HOSM)

Robust Differentiation Estimator (RDE) is used on this works to deduce our estimations. Consider a smooth dynamics function $s(x) \in \square$. The loop containing this variable may be closed by some possibly-dynamical discontinuous feedback where the control task may be to keep the output $s(x(t)) = 0$.

Then provided that successive total time derivatives $s, \dot{s}, \ddot{s}, \ldots, s^{(r-1)}$ are continuous functions of the state space variables, and the sliding point set is non-empty and consist locally of Filippov trajectories.

$$s = \dot{s} = \ddot{s} = \ldots = s^{(r-1)} \quad (1)$$

The motion on set is called r-sliding mode (rth-order sliding mode). The HOSM dynamics converge toward the origin of surface coordinates in finite time always that the order of the sliding controller is equal or bigger than the sum of a relative degree of the plant and the actuator. To estimate the derivatives s_1 and s_2 we will use the 2nd-order exact robust differentiator of the form.

$$\begin{aligned} \dot{z}_0 &= v_0 = z_1 - \lambda_0 \left| z_0 - s_w \right|^{\frac{2}{3}} sig\ (z_0 - s_w) \\ \dot{z}_1 &= v_1 = -\lambda_1 \left| z_1 - v_0 \right|^{\frac{1}{2}} sig\ (z_1 - v_0) + z_2 \\ \dot{z}_2 &= -\lambda_2 sign(z_2 - v_1) \end{aligned} \quad (2)$$

where z_0, z_1 and z_2 are the estimate of s_w, s_1 and s_2, respectively, $\lambda_i > 0, i = 0,1,2$. Under condition $\lambda_0 > \lambda_1 > \lambda_2$ the third order sliding mode motion will be established in a finite time.

The obtained estimates are $z_1 = s_1 = \dot{s}_w$ and $z_2 = s_2 = \ddot{s}_w$ can be used in the estimation of the state variables and also in control.

2.2. Cascaded Observers and Estimator

This work uses the previous approach to build the observer and obtain an estimation scheme in 20-Sim.

We produce the estimation in steps using like input the wheel angular position and the longitudinal body speed. The inputs are considered available for measurements.

The robust differentiation observer is used for estimation of the velocities and acceleration of the four wheels.

1st Step: produces estimation of angular velocity of the wheel. The convergence of these estimates is guaranteed in finite time t_0.

$$\begin{aligned} \dot{\theta}_0 &= v = \hat{\omega} - \lambda_0 \left| \theta - \theta_0 \right|^{2/3} sign(\theta - \theta_0) \\ \dot{\hat{\omega}} &= v_1 = \dot{\hat{\omega}} - \lambda_1 \left| \hat{\omega} - v_0 \right|^{1/2} sign(\hat{\omega} - v_0) \\ \ddot{\hat{\omega}} &= -\lambda_2 sign(\dot{\hat{\omega}} - v_1) \end{aligned} \quad (3)$$

2nd Step: Estimation of the forces F_x (longitudinal) and F_z (vertical).

To estimate the F_x we used the following equation,

$$\hat{F}_x = \frac{(T - J.\dot{\hat{\omega}})}{R_{ef}} \quad (4)$$

The torque could be also estimated by means of use additional equation from engine behavior or measured.

To estimate the F_z we use the following equation,

$$\begin{aligned} \hat{F}_{zf} &= \frac{m}{2.(l_f + l_r)}.(g.l_r - h.\dot{\hat{v}}_x) \\ \hat{F}_{zr} &= \frac{m}{2.(l_f + l_r)}.(g.l_f + h.\dot{\hat{v}}_x) \end{aligned} \quad (5)$$

where (21) is for the front axis, and (22) is for the rear axis. $\hat{v}_x$ and $\hat{\omega}_x$ are produced by the Robust Estimator (RE).

3rd Step: We estimated the friction coefficient or pneumatic adherence.

$$\hat{\mu}_x = \frac{\hat{F}_x}{\hat{F}_z} \quad (6)$$

4th Step: We estimated the Rolling Resistance and Aerodynamics Resistance Force.

$$\begin{aligned} \hat{R}_{rolling} &= 0.01.(1 + \frac{3.6.}{160}\hat{v}_x).\hat{F}_{zi} \\ i &= 1..4 \\ \hat{F}_x^{Aer} &= \frac{1}{2}\rho_{air} C_x A_f . \hat{v}_x^{\ 2} \end{aligned} \quad (7)$$

5th Step: We estimated the Slope Angle.

$$\hat{\alpha} = \arcsin\left(\frac{\hat{F}_x - 1/2\rho_{air} A_{frontal} \hat{v}_{xprom} - \hat{R}_{rolling})}{M_v g}\right. \quad (8)$$

where $\hat{F}_x = \hat{F}_{x1} + \hat{F}_{x2} + \hat{F}_{x3} + \hat{F}_{x4}$

$\hat{R}_{rolling} = \hat{R}_{rolling1} + \hat{R}_{rolling2} + \hat{R}_{rolling3} + \hat{R}_{rolling4}$

are the total traction forces and rolling resistance (each by tire-four wheels).

2.2 Implementation in 20-Sim

The observer is build in 20-Sim with blocks element. 20-Sim admits the interaction between bond graph and signal components (Getting Started with 20-Sim 3.6", *2005)*. We used the bond graph vehicle model to get the powers variables using sensors (flow and effort sensors), and are introduced in the observer and estimator modules like signal mode. The inputs in the observer module are the tire angle position and the measured longitudinal body velocity.

The observer module produces the estimated variables used in the estimator module to get the longitudinal and vertical forces. The friction coefficient is also estimated. Torque is measured with a sensor on the bond graph model, it is an input for the estimator module.

Figure 7 shows the general observer module where the inputs are the measured variables and the outputs are the first and the second derivative estimated. In our work the inputs are the angular position and the longitudinal body velocity, and the outputs are the angular acceleration estimated and longitudinal acceleration estimated. It is possible to get the angle velocity estimated and then the estimated angle acceleration.

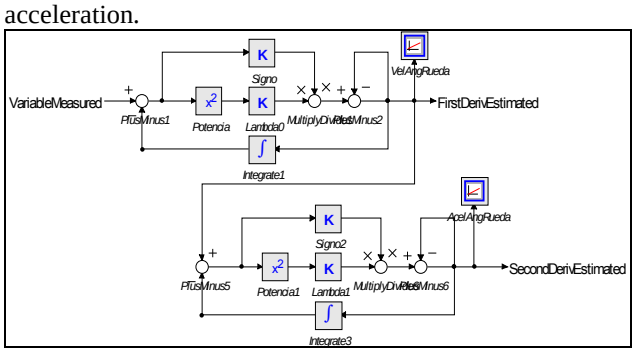


Figure 1: General Observer module with blocks diagram

Figure 8 shows the F_x estimated module, where the inputs are the estimated tire angular acceleration (estimated in the observer module) and the torque measured (or estimated). On this module other parameters are needed to estimate the longitudinal force (F_x) by the equations (18).

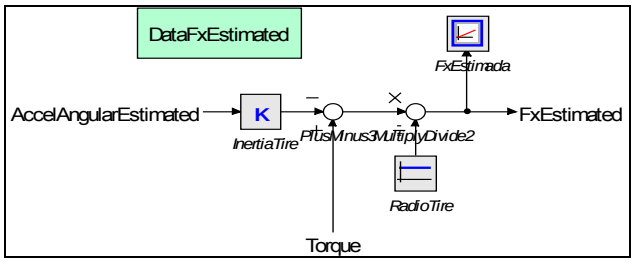


Figure 2: Fx Estimate Module

Figure 9 shows the F_z estimated module, where it has as input the longitudinal acceleration estimated (estimated in the observer module). On this module other parameters are needed to estimate the vertical force (F_z) by the equations (19) and (20).

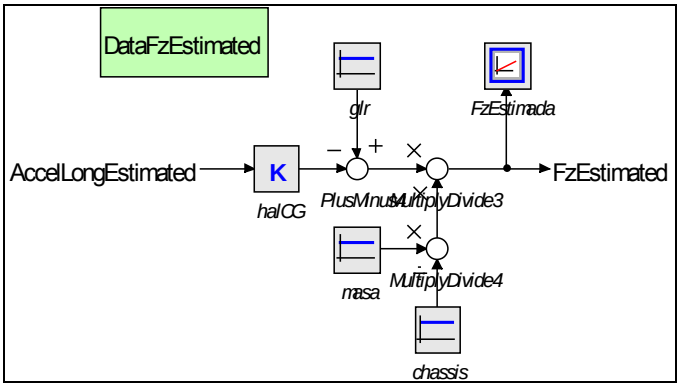


Figure 3: Fz Estimate Module

Figure 10 shows the μ estimated module, where it has as inputs the Longitudinal and Vertical Forces estimated ($\hat{F}_x, \hat{F}_z$) by the equation (21).

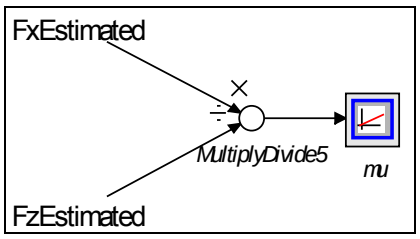


Figure 4: μ Estimate Module

Figure 11 shows the Rolling resistance estimated module, it has as inputs the Longitudinal Velocity measured and Vertical Forces estimated ($\hat{v}_x, \hat{F}_z$) by the equation (22).

This module is used in each wheel to estimate all different rolling resistance. These modes of use make an optimal estimated slope angle, so we can make individual estimation for each wheel.

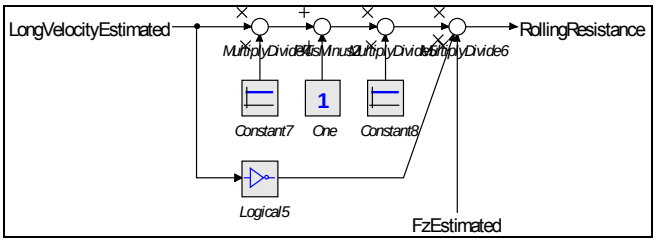


Figure 5: Rolling Resistance Estimate Module

Figure 12 shows the Aerodynamics Resistance estimated module, where it has as input the Longitudinal Velocity ($\hat{v}_x$) by the equation (23). Other coefficients like C_x, Area, ρ_{air} are inputs for the module.

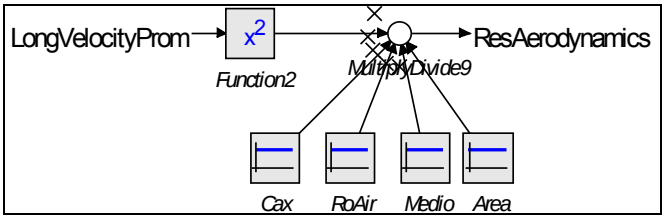


Figure 6: Aerodynamics Resistance Force Estimate Module

Figure 13 shows the complete estimated module for the slope angle, where the inputs are the estimated variables from each wheel. Then we estimate each rolling resistance, and the aerodynamics resistance force. Finally we can obtain the Slope angle estimated by the equation (24). Other coefficients are necessary like inputs for the estimated module.

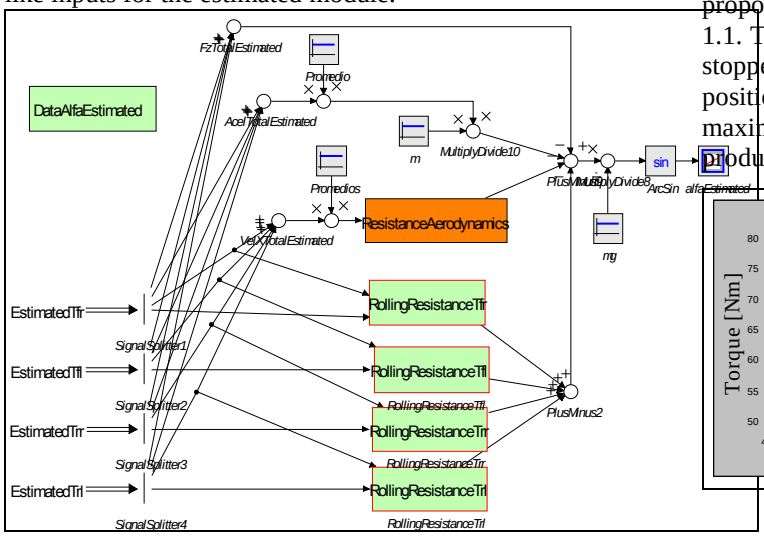


Figure 7: Complete Slope Angle Estimate Module

Figure 14 shows the complete bond graph model for the suspension system, where the mechanical suspension system is modeled with bond graph, sensor placed on the bond graph model and the observer and estimated modules with their signal inputs.

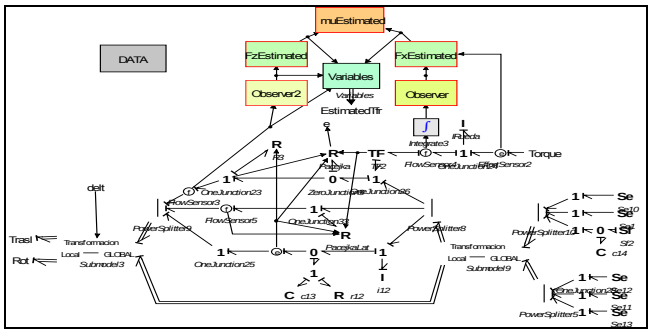


Figure 8: Suspension System with Observer and Estimate modules

Figure 15 shows the complete bond graph model for the vehicle, where we have the BG vehicle model and the Slope Angle Estimated Module.

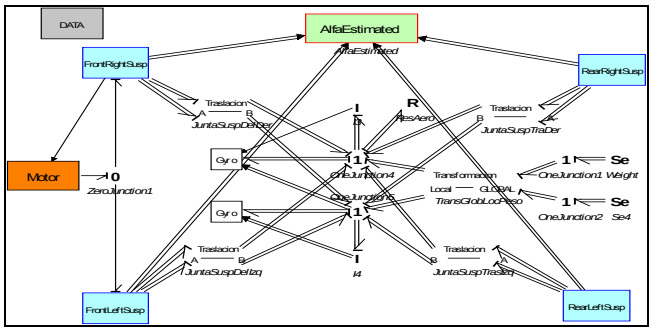


Figure 9: Vehicle BG Model and Slope Estimate Module

3. SIMULATION

In this part we show the results obtained with simulation on 20-Sim. The simulation shows us the behavior of vehicle dynamic and validates our approach and the proposed observers. The state and forces are generated by the Bond Graph Vehicle dynamics proposed. The data used are from a car Renault Clio RL 1.1. The simulation begins with zero velocity (vehicle stopped). On the second 8, the accelerator changes his position from 0 to 1 (the butterfly valve obtains maximum position) and begins the acceleration. It produces a torque from the engine to the traction tires.

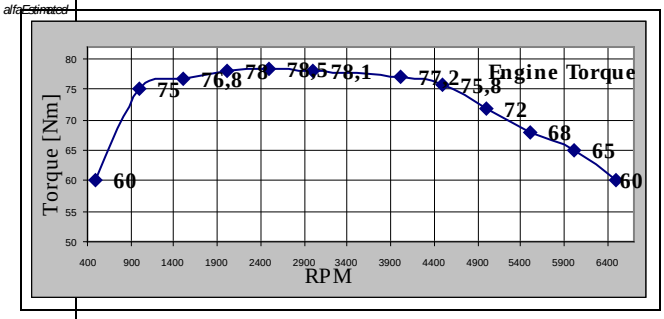


Figure 10: Engine Torque

At time=100 second, the accelerator passes from 1 to 0 causing the deceleration of the vehicle. Figure 17 shows the accelerator profile between 0 and 1 position. We propose a variable slope angle as it presented in table 1.

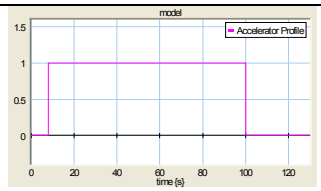


Figure 11: Accelerator Profile

Time	Slope Angle Proposed
0s-10s	0°
10s-20s	5°
20s-50s	10°
>50s	3°

Table 1: Slope Angle Proposed

Figure 18 shows the engine behavior, with the RPM, Torque and speed gearbox. Figure 19 shows the principal behavior of the vehicle, with Speed on axis 'X', Speed on axis 'Z', Pitch Angular Speed, Longitudinal and Vertical Position of the Centre Gravity. Figure 20 shows the Chassis Position on the Z global axis.

Figure 21 shows the tire angular velocity estimated and measured. The figure shows the good convergence to tire angular velocity.

Figure 22 shows the angular acceleration estimated and linear acceleration estimated. The 2^{dn} step on the observer model gives us the estimated longitudinal force F_x and the vertical estimated force F_z. Figure 23 shows the F_x estimated and real (vehicle simulation).

Figure 24 shows the F_z estimated and real (vehicle simulation).

The 3rd step on the observer model gives us the friction coefficient estimated.

Figure 25 shows the friction coefficient tire estimated and the friction coefficient obtained with the Pacejka model.

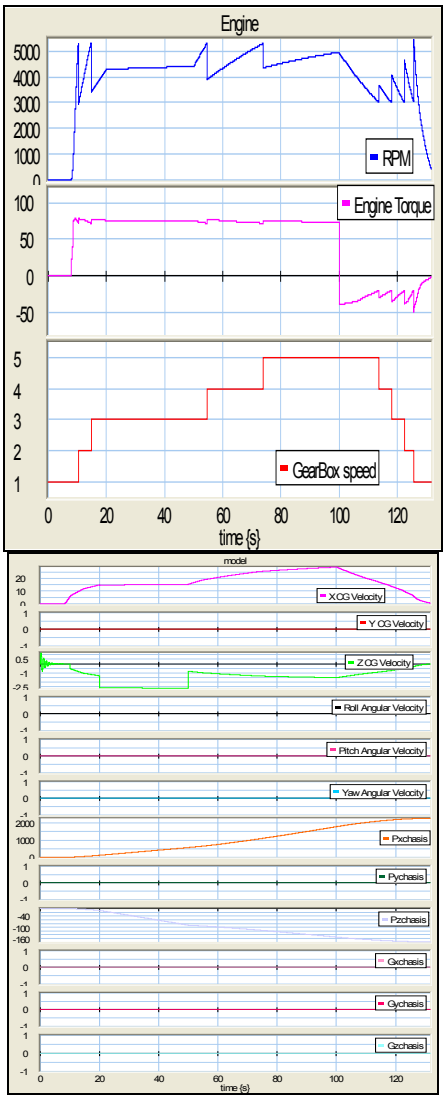


Figure 12: Principal behavior Centre of Gravity, Figure 13: Engine Behavior

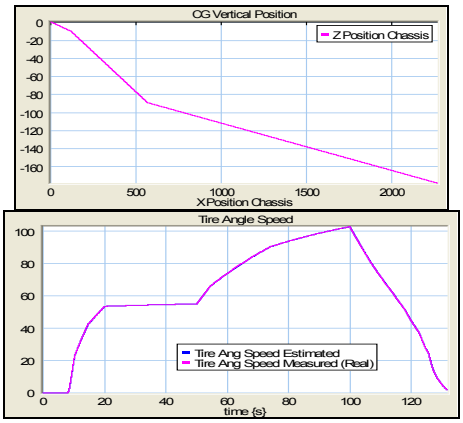


Figure 14: Z Chassis Position Figure 15: Tire Angular velocity Estimated and Measured

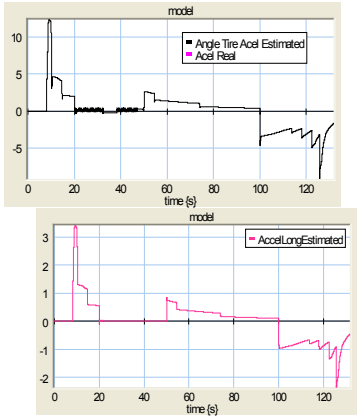


Figure 16: Angular Acceleration and Linear Acceleration Estimated

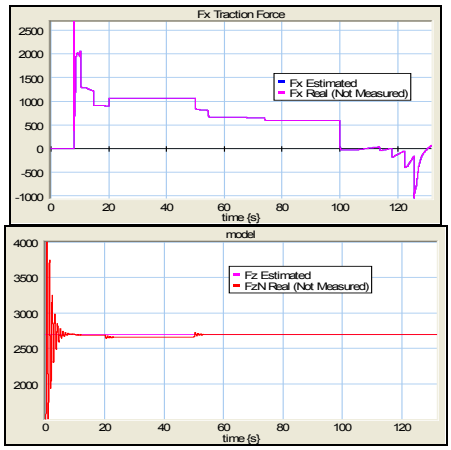


Figure 17: Longitudinal Force Estimated and Simulated; Figure 18: Vertical Force Estimated and Simulated

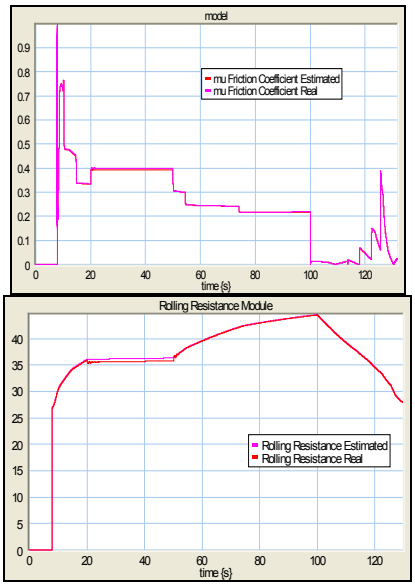
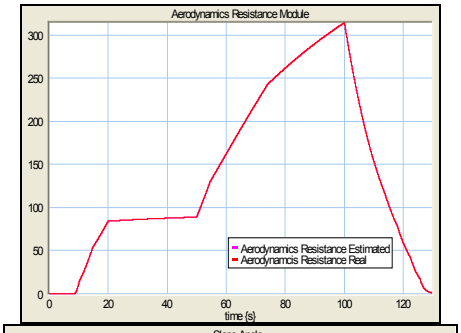


Figure 19: Friction Coefficient Estimated and with the Pacejka Model ; Figure 20: Rolling Resistance Estimated and Simulated

The 4th step on the observer model gives us the estimated rolling resistance and aerodynamics resistance force. Figure 26 shows the estimated and real Rolling resistance. Figure 27 shows the estimated and real Aerodynamics resistance force.



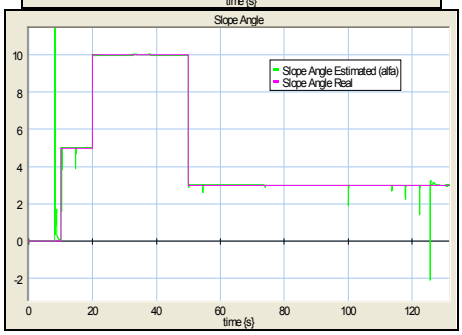


Figure 21: Aerodynamics Resistance Force Estimated and Real;
Figure 22: Slope Angle Estimated and Real

The 5th step on the observer model gives us the estimated slope angle.
Figure 28 shows the Slope Angle Estimated and Real proposed by table (1) to the vehicle model.

CONCLUSION

In this work we have developed a method to observe tire forces. The proposed estimation procedures are robust, and can then be used to improve the security by detecting some critical driving situation. This estimation can be used in several vehicles control system such Anti-lock Brake System (ABS), traction control system (TCS), diagnosis systems, etc. An observer was proposed to estimate the forces and friction coefficient. The estimations are produced using only the angular wheel position and longitudinal velocity as measurement and they are the input to the specially designed robust observer based on the Second Order Sliding Modes (SOSM). The method of estimation is verified through simulations using a contact model ,the "Pacejka Model" (Magic Formula) (H.B. Pacejka 1973).
This work presents a vehicle model composed by different modules. We have the chassis like rigid body, the suspension system, the pneumatics tire (Pacejka), transmission and engine. This model was constructed and simulated with the bond graph model in 20-Sim. Then we used the proposed efficient robust estimator, based on the second order sliding mode differentiator, to build an estimation scheme to identify the longitudinal and vertical forces, the friction coefficient, the tire rolling resistance and the slope angle. The estimation converging finite time produced allows us to obtain virtual measurements of model inputs, in five steps by cascade observers and estimators. Using the vehicle model we can compare the estimation variables with real variables (not measurable). These robust estimations on line are necessary for use on vehicle control dynamics. The observer and estimators were constructed and simulated in 20-Sim with the vehicle model.

This form of construction of models allows obtaining more finished and complex models, making possible to improve the definition of variables. It is possible to obtain estimations for different components specifically and individually, achieving a major precision in the estimation of variables. All this is possible with the relative simple construction of models. The plots obtained on the simulation in 20-Sim verify the correct works of the proposed robust observer for the estimation of variables.

REFERENCES

F. Aparicio Izquierdo, C. Vera Alvarez, V. Dias Lopez, 2004"*Teoria de los vehiculos automoviles*". U.P. de Madrid.

A. J. Blundell, 2002 "Bond Graphs for modeling Engineering Systems". *Department of Systems and Control Lanchester Polytechnic Coventry*

Cátedra DSF, 2006 "Introducción a la Modelización con Bond Graphs". Departamento de Electrónica – FCEIA – UNR.

G. Filippini, 2004 "Dinamica Vehicular mediante Bond Graph", Proyecto final de carrera de grado. Escuela de Ingenieria Mecanica, *Universidad Nacional de Rosario, Argentina*

L. Fridman. 2006 "Observadores en modos deslizantes", available : http://verona.fi-p.unam.mx/~lfridman/

T.D. Gillespie,2000 "*Fundamentals of Vehicle Dynamics*" - SAE.

I Hocine, 2003 "Observation d'états d'un véhicule pour l'estimation du profil dans les traces de roulement".

D.C.Karnopp, D.L.Margolis, R.C.Rosenberg, 2002 "Modeling and Simulation of Mechatronic System". *John Wiley & Sons, Inc.* - Third Edition.

J. Katz, "Race car Aerodynamics, Designing for Speed". *Bentley Publishers*

A. Levant, 1993 "Sliding order and sliding accuracy in sliding mode control," *International Journal of Control*, **vol. 58,** , pp 1247-1263

N.K. M'Sirdi, A. Naamane, A. Rabhi, 2007 "A nominal model for vehicle dynamics and estimation of input forces and tire friction". Marrakech CSC Nacer K. M'Sirdi.2003 Observateurs robustes et estimateurs pour l'estimation de la dynamique des véhicules et du contact pneu - route. *JAA. Bordeaux*, 5-6 Nov.

N.K. M'sirdi, A. Rabhi, N. Zbiri and Y. Delanne.2004 VRIM: Vehicle Road Interaction Modelling for Estimation of Contact Forces. TMVDA04. *3rd Int. Tyre Colloquium Tyre Models For Vehicle Dynamics Analysis* August 30-31, University of Technology Vienna, Austria.

H.B. Pacejka, “Tyre and Vehicle Dynamics”. *Delft University of Technology – Second Edition.*

. H.B.Pacejka, I.Besseling.1997 Magic Formula Tyre Model with Transient Properties. *2nd International Tyre Colloquium on Tyre Models for Vehicle Dynamic Analysis*, Berlin, Germany Swets and Zeitlinger.

H.B. Pacejka.1973: Simplified behavior of steady state turning behavior of motor vehicles, part 1: Handling diagrams and simple systems. *Vehicle System Dynamics* 2, , pp. 162 - 172.

H.B. Pacejka.1973: Simplified behavior of steady state turning behavior of motor vehicles, part 2: Stability of the steady state turn. *Vehicle System Dynamics* **2**, 1973, pp. 173 – 18

A. Rabhi, H. Imine, N.K. M’Sirdi, et Yves Delanne. 2004 Observers with Unknown Inputs to Estimate Contact Forces and Road Profile. pp 188-193. *AVCS’04. International Conference on Advances in Vehicle Control and Safety* Genova -Italy, October 28-31

G. Rill, 1995 “Vehicle Dynamics”. Fachhochschule *Regensburg University of Applied Sciences*, 2005.

W.F. Milliken, D.L. Milliken, *“Race Car Vehicle Dynamics”*. Society of Automotive Engineers, Inc.,

“Getting Started with 20-Sim 3.6”, *2005 Controllab Products* B.V., Enschede, Netherlands. Available: www.20sim.com,

AN RSVP MODEL FOR OPNET SIMULATOR WITH AN INTEGRATED QOS ARCHITECTURE

Sibel Tarıyan Özyer[(a)], Reza Hassanpour[(b)]

[(a)(b)]Department of Computer Engineering, Çankaya University, Ankara Turkey

[(a)]tariyan@cankaya.edu.tr, [(b)]reza@cankaya.edu.tr

ABSTRACT

Resource Reservation Protocol (RSVP) allows Internet real-time applications to request a specific end-to-end Quality of service (QoS) for data stream before they start transmitting data. In this paper firstly an overview of RSVP is presented. After that the different quality of services available and the relation between QoS and RSVP have been explained. The fundamentals of RSVP as a protocol is discussed. The performance issues and benchmarking for planned portion architecture at the department of Computer Engineering, Çankaya University has been given next. The experimental results and discussions conclude this paper. In this paper, OPNET network simulation tool has been used. Under given architecture and protocol, performance of quality of service implications has been carried out.

Keywords: RSVP, Quality of Service (QoS), Network flow.

1. INTRODUCTION

Internet allows the transmission of data between end points. In original design, it tries to transmit as quickly as possible but there is no guarantee to the timeliness and assurance of actual delivery. It provides its best effort service at end points. It may give qualitatively better service, but without the quantitative bounds of a guaranteed service it is far from expectations especially in an environment containing various services to be handled in the media. There is a great deal of interest in network applications. Accomplishment of best effort service for one single service is far from present day constraints. Due to demanding changes in end-point requirements, internet is affected to meet the quality of service (QoS) requirements (Tschudin 2001). There are several protocols for real time services (Video Conferencing, Internet TV and Internet Telephony, are rapidly growing and perfected) that support QoS of multimedia applications for IP networks such as Resource Reservation Protocol (RSVP), together with Real-Time Transport Protocol (RTP), Real-Time Control Protocol (RTCP), Real-Time Streaming Protocol (RTSP) (Moon and Aghvami 2001; Bouras 2007; Light et. Al. 2004), provides a working foundation for real-time services. Utilization of RSVP in a network with different perspectives is analyzed and simulations conducted are given in experiments. In this paper, OPNET network simulation tool (OPNET 2000) is used. OPNET is a network simulation tool that outputs the characteristics of a real time network utilizing different services with a priori parameters. Under given architecture and protocol, performance of quality of service implications is carried out.

There is no single technique provides efficient, dependable QoS in an optimum way. Instead a variety of techniques have been developed, with practical solutions often combining multiple techniques. Some of the techniques used to achieve QoS are: Over provisioning, Buffering, Traffic Shaping, the Leaky Bucket Algorithm, Token Bucket Algorithm, and Resource Reservation (Tanenbaum 2003).

RSVP (Resource Reservation Protocol) is a resource reservation setup protocol for the Internet. The RSVP protocol is used by hosts to obtain specific qualities of service from the network for particular application data streams or flows. It is also used by routers to deliver quality-of-service (QoS) requests to all nodes along the path of the flows and to establish and maintain state to provide the requested service (Rfc 2205). RSVP carries the request through the network, visiting each node the network uses to carry the stream. At each node, RSVP attempts to make a resource reservation for the stream. Some applications require reliable delivery of data but do not impose any stringent requirements for the timeliness of delivery. But applications such as video- conferencing, IP telephony, NetRadio require almost exact opposite: Data delivery must be timely but not necessarily reliable. Thus, RSVP was intended to provide IP networks with the capability to support the divergent performance requirements of differing application types (Karsten et. al. 2001).

2. DETAILED BACKGROUND OF RSVP SYSTEM

The RSVP protocol performs a reservation for each flow requiring QoS services; a flow is defined by five tuples (source IP address, destination IP address, transport protocol, source port, and destination port). Each flow needs several RSVP messages, to request, maintain and release the required resources.

With an RSVP based quality of service architecture there are two basic elements: sources and destinations, all of them run RSVP daemons that participate in RSVP protocol and exchange RSVP

messages on behalf of their hosts. They exchange basically two types of messages: PATH and RESV. The RSVP source sends a PATH message which is encapsulated in IP or UDP datagrams (Stallings 2004). The message travels through the network to the destination. When it is received by the destination, if it wants to make a reservation for the particular RSVP flow, it responds with a RESV message and it traverses the reverse path back to the sender. Otherwise, a RESV ERROR message is issued and is sent back to the receiver. An end-to-end reservation is successfully established when the RESV message reaches the sender and is successfully processed by the RSVP daemon on the sender and in all the other nodes in the middle.

A multicast reservation session can also be made. In this case the sender sends the PATH messages to a multicast group address. As in the case of unicast, the path messages travels through the network to all the members of the multicast group. When PATH messages reach the receivers, each receiver independently decides if it wants to request a reservation for the session. Each receiver can potentially request for different reservations for the same session (Barzilai et. al. 1996). Figure1 shows an example, where S1 and S2 are sources, and D1, D2 and D3 are destinations of data. D1, D2 and D3 are members of the same multicast group and S1 and S2 send messages to this multicast group. But we can see in the figure that not all the destinations make the same reservation. D3 sends RESV messages to S1 and S2, so it accepts all the reservations. D2 only accepts one of the reservation requests sending one RESV message to S1. But D1 does not want to make any reservation and does not send any RESV message to S1 and S2.It has to be clear that RSVP is a protocol to negotiate a quality of service for a specific application and it is not a routing protocol. It uses the routing table in routers to determine routes to the appropriate destinations. So it was designed to interoperate with existing unicast and multicast IP routing protocols (3Mallofre 2003).

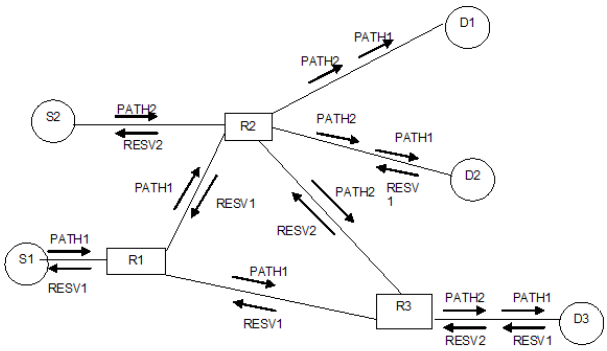


Figure 1: PATH and RESV messages flows in RSVP (Barzilai et. al. 1996)

2.1. Reservation Styles

Reservation style indicates to the network element that an aggregation of reservation request is possible for a multicast group. Resource reservation controls how much bandwidth is reserved, whereas reservation filter determines the packets that can make use of this reservation. RSVP supports three styles of reservation. A description of these styles is provided in the following subsections. If we have different senders for the same RSVP session, then we have two modes:

- Distinct Reservation: creates a different reservation for each upstream sender.
- Shared Reservation: creates a shared reservation for specified senders.
- But we have another option that controls the set of senders. With this option there are also two options:
- Explicit: select a list of the senders.
- Wildcard: selects all the senders for the session. And now if we mix these modes, then as we can see in Figure 2, there exist the Fixed-Filter style (FF), the Shared-Explicit style (SE), and the Wildcard-Filter (WF):

Sender Selection	Reservations: Distinct	Shared
Explicit	Fixed-Filter (FF) style	Shared-Explicit (SE) Style
Wildcard	(None defined)	Wildcard-Filter (WF) Style

Figure 2: Reservation attributes and styles (Rfc 2205)

The Wildcard-Filter (WF) style creates a single reservation shared by all flows from all upstream senders. The Shared-Explicit style (SE) creates a single reservation shared by selected upstream senders, so is the same than the Wildcard-Filter but with not all the senders. And the last style is the Fixed-Filter, which creates a distinct reservation for data packets from a particular sender. This is the last style because there is no defined style for a distinct reservation in a Wildcard sender selection. WF and SE are appropriate for multicast applications in which multiple data sources are unlikely to transmit simultaneously.

2.2. Incorporating RSVP with OPNET

OPNET allows for a very large number of potential statistics. For this reason, collection mechanisms are deactivated by default when a simulation is executed. However, OPNET provides a mechanism to explicitly activate statistics of particular interest, which are recorded in appropriate output files. This is accomplished by specifying a list of probes when running a simulation, which indicate the particular statistic that should be collected.

In order to investigate the RSVP in the network environment, the following statistics have been specified and studied that allowed for the system's behavioral study and validation.

- Link Delays, Throughput and Utilization (for identifying any congested links).
- RSVP Control Traffic sent and received (for verifying the protocol's correctness of functionality).
- Application Traffic sent and received (for verifying the correct traffic flow through the system's components).

2.2.1. The Scenario

Utilization of RSVP in a network with different perspectives is analyzed and simulations conducted are given in experiments for future network topology planned part at the department of Computer Engineering, Cankaya University. In this paper, OPNET network simulation tool is used. OPNET is a network simulation tool that outputs the characteristics of a real time network utilizing different services with a priori parameters. Under given architecture and protocol, performance of quality of service implications is carried out.

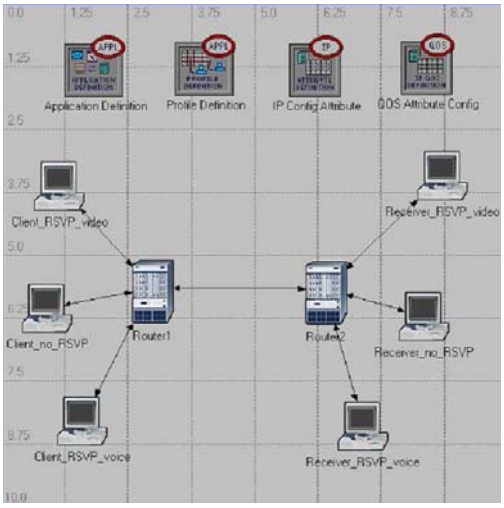


Figure 3: Rsvp enabled network

In order to evaluate the performance of the RSVP model, this network model has been created (see Figure 3).The Ethernet network consists of three clients sending traffic to associated receivers via routers. All the nodes in the network are connected with PPP DS0 links with a 64 Kbps data rate. Client_RSVP_video node and Client_no_RSVP video node are video applications while Client_ RSVP_voice node is voice application. Two video conferencing sessions are competing for the same resources. Traffic between Client RSVP video and Receiver RSVP video uses RSVP to reserve resources also traffic between Client_RSVP_voice and Receiver_RSVP_voice uses RSVP to reserve resources while traffic between Client_no_RSVP and Receiver_no_RSVP uses best effort service. There is one session for each source-destination pair in-place for the duration of the simulation. The reservation will be made for traffic in both directions. The traffic generated by each client is described as having a bandwidth of 5,000 bytes/sec and a buffer size of 5,000 bytes. These parameters will be used for the reservation. Also reservation style is selected Wild Card. OPNET supports five different QoS policies: RSVP Protocol, Committed Access Rate, (CAR), Custom Queuing (CO), Priority Queuing (PQ), and Weighted Fair Queuing (WFQ). In addition OPNET's RSVP model supports Controlled Load service. This service is supported for WFQ and Custom Queuing schemes. This scenario based on WFQ and RSVP.

Configuring Applications Attributes describing RSVP parameters set by the application are defined in two objects: the QoS Attribute Configuration object and the Application Attribute Configuration object. To run an RSVP simulation, both objects must be included in the scenario. Also Profile Definition Attributes and IP Configuration Attributes can be seen for the general network configuration. Figure 4 illustrates the Profile definition Attributes are video reserved, Video unreserved and voice reserved which are defined in Profile Definition. Figure 5 shows the IP Configuration Attributes with default values.

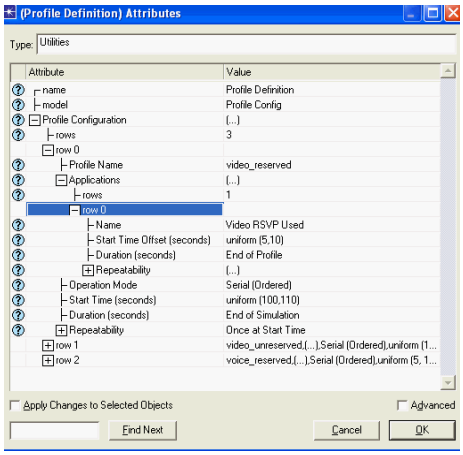


Figure 4: Profile Definition Attributes

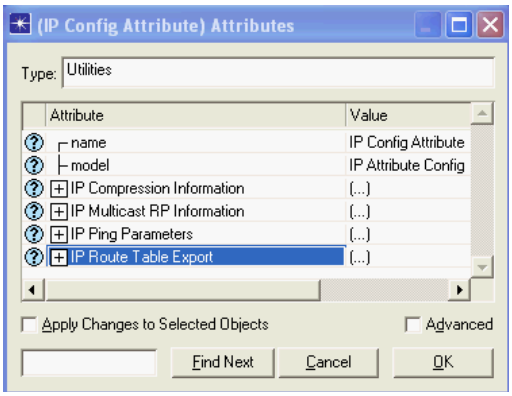


Figure 5: IP config Attribute

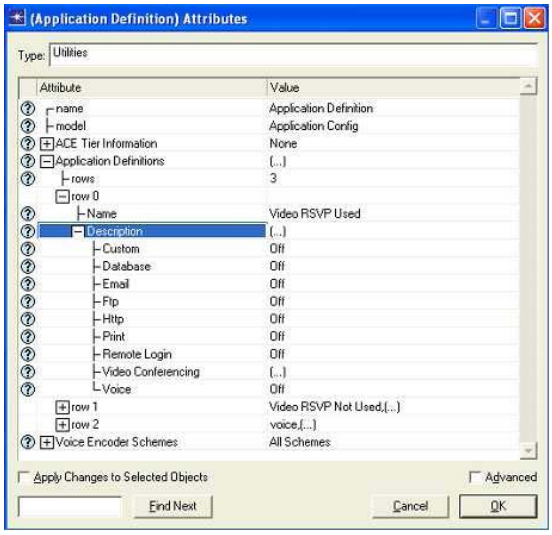


Figure 6: Application definition attributes

Figure 6 shows the Video conferencing and voice applications are defined in Application Definition. For Video Applications, video conferencing status is "On" and other applications are "Off". For voice Application, voice status is "On" and other applications are "Off".

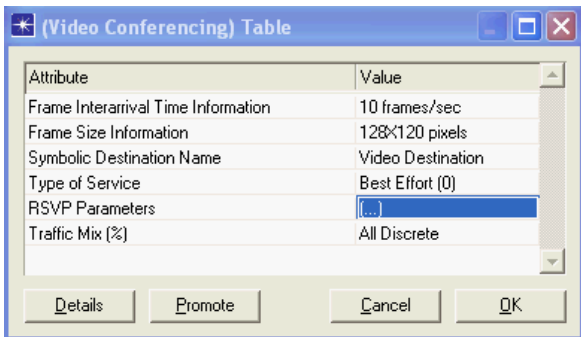


Figure 7: Video conferencing table

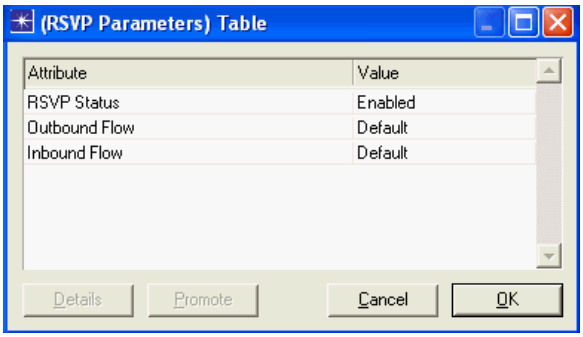


Figure 8: RSVP parameters table

The following Figures 7 and 8 show the RSVP parameters on Video conferencing application and RSVP status of this application.

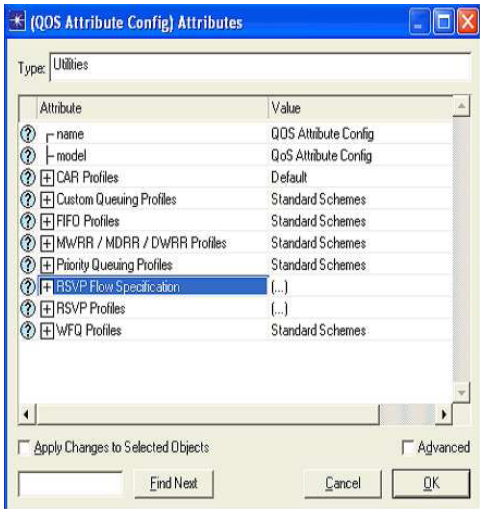


Figure 9: QoS attribute config attributes

Figure 9, Figure 10, and Figure 11 illustrate the RSVP flow specification and RSVP profiles. The traffic generated by each client is described as having a bandwidth of 5,000 bytes/sec and a buffer size of 5,000 bytes. These parameters will be used for the reservation. Also reservation style is selected Wild Card.

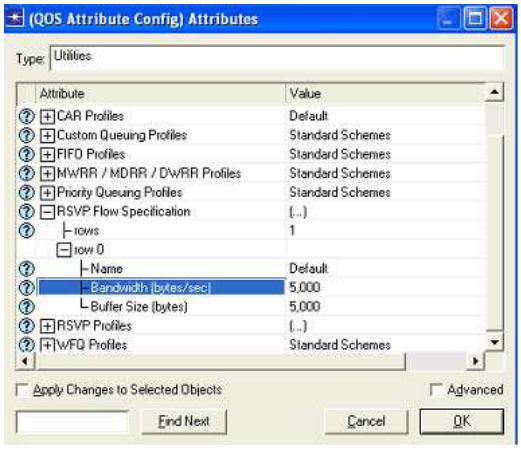


Figure 10: QoS flow spec atribute

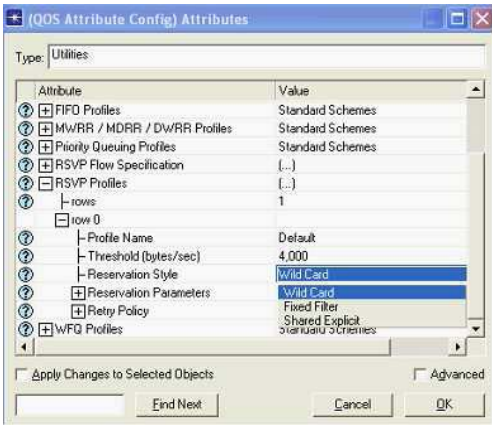


Figure 11: QoS RSVP reservation style

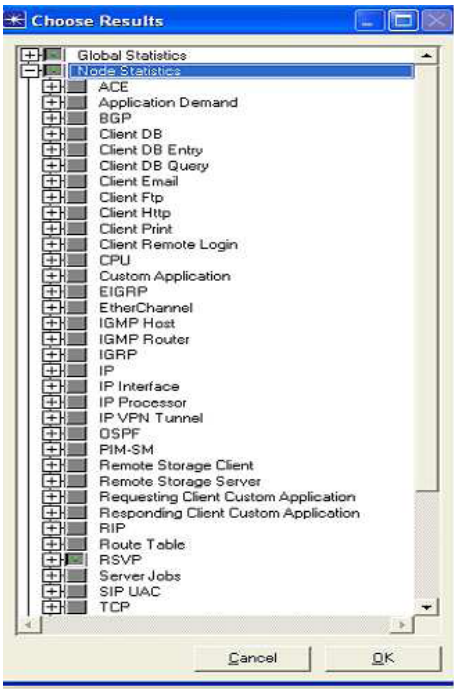


Figure 12: Statistics collection

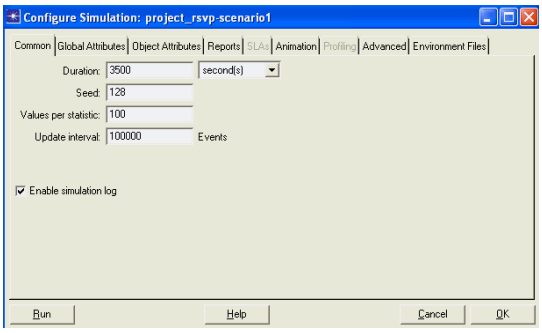


Figure 13: Simulation runtime settings

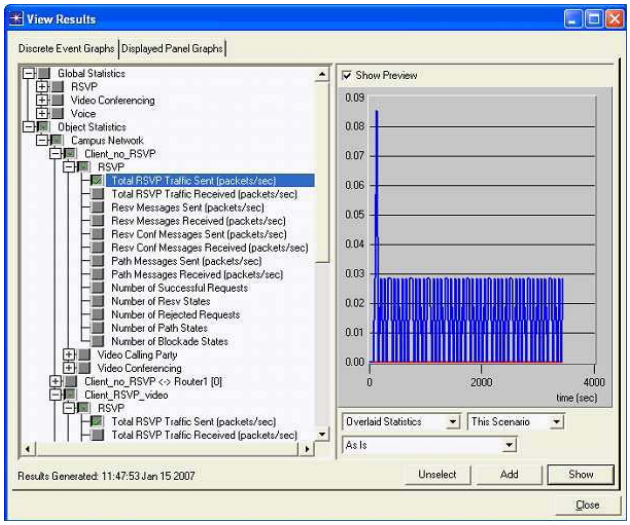


Figure 14: Selected simulation results

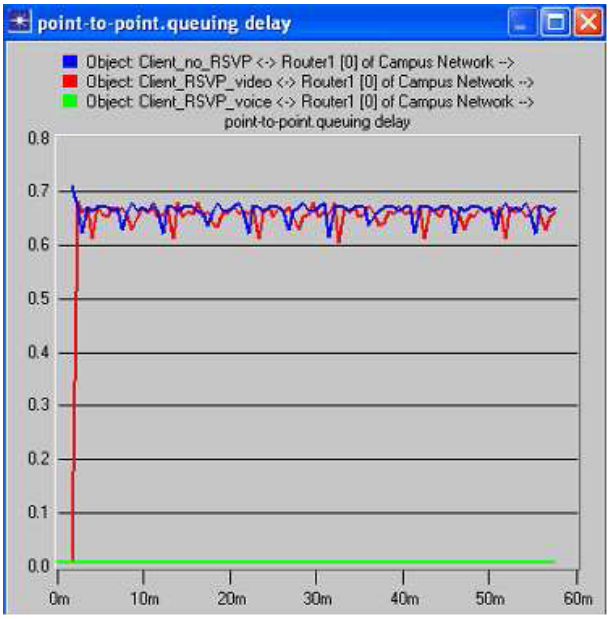


Figure 15: P2P queuing delay

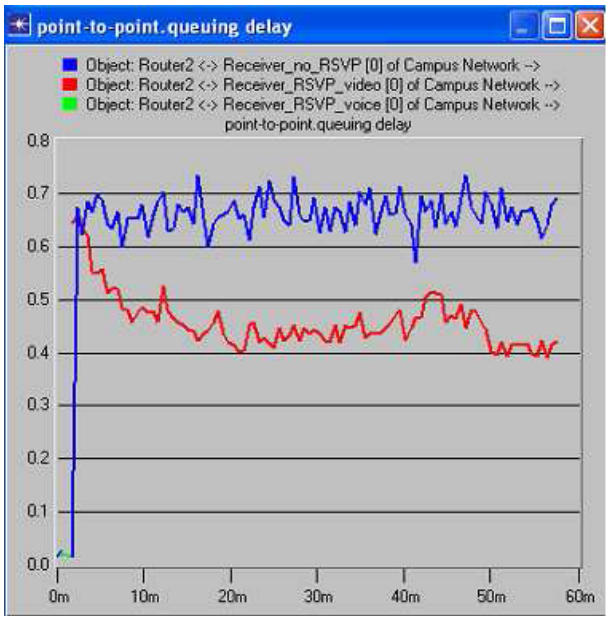


Figure 16: P2P queuing delay

Figure 15 and Figure 16 compare the point-to-point queuing delay experienced using RSVP with the queuing delay experienced not using RSVP between the Clients to Router1 and Router2 to Receivers. As expected, traffic using RSVP reservation experienced less queuing delay.

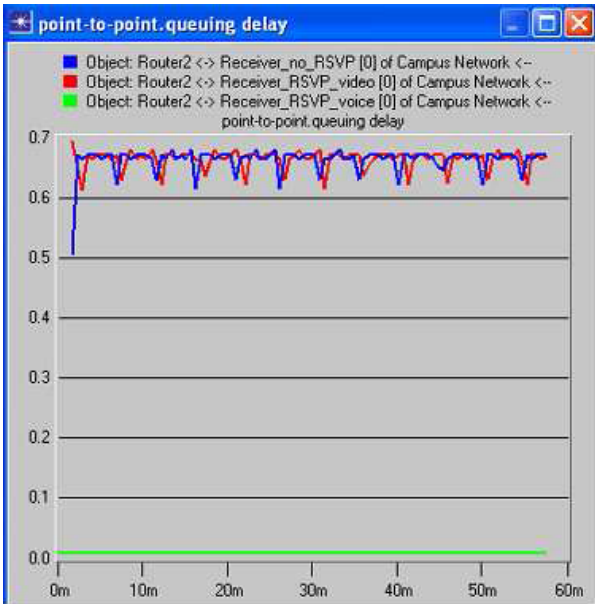


Figure 17: P2P queuing delay

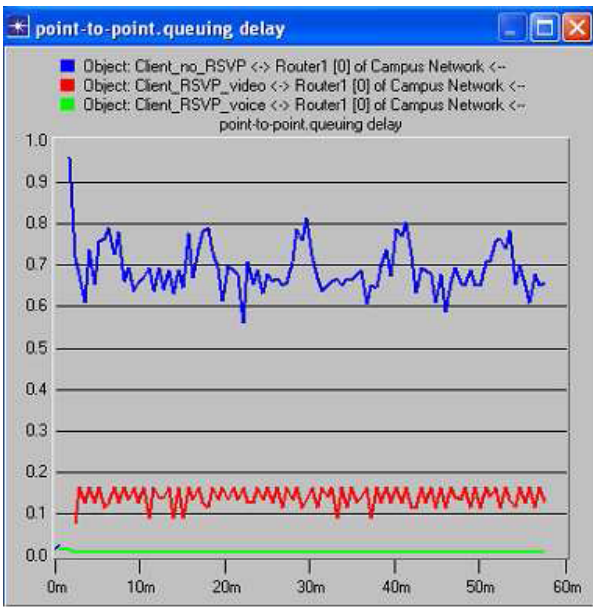


Figure 18: P2P queuing delay

Figure 17 and Figure 18 compare the point-to-point queuing delay experienced using RSVP with the queuing delay experienced not using RSVP between the Receivers to Router2 and Router1 to Clients. As expected, traffic using RSVP reservation experienced less queuing delay.

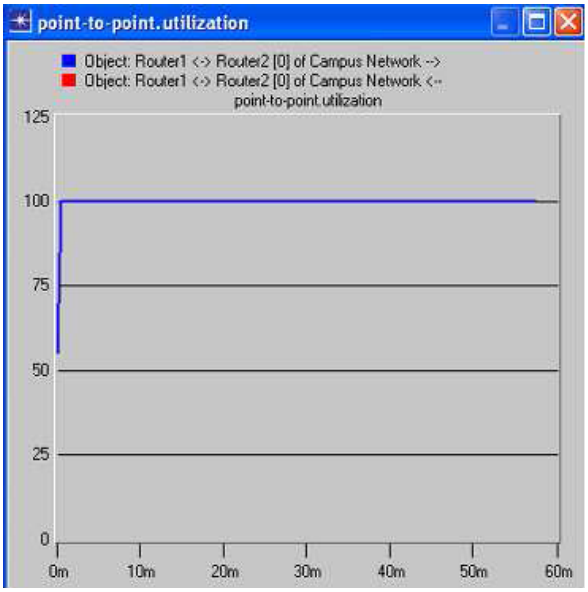


Figure 19: Link utilization

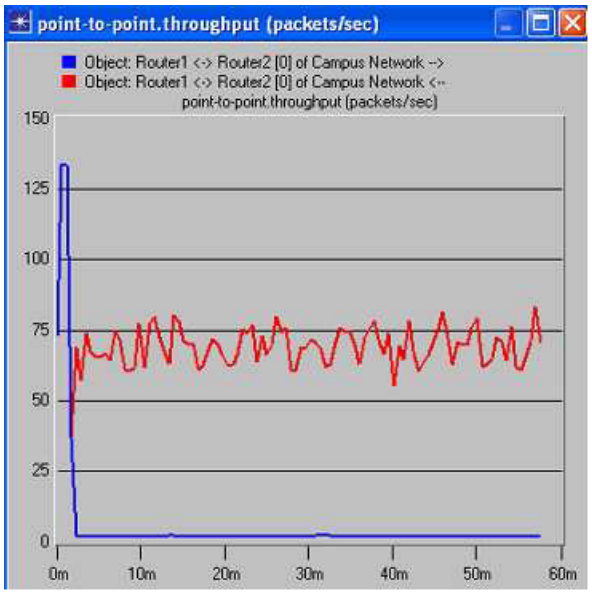


Figure 20: Link utilization

In Figure 19, utilization of the link between Router1 and Router2 is shown. RSVP is used between Router1 and Router2. Maximum Reservable Bandwidth is a percentage of the link bandwidth that RSVP can use. Maximum Reservable Bandwidth is configured to 100%. As expected, traffic using RSVP reservation experienced using full bandwidth. In Figure 20, throughput of the link between Router1 and Router2 is shown. It clearly shows how sufficient the link between Router 1 and Router 2 is for this load.

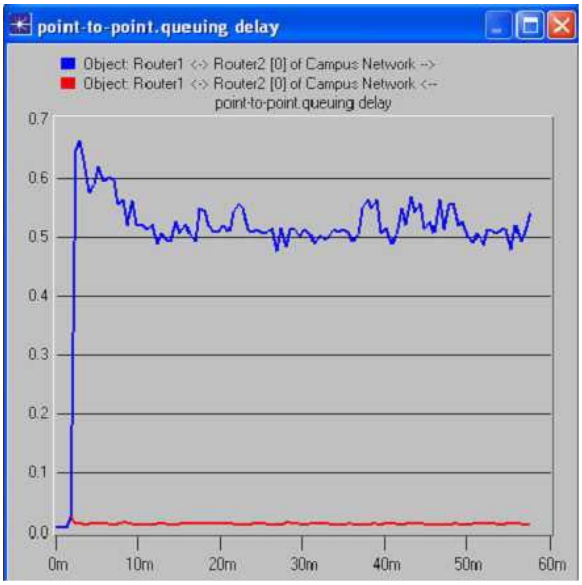


Figure 21: Link Queuing Delay

In Figure 21 queuing delay of the link between Router1 and Router2 is shown. RSVP is used between Router1 and Router2. In our scenario all clients are connected to Router1. Therefore Router1 sends more confirmation messages. As expected, the outgoing link between Router1 and Router2 experienced higher queuing delay.

Figure 22 and Figure 23 show the Total RSVP Traffic Sent and Received on Router1 and Router2 experienced using RSVP. As expected, the network is used by Router2 all the time so there is no too much change. The Router1 uses the network some times. Therefore fast change can be seen.

These two figures (Figure 24 and 25) show the Total RSVP Resv Messages Sent and Received on Router1 and Router2 experienced using RSVP. As expected, the network is used by Router2 all the time so there is no too much change. The Router1 uses the network some times. Therefore fast change can be seen.

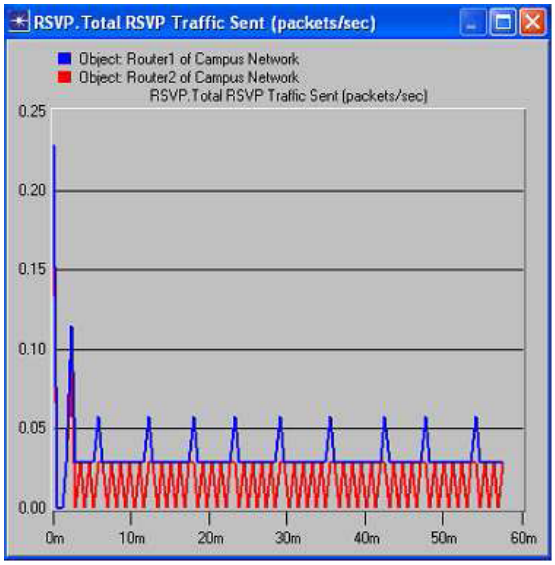


Figure 22: RSVP traffic sent

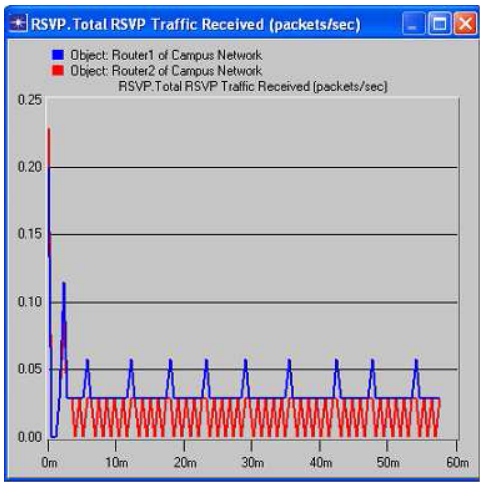


Figure 23: RSVP traffic received

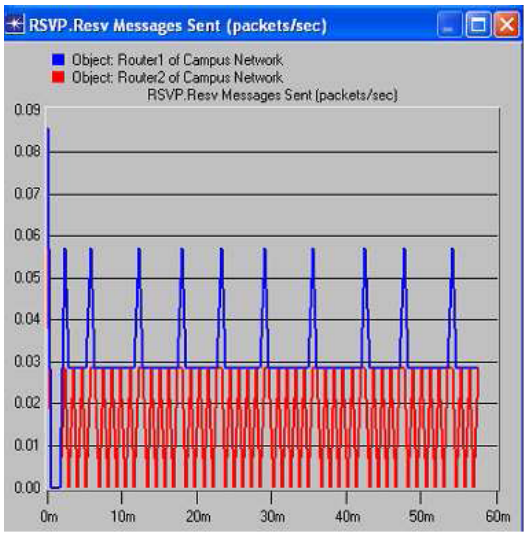


Figure 24: RSVP traffic sent

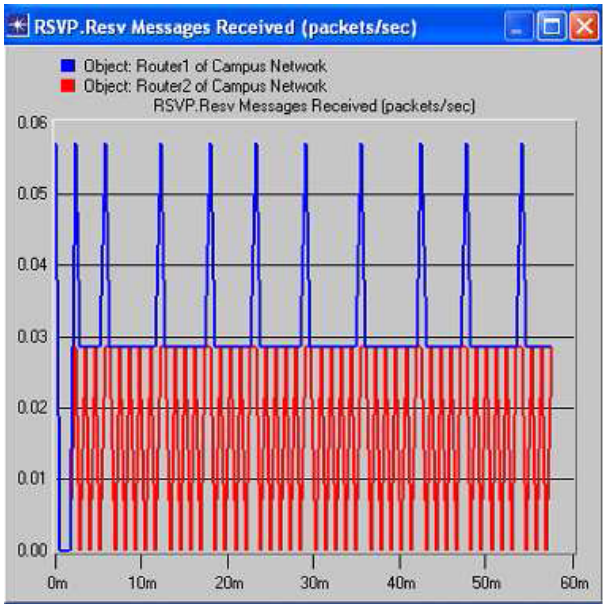


Figure 25: RSVP traffic received

3. CONCLUSIONS

We conducted some experiments by using OPNET IT Guru Academic Edition 9.1. A scenario is established. Clients and receivers with/without RSVP are used in our simulation. Some results are obtained link utilization, throughput, point to point delay, queuing delay, IP traffic are measured. RSVP protocol does allocation of bandwidth before transmission. If allocation is not done, data transmission does not occur.

In summary RSVP has the following attributes (Rfc 2205):

- RSVP makes resource reservations for unicast and multicast applications.
- RSVP sessions are simplex. Thus, a bidirectional exchange of data between a pair of machines actually constitutes two separate RSVP simplex sessions.
- RSVP is receiver-oriented. The receiver of a data flow initiates and maintains the resource reservation used for that flow.
- RSVP maintains soft state in routers and hosts, providing graceful support for dynamic membership changes and automatic adaptation to routing changes.
- RSVP is not a routing protocol but depends upon present and future routing protocols. RSVP transports and maintains traffic control and policy control parameters that are opaque to RSVP.
- RSVP provides several reservation models or styles to fit a variety of applications.
- RSVP provides transparent operation through routers that do not support it.
- RSVP supports both IPv4 and IPv6.

REFERENCES

Barzilai, T., Kandlur, D., Mehra, A. and Saha, D., 1996. Design and implementation of An RSVP Based Quality of Service Architecture for Integrated Services Internet, *International Conference on Distributed Computing Systems*.

Bouras, C., Gkamas, A., Primpas, D. and Stamos, K., 2007. Real-Time Protocols RTP/RTCP, *Encyclopedia of Internet Technologies and Applications, IDEA Group publishing*

Karsten, M., Schmitt, J. and Steinmetz, R., 2001. Implementation and Evaluation Of The Kom Rsvp Engine, *IEEE Infocom*.

Light, J. and Bhuvaneshwari, A., 2004. Performance Analysis of Audio Codecs over Real-Time Transmission Protocol (RTP) for Voice Services over Internet Protocol, *Communication Networks and Services Research*

Mallofre, G. R., 2003. Resource Reservation Protocol (RSVP), *Seminar on Transport of Multimedia Streams in Wireless Internet, Department of Computer Science University of Helsinki, Finland*.

Moon, B., Aghvami, H., 2001. RSVP Extensions for Real-Time Services in Wireless Mobile Networks, *Communications Magazine, IEEE*, Volume 39, Issue 12, pp. 52 - 59.

OPNET, 2000. Application Model Description, *OPNET General Models Manual*, pp. 1-16.

RFC 2205 http://www.ietf.org/rfc2205.txt.

Stallings, W., 2004. Computer Networking with Internet Protocols and Technology, *Prentice Hall, NJ*.

Tanenbaum, A., 2003. Computer Networks, *Prentice Hall 4th Edition*

Tschudin, C., 2001. Overiew: Qos and Rsvp, *Data Networks II*, Volume 15, pp. 1–29.

Zhang, L., Deering, S., Estrin, D. and Shenker, S., 2002. RSVP: A New Resource ReSerVation Protocol, *IEEE Communications Magazine* 40:5, pp. 116-127.

AUTHORS BIOGRAPHY

Reza Hassanpour received B.S and M.S degrees in Computer Engineering from Shiraz University and Tehran Polytechnic University in 1995 and 1998 respectively. He received his Ph.D.degree in Computer Engineering from the Middle East Technical University in 2003. Currently, he is assistant professor at Computer Engineering Department of Çankaya University, Ankara, Turkey. His research interests are 3D Computer Vision, 3D Modeling and Visualization.

Sibel Tarıyan Özyer received his BSc. and MSc. degrees from the Computer Engineering Department of Çankaya University, Ankara, Turkey in 2004 and 2007. Currently, she is PhD candidate at the Modeling and Design of Engineering Systems Department of Atılım University, Ankara, Turkey. She is also working as a Research Assistant at the Computer Engineering

Department of Çankaya University. Her research interests are computer networks, and wireless sensor networks.

ELABORATION OF A COMPOSITE SCORE FOR DESIGNING AND CHOOSING SUSTAINABLE "OFFICIAL CAR POLICY"

Rigo Nicolas [(a)], Ndiaye Alassane Ballé [(b)]

[(a) (b)] QALINCA - Chair of Innovation for Quality, Logistics and Supply Chain Management, Department BEAMS, Faculty of Engineering Sciences, University of Brussels, Belgium,

[(a)] nicolas.rigo@ulb.ac.be, [(b)]abndiaye@ulb.ac.be

ABSTRACT

A lot of companies provide their employees with official cars in order to get some fiscal advantages while simultaneously satisfying the employee himself. It is a common "win-to-win" situation in most of the cases. However, these official car policies can lead to numerous impacts such as road congestion increase, air pollution as well as climate change. Therefore, it is of key importance to be able to manage that situation in a sustainable way. It is the aim of this paper to develop a methodology for designing and choosing sustainable alternatives in matter of official car policy management. The methodology is based on the elaboration of composite scores integrating the environmental footprints, the costs and the social impacts of proposed alternatives. These scores allow prioritizing official car policies while simultaneous sensitivity analyses make it possible to provide sustainable recommendations.

Keywords: official car policy, sustainable management, composite score.

1. INTRODUCTION

When observing the ways of managing commuting of employees, it is obvious that the "official car policies" are very widespread among local, national or international companies. In one hand, it is a fiscal advantage for the company while in the other hand it represents substantial expenses that the employee does not have to take in charge for the use of a private vehicle. It has to be noted that 50% of annual car registrations are made by companies in Belgium. However, the negative impacts of such a situation have to be highlighted: emissions of toxic substances, impacts on climate change, increase of road congestion, etc.

So, it is essential to be able to manage and control "official car policies" in order to minimize the above mentioned impacts.

It is the aim of this paper to propose the elaboration of a composite score to help companies for designing and choosing sustainable official car policies; in other words, cost-, social-, and environmental-friendly policies.

In Section 2, the authors review a number of books, articles, scientific papers and European Directives for elaborating the proposed composite score.

Section 3 describes the structure of the composite score and the underlying aggregation methods. First, major pollutants compose the environmental footprint made up of carbon dioxide, carbon monoxide, sulfur oxides, particle matters and nitrogen oxides. Based on these pollutants, the emissions of current official car policies and alternatives can be evaluated and aggregated in an environmental footprint. Secondly, the global cost is made up of the leasing and fuel costs for official cars while only subscription fees are considered in case of public transport use. Thirdly, the social impact is elaborated on the basis of a qualitative indicator: the comfort, while a second impact is quantified: the travel time. Then, the three above mentioned "super indicators" are aggregated into a final composite score expressing the sustainable performance of current official car policies and alternatives. Based on this score, the various scenarios can be compared and prioritized. Moreover, sensitivity analyses can be performed for testing the robustness of the recommended solutions.

The applicability of this methodology is demonstrated in Section 4 by analyzing and solving a case study.

Finally, some conclusions and perspectives of development are presented in Section 5.

2. LITERATURE REVIEW

Number of publications, articles, European directives and other reports were essential for developing the proposed methodology and the study case.

For the elaboration of the composite score, the major literature source is (Rigo et al. 2008) presented during the MAS conference 2008 where the authors developed a similar approach. This approach was improved and adapted to the problem of "official car policy". Moreover, the authors reviewed number of other references in order to validate the methodology used, such as (Roy and Bouyssou 1993); (Roy 1985); (Rigo, Ndiaye, Dreyer, Zomer, Pinon and Tremeac 2007); (Brans and Mareshal 2005).

For the elaboration of the database used for applying the proposed methodology and solving the study case, (Van Essen et al. 2003) proposes a lot of useful "top-down" data regarding the environmental performances of freight and passenger transport for different modes.

All in all, this review led to the elaboration of the composite score and the demonstration of its applicability in practical situations.

3. METHODOLOGY

This section develops in details the scheme used for the elaboration of the sustainable composite score.

This composite score is made up of various types of indicators grouped into three categories respectively linked to the pillars of the sustainable development. For each category, a list of indicators is elaborated. A first aggregation makes it possible to calculate the environmental footprint, the global cost and the social impact via three scores integrating all the information contained in the "first level" indicators.

These "super indicators" can be aggregated into a sustainable composite score providing the actual sustainable performance of the analyzed official car policy and related possible altenatives.

Since the authors use the PROMETHEE method for elaborating the "super" environmental and social indicator as well as the sustainable composite score, the next section recalls the main steps of this pair wise based multi criteria decision aiding method.

Then, the ways of calculating the impacts on the indicators are detailed as well as the methods used to aggregate them in super indicators.

Finally, the elaboration of the sustainable composite score is explained.

3.1. The PROMETHEE methodology

Since the PROMETHEE method is used for the aggregation of the environmental and social impacts as well as for the elaboration of the composite score, the following modeling is recalled.

First, let us consider a set of criteria, $\{g_1(.), g_2(.), g_3(.),..., g_m(.)\}$ and a set of scenarios to compare $A = \{a_1, a_2, a_3,..., a_n\}$. Let us define $g_j(a_i)$ the evaluation of scenario a_i on the axis *j*.

Let us consider the deviation of impacts of two actions on a criterion:

$$d_j(a,b) = g_j(a) - g_j(b); \quad \forall a,b \in A \tag{1}$$

In order to delete the possible scale effects related to the units of criteria, let us define the following function in the case of a criterion *j* to maximize,

$$0 < F_j[d_j(a,b)] \le 1; \quad \forall a,b \in A \tag{2}$$

Where:

$$F_j[d_j(a,b)] > 0 \Rightarrow \; F_j[d_j(b,a)] = 0; \quad \forall a,b \in A \tag{3}$$

If the criterion *j* has to be minimized, the following relation is considered,

$$0 < F_j[-d_j(a,b)] \le 1; \quad \forall a,b \in A \tag{4}$$

The pair $\{g_j(.); F_j[d_j(a,b)]\}$ is called the generalized criterion associated to the criterion g_j or the preference function related to the criterion g_j.

Various preference functions are available and can be varied to an infinite number of solutions corresponding to the needs of the users. Two examples are described here after:

The usual generalized criterion is defined as follows:

$$F_j[d_j(a,b)] = \begin{cases} 0 & if \quad d_j(a,b) \le 0 \\ 1 & if \quad d_j(a,b) > 0 \end{cases} \; ; \; \forall a,b \in A \tag{5}$$

The U-shape generalized criterion can be defined as follows:

$$F_j[d_j(a,b)] = \begin{cases} 0 & if \quad d_j(a,b) \le q \\ 1 & if \quad d_j(a,b) > q \end{cases} \; ; \; \forall a,b \in A \tag{6}$$

Where *q* is the preference threshold.

On the basis of these generalized criteria, let us calculate the aggregated preference indices as follow:

$$\begin{cases} \pi(a,b) = \sum_{j=1}^{m} F_j[d_j(a,b)] \times \omega_j \\ \pi(b,a) = \sum_{j=1}^{m} F_j[d_j(b,a)] \times \omega_j \end{cases} \; ; \; \forall a,b \in A \tag{7}$$

Where w_j is the weight allotted to the criterion *j*. Then, the positive and negative outranking flows are calculated as follows:

$$\phi^+(a) = \frac{1}{n-1} \sum_{x \in A} \pi(a,x) \tag{8}$$

$$\phi^-(a) = \frac{1}{n-1} \sum_{x \in A} \pi(x,a) \tag{9}$$

The PROMETHEE II complete ranking is based on the following outranking net flow:

$$\phi(a) = \phi^+(a) - \phi^-(a) \tag{10}$$

3.2. Towards the "super indicators"

Three main fields have to be considered in order to develop a sustainable composite score: the economic aspects focusing on the direct costs for the company, the environmental impacts dedicated to the pollutant

emissions and the social aspects highlighting the well-being of the employee during the journey.

3.2.1. The environmental indicators

When looking at the environmental aspects, the authors focus on the air pollution and selected a list of six pollutants usually considered in the frame of transport activities. The "Well to Wheel" framework was chosen in order to depict a global picture of the emissions caused by transport of persons. This means that both "Well to Tank" (*WTT*) and "Tank to Wheels" (*TTW*) emissions are considered.

Among these chemical substances, the carbon dioxide (CO_2) resulting from fuel combustion has a global impact on climate change and is certainly the major greenhouse effect gas. The nitrogen oxides (NO_x) are generated by 'high-temperature' combustions and contribute to the creation of ozone impacting on human health and vegetation. The sulfur oxides (SO_x) and among others, the sulfur dioxide originates in the sulfur contained in the fuel. SO_2 leads to acidification impacting on the public health and the crops. The particle matters (PM) such as soot and ashes, present in oil causes local pollution, impacting seriously on human health. The carbon monoxide (CO) results from an incomplete combustion of fuel and causes a local acidification and the creation of ozone on the ground level.

Then, it is the aim of this paper to evaluate the impacts of "official car policies" on these emissions. Since most of these data are not available from the companies, a top down approach is used in order to provide realistic estimations. So, in one hand data are available for type vehicles: the energy consumption expressed in MJ/km and the emission factors per vehicle in g/km are available for the above-mentioned pollutants. In the other hand, data are also available for refining and electricity production emissions and expressed in g/MJ.

Therefore, "Well to Wheels" emissions can be quantified for official car policies and designed alternatives leading to a detailed estimation of the toxic pollutants emitted in the atmosphere.

Five transport modes are considered in the frame of this paper; namely the cars, buses, trams, metros and trains. If we look at the "Tank to Wheels" emissions, only cars and buses are concerned since trains, metros and trams only lead to electricity production emissions.

So, let us consider the following factors:

- $EF_{p,HW,v}$, the emission factor of pollutant *p* by a vehicle *v* of the fleet *V* on highway section expressed in g/km;
- $KM_{HW,v}$, the distance covered on highways by the vehicle *v*;
- $EF_{p,CR,v}$, the emission factor of pollutant *p* on motorway section in g/km;
- $KM_{CR,v}$, the distance covered on city roads by the vehicle *v*;

In this paper, the authors do not consider car pooling scenarios.

So, the following formulas provide the *TTW* emissions of pollutant *p* in *g* for the fleet *V* on highway and city road sections per day:

$$TTW_{p,V,HW} = \sum_{v \in V} (EF_{p,HW,v} * KM_{HW,v}) \quad (11)$$

$$TTW_{p,V,CR} = \sum_{v \in V} (EF_{p,CR,v} * KM_{CR,v}) \quad (12)$$

The authors want to emphasize the rural and urban area and the related pollution impacts. That is why the *TTW* emissions are split between highway and city road sections making it possible to use specific urban and non urban societal costs when elaborating the environmental footprint, as explained at the end of this section.

When applying these equations in a practical case, a specific time window has to be fixed for quantifying the emissions. In the study case developed in this paper, the authors proposed to consider 22 working days equivalent to one month. Therefore, the *TTW* emissions will be multiplied by 22 since the daily covered distance is supposed to be the same.

When looking at the *WTT* emissions, let us consider the following factors for diesel engine vehicles:

- ER_p , the emission of refining pollutant *p* in g/MJ ;
- $EU_{p,CR,v}$, the energy use in MJ/km for a vehicle *v* on city roads;
- $EU_{p,HW,v}$, the energy use in MJ/km for a vehicle *v* on highways.

Then the next equation provides the *WTT* emissions of a pollutant *p* for diesel engine vehicles of the fleet *V*:

$$WTT_{p,V} = \sum_{v \in V} (EU_{p,HW,v} * KM_{HW,v} + EU_{p,CR,v} * KM_{CR,v}) * ER_p \quad (13)$$

When evaluating the *WTT* emissions, the authors considered that it is very difficult to emphasize the crossed urban and non urban area since the toxic substances are emitted during the feedstock extraction, storage, distribution and transport of the fuel, tasks taking place far away from the actual transport route used by the employees.

When looking at *WTT* emissions, electric modes have to be considered too. The authors proposed to exclude the nuclear power source of electricity production since the natural human behavior reluctantly

accepts the high risks and waste aspects linked to that electricity source.

Then, let us consider the following factors:

- $EP_{p,v}$, the emission factor of pollutant *p* by a vehicle *v* expressed in g/km;
- ER, the energy return more or less equal to 38% regarding the electricity production and distribution rates of 42% and 90%;
- EU_v, the energy use for a vehicle *v* expressed in MJ/SKM (Seat Kilometers because of the consideration of electric public transport);
- KM_v, the total distance covered by the vehicle *v*.

So, the *WTT* emissions of the electric public transport used for one seat *s* can be expressed as follows:

$$WTT_{p,s} = \sum_{v \in V} EP_{p,v} * ER * EU_v * KM_v \qquad (14)$$

Then, by multiplying this figure by the number of seats occupied by the new employees using public transport, we obtain the marginal emission due to the public transport use.

The assessment of the abovementioned indicators for each pollutant makes it possible to elaborate the evaluation table used to calculate the environmental footprint. This evaluation table is based on the results of equations (11), (12), (13) and (14).

Since the authors used the PROMETHEE method for aggregating these impacts, and as explained in Section 3.1, weights have to be defined.

The authors propose to use the societal costs in urban and non urban sections for the *TTW* emissions while the average of the urban and non urban costs are considered for the *WTT* emissions.

It has to be noted that no urban/rural distinction is available for CO_2 and SO_2.

Table 1: Societal costs of toxic substances

Substance	Societal cost (E/ton)
CO_2	50
SO_2	3000
NOx (urban)	7000
NOx (Non urban)	5000
PM10 (Urban)	225000
PM10 (Non urban)	50000
CO (Urban)	500
CO (Non urban)	100

3.2.2. The costs

In the frame of this paper, direct costs falling to the company namely the leasing and the fuel consumption are considered. The leasing covers the depreciation and the insurance costs as well as repairs and maintenance including tires. In the case of public transport, only subscription fees are taken into account.

Then, the total cost of the fleet can be evaluated and expressed in € per month. This does not require specific modelling. It is detailed in the analysis of the proposed case study.

3.2.3. The social aspects

The third main pillar of the sustainable development is the social aspects. Comfort and time gains can vary strongly due to the congestion of the road network.

First, the comfort is considered and evaluated on the following qualitative scale: High, Medium and Low.

The time gain can be quantified. It is quite easy to estimate the average time for travelling by cars or by using public transport in usual situations (excluding strikes or exceptional accidents). All in all, these indicators are aggregated into one social score by using the PROMETHEE method allowing the user to combine quantitative and qualitative approaches. This is detailed in the analysis of the study case.

3.3. Towards the sustainable composite score

The composite score is made up of the three previous super indicators related to the costs, the pollution and the social aspects aggregated by using PROMETHEE.

Based on this composite score, current official car policies can be compared with new designed alternatives as it is demonstrated in the study case. Moreover, this composite indicator makes it possible to perform sensitivity analyses ensuring robust recommendations.

The above mentioned aggregations are implemented in a software based solution used in the case study while the actual evaluations are performed in spreadsheets and directly imported into the software.

4. STUDY CASE

The case studied in the frame of this paper revolves around a small company located in Brussels, providing official cars to their 20 senior executives. The monthly fees include the vehicle hiring as well as the insurances, the repair and maintenance. For the fleet currently used, the monthly fee is about €432,- per vehicle.

Among the senior executives, five come from Tienen to the company headquartered in the center of Brussels, 10 come from Gent and 5 from Borgworm. The routes are characterized as follows:

Table 2: Route description, Scenario 0

Routes	Highway	City road	Total
Tienen – Brussels	38 km	9 km	47 km
Borgworm – Brussels	68 km	10 km	78 km
Gent - Brussels	45 km	11 km	56 km

Two alternatives are considered. Firstly, it is proposed to replace the current fleet by new vehicles emitting less toxic substances. This alternative is characterized by the same route as Scenario 0 and the monthly related fees are about €451,- per vehicle.

The second alternative aims at convincing the senior executives coming from Tienen and Borgworm to use rail mode since a door-to-door service is provided by rail network between Tienen, Borgworm and Brussels while the vehicles used for "Gent-Brussels" are replaced by new vehicles with lower emissions.

Regarding this alternative, the route lengths are as follows:

Table 3: Route description, scenario 2

Routes	Highway	City roads	Total
Tienen – Brussels	50 km		50 km
Borgworm – Brussels	80 km		80 km
Gent - Brussels	45 km	11 km	56 km

Since the authors did not emphasize the urban and non urban sections for the *TTW* emissions, it is not necessary to split the length of the routes Tienen - Brussels and Borgworm - Brussels.

4.1. The environmental footprint

Based upon the data presented in (Van Essen et al., 2003), the authors performed the calculations for evaluating the *WTT* as well as the *TTW* emissions in urban and non urban sections by using the equations detailed in Section 3.2.1.

The next tables present the data used for performing the environmental evaluations.

Table 4: Data base, Current Fleet, TTW emissions

	Current Fleet	
	City Road	Highways
CO2 [g/km]	222	154
NOx [g/km]	0,86	0,52
CO [g/km]	1,31	0,19
SOx [g/km]	0,048	0,032
PM [g/km]	0,144	0,071
Energy Use [MJ/km]	3,03	1,75

Table 5: Data base, New Fleet, TTW emissions

	New Fleet	
	City Road	Highways
CO2 [g/km]	192	111
NOx [g/km]	0,41	0,25
CO [g/km]	0,33	0,05
SOx [g/km]	0,006	0,004
PM [g/km]	0,046	0,031
Energy Use [MJ/km]	2,62	1,51

Table 6: Data base, Diesel engine, WTT emissions

Emissions of Refining (g/MJ)					
	CO2	NOx	CO	SO2	PM10
Diesel	6,8	0,036	0,005	0,052	0,001

Table 7: Data base, Electric engine, WTT emissions

Electric engines	
CO2 [g/MJ]	178
NOx [g/MJ]	0,45
CO [g/MJ]	0,03
SOx [g/MJ]	1,04
PM [g/MJ]	0,05
Energy Return	0,38
Energy Use [MJ/SKM]	0,31

Table 8 and Table 9 present the valuation of *TTW* and *WTT* emissions for Scenario 0 revolving around the use of the present fleet. It has to be noted that the total emissions are obtained by considering the round trip. HW states for Highways and CR for City roads.

Table 8: Scenario 0, TTW emissions [g]

	Gent Brussels	Tienen Brussels	Borgworm Brussels	TOTAL (x2)[g]
CO2 HW	1524600	643720	1151920	6640480
CO2 CR	537240	219780	244200	2002440
CO HW	1881	794,2	1421,2	8192,8
CO CR	3170,2	1296,9	1441	11816,2
NOx HW	5148	2173,6	3889,6	22422,4
NOx CR	2081,2	851,4	946	7757,2
SOx HW	316,8	133,76	239,36	1379,84
SOx CR	116,16	47,52	52,8	432,96
PM HW	702,9	296,78	531,08	3061,52
PM CR	348,48	142,56	158,4	1298,88

Table 9: Scenario 0, WTT emissions [g]

	Gent Brussels	Tienen Brussels	Borgworm Brussels	TOTAL (x2)[g]
CO2	117843,33	49769,27	89042,3	513309,8
CO	119,955	63,845	95,75	559,1
NOx	657,03	290,61	501,54	2898,36
SOx	934,23	407,65	710,98	4105,72
PM	50,655	34,585	43,39	257,26

The *TTW* and *WTT* emissions of Scenario 1 are presented on Table 10 and Table 11. The use of stricter emission standard engines leads to an important pollution reduction.

Table 10: Scenario 1, TTW emissions [g]

	Gent Brussels	Tienen Brussels	Borgworm Brussels	TOTAL (x2)[g]
CO2 HW	1098900	463980	830280	4786320
CO2 CR	464640	190080	211200	1731840
CO HW	495	209	374	2156
CO CR	798,6	326,7	363	2976,6
NOx HW	2475	1045	1870	10780
NOx CR	992,2	405,9	451	3698,2
SOx HW	39,6	16,72	29,92	172,48
SOx CR	14,52	5,94	6,6	54,12
PM HW	306,9	129,58	231,88	1336,72
PM CR	111,32	44,54	50,6	412,92

Table 11: Scenario 1, WTT emissions [g]

	Gent Brussels	Tienen Brussels	Borgworm Brussels	TOTAL (*2)[g]
CO2	101682,02	42943,82	76830,84	442913,36
CO	103,565	55,139	82,674	482,756
NOx	566,984	250,8048	432,8128	2501,2032
SOx	806,168	351,7936	613,5296	3542,9824
PM	43,769	29,8918	37,4948	222,3112

Finally, the *TTW* and *WTT* emissions of Scenario 2 are presented in Table 12 and Table 13. It has to be noted that the *TTW* emissions are only due to the route Gent-Brussels since the railway is used by the employees coming from Tienen and Borgworm.

Table 12: Scenario 2, TTW emissions [g]

	Gent Brussels	TOTAL (x2)[g]
CO2 HW	1098900	2197800
CO2 CR	464640	929280
CO HW	495	990
CO CR	798,6	1597,2
NOx HW	2475	4950
NOx CR	992,2	1984,4
SOx HW	39,6	79,2
SOx CR	14,52	29,04
PM HW	306,9	613,8
PM CR	111,32	222,64

Table 13: Scenario 2, WTT emissions, [g]

	Gent Brussels	Tienen Brussels	Borgworm Brussels	TOTAL (x2)[g]
CO2	101682,02	5242,1	8387,36	230622,96
CO	103,565	0,8835	1,4136	211,7242
NOx	566,984	13,2525	21,204	1202,881
SOx	806,168	30,628	49,0048	1771,6016
PM	43,769	1,4725	2,356	95,195

Tables 12 and 13 highlight the really substantial impacts of using electric public transport.

Then, the authors used the developed related software to elaborate the environmental footprint on the basis of these evaluations and the societal weights. Regarding the preference functions, the authors propose to use the usual function for conserving all the information during the decision process. The next figure presents the environmental footprint of each alternative.

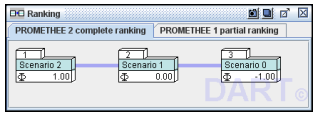


Figure 1: The environmental footprint

Scenario 2 obtains the best environmental footprint expressed by an outranking net flow equal to 1.

4.2. The global cost

The monthly leasing cost is about €432,- per vehicle for the current fleet while it is equal to €451,- per new vehicle. Then, the monthly leasing cost of the current scenario is about €8640,- while the monthly leasing cost of the proposed new fleet would be about €9020,- since 20 vehicles are concerned.

Regarding the second alternative, the price of a monthly subscription is about €146,- for those coming from Tienen and Borgworm since the travelled distances are in a same price category.

Then, the cost of Scenario 2 would be equal to €5970,- per month.

The global cost is made up of another compound; namely the fuel cost. The fuel consumption of the two fleets is shown on the next table:

Table 14: Fuel consumption

	Urban area	Non urban area
S0 (current fleet)	7,5 l/100km	4,6 l/100km
S1 and S2 (new fleet)	6,5 l/100km	4,3 l/100km

Then, regarding the route description detailed in Table 2 and Table 3, the monthly consumption of the current fleet is about 2660,02l while the monthly consumption of the proposed new fleet is about 2440,46l. Scenario 2 is characterized by a monthly fuel consumption of 1166 liters.

When observing the current situation, we can use an average of the fuel price based on the last two years such as €1,2 per liter while some estimations lead to a weak fuel price increase in the coming years, compared to the present situation, to €1,- per liter. Indeed, the possible replacement of the fleet requires some time for its actual implementation so that one can consider that the fuel price would be changed by that horizon.

So, the fuel cost of the current scenario is about €3192,03 per month and the fuel cost of Scenario 1 is about €2928,55 per month due to the fuel consumption and fuel price decrease. The fuel price of Scenario 2 is about €1166,-.

All in all, the global cost of the current scenario is about €11832,03 per month while the monthly cost of Scenario 1 is about €11948,55 and Scenario 2 is about €7136,-. Therefore, Scenario 2 is the cheapest scenario due to the low price of public transport subscription in this particular case characterized by a door-to-door service ensured by train.

4.3. The social score

Regarding the social indicator integrating the travel time as well as the comfort during the journey, the authors proposed to use the PROMETHEE method since it allows combining qualitative and quantitative assessments. The comfort is based on a qualitative "three levels" scale while the travel time is well known for both train and road transport modes. It has to be noted that for road transport, an additional travel time of 30 minutes is considered compared to the usual travel time due to rush hours in the morning and in the afternoon.

70 minutes are necessary to reach Brussels from Tienen while 85 minutes are required from Borgworm and 77 minutes from Gent. Regarding the public transport, 38 and 58 minutes are needed for reaching Brussels from Tienen and Borgworm respectively.

Regarding the second indicator, three linguistic levels are proposed: Low, Medium and High. The authors supposed that comfort – inversely proportional to stress levels - is high when using train since the employee does not need to concentrate himself on road traffic, congestion or risk of accidents while it is only low when driving his car.

The next figure illustrates the social evaluation table while the social composite score is highlighted on Figure 3.

Evaluation

	Confort T-B	Confort B-B	Confort G-B	Time T-B	Time B-B	Time G-B
Scenario 0	Low	Low	Low	70.0	85.0	77.0
Scenario 1	Low	Low	Low	70.0	85.0	77.0
Scenario 2	High	High	Low	38.0	58.0	77.0

Figure 2: Social evaluation table

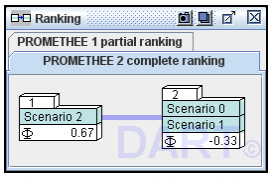


Figure 3: Social ranking

It has to be noted that the authors allotted the same weight to each of the 6 social criteria since it is impossible to give more or less importance to employees of the three different sections. Scenario 2 is the best alternative regarding the well-being of the employees with a social score about 0,67.

4.4. Towards the sustainability

Based upon the three above calculated "super-indicators", the authors proposed to go a step further by aggregating them into a final composite score integrating the three major pillars of the sustainable development.

So, the next figure summarizes the three "macro evaluations" leading to the final ranking based on a composite sustainable score.

Evaluation

	ENV	ECO	SOC
Scenario 0	-1.0	11693.82	-0.3333333333
Scenario 1	0.0	11331.76	-0.3333333333
Scenario 2	1.0	7136.0	0.6666666666

Figure 4: The final evaluation table

After using once again the PROMETHEE method, the authors obtained the following final ranking integrating all the above mentioned information. Scenario 2 obtains the best sustainable composite score. This means that scenario 2 is the best alternative regarding the sustainable management of this particular official car policy.

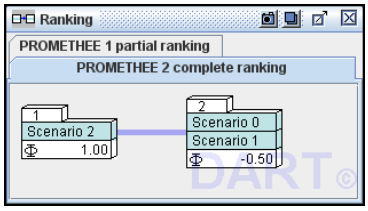


Figure 5: Sustainable composite score

The authors performed sensitivity analyses to test the robustness of the recommendations based on Figure 5. In this particular case, and as already observed in the three super indicators, Scenario 2 is the best alternative whatever the allotted weights or preference functions are since it is the best regarding all the indicators.

However, it is interesting to observe that no distinction is made between Scenario 0 and Scenario 1. It is due to the usual preference function used in the last aggregation step for the macro economic indicator. Indeed, if the U-shape function is used for the global cost with a preference threshold equal to €200,- which is slightly bigger than the deviation between the global costs of S0 and S1, the following composite scores are obtained.

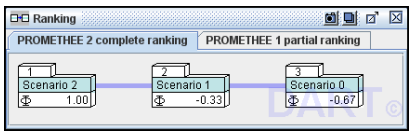


Figure 6: Modified sustainable composite score

This means that moving the preference thresholds could lead to different recommendations according to the preferences of the decision makers.

5. CONCLUSION

The goal of this paper was to elaborate a composite score for helping decision makers in designing and choosing sustainable official car policies. Such a score would help each company during the decision process. Moreover, the idea was to develop a software based solution able to support the use of such a score in order to perform automatically all the calculations and providing graphs for helping the communication.

To achieve this goal, the authors review many books, articles and EU directives in order to define precisely the framework and the objectives to meet.

Then the composite score was elaborated on the basis of the three main pillars of the sustainable development, namely the costs, the pollution and the social aspects. A top down approach was used to evaluate the impacts of "official car policy" since actual data are not systematically available in companies. This led to a demonstration of the applicability of the elaborated composite score and the related software.

Finally, it has to be noted that perspectives of development are identified in the frame of this paper. Indeed, in (Maibach et al. 2007), a handbook on the estimation of external costs of transport activities is presented. Then, it would be useful to incorporate these external costs such as road congestion, traffic safety, etc in the proposed composite score to provide a more global picture of the official car policy impacts.

ACKNOWLEDGMENTS

The authors of this paper want to thank the Brussels Government and IRSIB "Institut d'encouragement de la Recherche Scientifique et de l'Innovation de Bruxelles".

REFERENCES

Azibi, R. & Vanderpooten, D., 1997, Elaboration de critères agrégeant des conséquences dispersées : deux exemples concrets , In: Colorni A., Paruccini M., Roy B., eds. *The EURO working group on multiple criteria decision aiding, A-MCD-A ; Multi Criteria Decision Aiding*, pp 13-30.

Blaauw, H.G., de Wilde, H., Hekkenberg, R.G., 2006, Environmental impact in relation to economics of inland navigation, Proceedings of the *PIANC conference*, 2006, Estoril.

Brans, J.P., & Mareschal, B, 2005, PROMETHEE Methods, In: J.Figuera, S. Greco and M. Ehrgot eds. *Multiple Criteria Decision Analysis: State of the Art Surveys*, Berlin, Springer-Verlag , Berlin, pages 163-189.

Cescotto, S., Roubens, M., Rigo, N., Gao, S., Wang, X., Zhang, A., Lourenco, N., Zhou, J., Xiang, X., Ferreira, JPL. Management of Water Pollutants Based on Multi-Criteria Analysis and Fuzzy Logics, *Computational Methods in Engineering and Science EPMESC X*, Augustus, 2006, Sanya, Hainan,China, Tsinghua University Press & Springer-Verlag.

Directorate-General for Energy and Transport, 2006, *White Paper, European transport policy for 2010: time to decide*, Brussels, EU Commission

Directorate-General for Energy and Transport, 1994, *Inventory of the state of the art for cost benefit and multicriteria methods for projects in inland waterway*s, Brussels, EU Commission

Directorate-General for Energy and Transport, 2007, *Preparation of an impact assessment on the internalisation of external costs*, Brussels, EU Commission

Directorate-General for Energy and Transport, 2006, *EU Energy and Transport in figures, Statistical pocketbook 2006*, Brussels, EU Commission.

Directorate-General for Energy and Transport, 2007, *EU energy and transport tin figures, Statistical Pocket Book 2007/2008*, Brussels, EU Commission.

Directorate-General for Energy and Transport, 2008, *EU energy and transport tin figures, trends to 2030*, Brussels, EU Commission.

Haenen, M., Becker, J., Burgess, A., 2006,CREATING, Deliverable WP2, Final Report, *Innovative logistics concepts for Integrated Transport Chains*, Brussels.

Hekkenberg, R.G., Ndiaye, A.B., Lundoluka, F., Rigo, N., Unified performance indicators for inland waterway transport, Proceedings of the *PIANC conference*, 2006, Estoril.

Maibach, M., Schreyer, C., Sutter, D., Van Essen, H.P., Boon, B.H., Smokers, R., Schroten, A., Doll, C., Pawlowska, B., Bak, M., 2007. *Handbook on estimation of external cost in the transport sector, International measures and policies for all external costs of transport*, Delft, CE.

Ndiaye, A.B., Rigo, N., Lundoluka, F., Hekkenberg, R.H., Hadhazi, D., Simongati, Gy., Hargitai, Cs., 2007. CREATING, Deliverable WP4, Final Report, *Performance Assessment*, Delft.

Ndiaye, A.B., Glansdorp, C. Prunieras, J., Willems, C., 1993. A particular approach for the assessment of River Information Services (RIS) in Europe, *International Navigation Association ISSN 1680-1059*, n°114, pp79-92.

Pomerol, J.C. et Barba-Romero, S., 1993. *Choix multicritères dans l'entreprise*, Paris, Hermès.

Rigo, N., 2008. *An advanced integrated assessment methodology of multimodal transport impacts: elaboration and application to European cases*, Ph. D. thesis, University of Liège.

Rigo, N., Ndiaye, A.B., Dreyer, B., Zomer, G., Pinon, H., Tremeac, Y., 2007. Elaboration of an integrated assessment framework for the global evaluation of E-maritime operations, Proceedings of the *BIVEC-GIBET Transport Research Day*, pp 281-293, April 2007, Rotterdam.

Rigo, N., Hekkenberg, R.G., Hadhazi, D., Ndiaye, A.B., Simongati, Gy., Hargitai, Cs., Halet, F., 2007. Integrated Indicator for the sustainable assessment of intermodal chains; Application to the transport

of cargos from Frankfurt am Main (Germany) to Sofia (Bulgaria), Proceedings of the *European Inland Waterways Navigation conference*, June 2007, Visegrad.

Rigo, N., Hekkenberg, R.G., Hadhazi, D., Ndiaye, A.B., 2007. Performance assessment for intermodal chain, Proceedings of the *BIVEC-GIBET Transport Research Day*, pp 707-723, April 2007, Rotterdam.

Rigo, N., Hekkenberg, R., Ndiaye, A.B., Hadhazi, D., Simongati, Gy. 2007. Performance Assessment for intermodal chains. In : *European Journal of Transport and Infrastructure Research*.

Rigo, N., Ndiaye, A.B., 2008, Evaluation des performances d'une chaîne de transport; développement et utilisation d'un modèle intégré d'aide à la décision dans un cas d'étude, Forum de Recherche en Transport et Logistique, Sousse, April 2008.

Rigo, N., Ndiaye, A.B., 2008, Decision support system applied to combined freight transport. Proceedings of the *7th International Workshop on Modelling and Applied Simulation*, pp 106-115, September 17-19 2008, Campora S. Giovanni (Amantea), Italy.

Roy, B. et Bouyssou, D., 1993. *Aide multicritère à la décision : Méthodes et cas*, Paris, Economica, coll. « Gestion ».

Roy, B., 1985. *Méthodologie multicritère d'aide à la décision*, Paris, Economica , coll. « Gestion ».

Schweigert, D., 1995. Proceedings of the *5th Workshop of the DGOR-Working Group – Multicriteria Optimization and Decision Theory*, FB Mathematik Universität, Kaiserslautern.

Timmermans, J-M., Matheys, J., Van Mierlo, J., Lataire, P., 2006. Environmental rating of vehicles with different fuels and drive trains: a univocal and applicable methodology, *European Journal of Transport and Infrastructure Research*, no. 4, pp. 313-334.

Van Essen, H., Bello, O., Dings, J., Van den Brink, R. *To shift or not to shift, that's the question; The environmental performance of the principal modes of freight and passenger transport in the policy-making concept*, Delft, CE, March 2003.

AUTHORS BIOGRAPHY

Dr. **Alassane B. NDIAYE** is a Civil Engineer, Master of Business Administration, Master of Transport Sciences and Doctor of Applied Sciences. Educated in Supply Chain Management strategies at Stanford University (California, USA) and at the Massachusetts Institute of Technology, (MIT, Boston, USA), he is the holder of the Professorship Chair of Innovation for Quality, Logistics, Transport and Supply Chain Management at the Polytechnical Faculty of the Universite Libre de Bruxelles (Belgium). Former Expert adviser to the Ministry of Transport, he is the Executive Director of the Belgian Interuniversity Center for Mobility, Transport & Logistics, a plateform founded 10 years ago by 8 Belgian universities aiming, among other things, at enforcing the cross sectional and interdisciplinary approach to logistic and transport problems (both in education and R&D). Member of prestigious international scientific organizations, Dr. Ndiaye is a guest professor in many universities in Africa, South America and South East Asia.

Dr. **Nicolas Rigo** is a Bachelor in Mathematics, Master of Business Administration, Master of Transport Sciences and Doctor in Engineering. He works with Dr. Ndiaye in the frame of the Chair of Innovation for Quality, Logistics, Transport and Supply Chain Management at the Polytechnical Faculty of the Universite Libre de Bruxelles (Belgium) and is responsible for the development of a Spin-Off at the Universite de Bruxelles (Junior Consultancy Company) involved in the design and choice of sustainable logistics solutions. Dr. Rigo is a guest professor in universities in South-America and Africa.

MODELLING OF TRANSPORT PROJECT UNCERTAINTIES: RISK ASSESSMENT AND SCENARIO ANALYSIS

Kim Bang Salling[(a)], Steen Leleur[(b)]

[(a), (b)] Department of Transport
Technical University of Denmark (DTU)
Build. 115, DTU, DK 2800 Kgs. Lyngby
Denmark
[(a)]kbs@transport.dtu.dk, [(b)]sl@transport.dtu.dk

ABSTRACT
This paper proposes a new way of handling the uncertainties present in transport decision making based on infrastructure appraisals. The paper suggests to combine the principle of Optimism Bias, which depicts the historical tendency of overestimating transport related benefits and underestimating investment costs, with a quantitative risk analysis based on Monte Carlo simulation and to make use of a set of exploratory scenarios. The analysis is carried out by using the CBA-DK model representing the Danish standard approach to socio-economic cost-benefit analysis. Specifically, the paper proposes to supplement Optimism Bias and the associated Reference Class Forecasting (RCF) technique with a new technique that makes use of a scenario-grid. We tentatively introduce and refer to this as Reference Scenario Forecasting (RSF). The final RSF output from the CBA-DK model consists of a set of scenario-based graphs which function as risk-related decision support for the appraised transport infrastructure project.

Keywords: decision support, risk analysis, reference class forecasting, reference scenario forecasting

1. INTRODUCTION

This paper sets out a new methodology for examining the uncertainties relating to transport decision making based on infrastructure appraisals. The approach proceeds by combining the principle of Optimism Bias, which depicts the historical tendency of overestimating transport related benefits and underestimating investment costs, with a quantitative risk analysis based on Monte Carlo simulation and by using a set of exploratory scenarios. The analysis is carried out by using the CBA-DK model representing the Danish standard approach to socio-economic cost-benefit analysis. Specifically, the paper proposes to supplement Optimism Bias and the associated Reference Class Forecasting (RCF) technique with a new technique that makes use of a scenario-grid. We tentatively introduce and refer to this as Reference Scenario Forecasting (RSF).

The paper is disposed as follows. In Section 2 a description is given of Optimism Bias and Reference Class Forecasting. Section 3 presents the applied Greenland case study and the calculations carried out in the CBA-DK model together with a set of altogether nine scenarios. For one of the scenarios, the Reference Scenario 5, the input probability distributions based on RCF are described and the results from a model run are given. In the following Section 4 the principles of Reference Scenario Forecasting are presented and illustrated by a set of model runs. These RSF results consist of a set of scenario-based graphs which function as risk-related decision support for the appraised transport infrastructure project. The final Section 5 gives a conclusion and a perspective on the further research.

2. OPTIMISM BIAS AND REFERENCE CLASS FORECASTING

The Optimism Bias approach is dealt with by using a well-established technique named Reference Class Forecasting (RCF). The theoretical background is made up by prospect theory developed by Kahneman and Tversky in 1979 (Daniel Kahneman received the Nobel prize in Economics in 2002 for his work in collaboration with Amos Tversky (1937-1996)). Prospect theory describes decisions between alternatives that involve risk, i.e. alternatives where the general outcome is uncertain but the associated probabilities are known. A reference class denotes a pool of past projects similar to the one being appraised. Herein a systematically collection of past errors is gathered for a range of projects comparing the deficiencies in the planning stage. Experience from past projects is then collected and compared so that "planning fallacy" can be avoided (Flyvbjerg and COWI 2004).

Reference Class Forecasting is established on the basis of information from a class of similar projects. The classification of reference classes have been explored in Flyvbjerg and COWI (2004), pp. 13-14, where three main groups of projects has been statistically tested for similarities, namely roads (highways and trunk roads), rail (metro, conventional rail and high speed rail) and fixed links (bridges and tunnels). Hence, RCF does not try to forecast specific uncertain events that will affect the particular project,

but instead it places the project to be evaluated in a statistical distribution of outcomes from this class of reference projects. Flyvbjerg et al. (2003) have built a large pool of reference class projects divided into three types of transport-related infrastructure investments, namely road, rail and fixed links projects. From the latter Salling (2008) has performed a data analysis uncovering a set of probability distributions that fit the data from Flyvbjerg et al. (2003) associated with transport infrastructure assessments see Table 1.

Table 1: Fitted distributions from Salling (2008)

Impact	**Distribution**
Travel time savings	Beta-PERT
Construction costs	Erlang

The two distributions depicted in Table 1 have been fitted against reference class projects concerning travel time savings (traffic demand forecasts) and construction costs. These two impacts make up the key components in most transport evaluation schemes (Leleur 2000), for which reason the following case study applies these distributions for a risk assessment study in Greenland.

3. THE GREENLAND CASE STUDY

The paper makes use of information comprised in Leleur et al. (2007), Salling (2008) and Salling and Banister (2009) in which an examination of a new international airport in Nuuk is presented by three project alternatives. These consist of two alternatives replacing the existing runway in Nuuk, i.e. increasing the current runway length to either 1799 metre (m) or 2200m, and as the third alternative the construction of a new, relocated airport to the south with a 3000m runway, consequently leading to a closing of the existing airport. Results from this study clearly pointed towards either of the two extension alternatives leaving the Nuuk 3000m alternative infeasible from a societal perspective. Finally, an article posted on the website of the Home Rule Authorities in Greenland, October 2007 outlined that the Nuuk 2200m alternative has been selected for implementation (Kristensen 2008). In light of this information, this paper examines the robustness of this decision based on combining a set of scenarios with risk analysis.

3.1. The CBA-DK Model

The CBA-DK model combines deterministic calculation based upon conventional cost-benefit analysis (CBA) with a stochastic calculation based on a quantitative risk analysis (QRA). This model is in accordance with the socio-economic analysis guidelines provided by the Danish Ministry of Transport (DMT 2003). It is developed on a Microsoft Excel platform forming the basis of the CBA, and the QRA is carried out with an add-in software from Palisade named @RISK which implements a standardized Monte Carlo simulation (Palisade 2007; Salling 2008). The deterministic calculation from CBA-DK produces the following decision criteria for the Nuuk case as shown in Table 2.

Table 2: Decision criteria from a deterministic CBA model run for Nuuk 2200m (Salling and Banister 2009)

Construction costs	1,059 MDKK
Benefit cost ratio (BCR)	2.5
Internal rate of return (IRR)	13.8%
Net present value (NPV)	1,706 MDKK
First year rate of return (FYRR)	19.8%

These criteria values show clearly that the 2200m alternative produces very good societal results with a significantly higher NPV (1€ ≈ 7.5 Danish Kroner (DKK)). However, the results only depict one set of possible outcomes. To provide strategic decision support the CBA-DK model is used on a set of exploratory scenarios that express external economic factors e.g. a deregulation regime combined with a specific socio-geographic development e.g. Nuuk getting higher importance as centre.

3.2. Scenarios

The scenarios in this study have been set up with respect to two main types of regimes: Three global regimes which deal with the overall international economic development and three regional/local regimes describing the future importance of Nuuk as centre

Global economic development

Nuuk importance as centre and regional growth pole in Greenland	Deregulation	Regulation	Stagnation
Clearly higher	Sc. 1	Sc. 4	Sc. 7
Moderately higher	Sc. 2	**Ref sc. 5**	Sc. 8
Same	Sc. 3	Sc. 6	Sc. 9

Increasing uncertainty (vertical)

Increasing uncertainty (horizontal)

Figure 1: Scenario-grid for imagined futures representing possible and plausible development

(adapted from Leleur et al. 2004). The regimes vary as depicted in Figure 1 where the horizontal axis outlines the global economic development and the vertical axis outlines the importance of Nuuk as centre and regional growth pole in Greenland. Uncertainty tendencies as relating to the regimes have also been indicated.

Altogether nine scenarios have been formulated, which are expected to have different influences on the feasibility of the Nuuk airport investment. The set of scenarios is expressing a range of possible and plausible developments, each of which could prevail as the context of the appraisal study. The influences are discussed below relating these both to the deterministic and stochastic CBA-DK calculations.

3.3. Reference scenario 5

To enhance the understanding of the uncertainties involved a Monte Carlo simulation is performed (Vose 2002; Salling 2008). Selecting appropriate probability distributions to acknowledge the embedded impact uncertainties presents the critical part of this calculation procedure. As previously presented in Table 1 two underlying transportation impacts are implemented in terms of an Erlang distribution and a Beta-PERT distribution.

3.3.1. Construction costs (Erlang distribution)

Construction costs for large infrastructure projects have a tendency to be underestimated, which means that socio-economic analyses become overoptimistic. These misinterpretations of ex-ante based costs, deliberate or otherwise, result in budget overruns. From the data derived from Flyvbjerg et al. (2003) a sample collection of 57 rail type projects revealed that 88% of the infrastructure projects experienced costs overruns. It has been assumed that the empirical results from rail projects can be applied to airport infrastructure projects. In the database no separate section on airport infrastructure projects was available but rail projects were judged to be the most suitable project class. Thus, the dataset as concerns rail infrastructure projects resulted in input parameters towards the Erlang distribution with a shape parameter of 9 and a scale parameter estimated on the data set (Lichtenberg 2000; Salling and Banister 2009). The worst observation from the data sample, with a cost overrun of 100% has been used as upper limit while a best case observation, cost underrun, occurred with -15% as a lower limit. In this context, a cost overrun of 100% means that the ex-ante based construction cost estimate in reality was exceeded by 100%.

3.3.2. Travel time savings (PERT distribution)

Typically, travel time savings are calculated on basis of current traffic flows provided by a traffic and forecast model. Hereby, future traffic flows are determined based upon a forecast rooted in e.g. past data information, expert judgments, empirical evidence, etc. However, such a futuristic demand forecast is extremely troublesome to make (Priemus et al. 2008). The same data material reveals a comparison between 27 rail projects depicting the inaccuracy for traffic demand forecasts. The overestimation of demand forecast, and hereby mis-calculations in terms of travel time savings, occurs in almost 85% of the cases. The worst observation from the data sample, with a demand underrun of -95% has been used as lower limit while a best case observation, with a demand overrun, occurred with 75% as an upper limit. In this context, a demand underrun of -95% means that the ex-ante developed forecast was under-exceeded by 95%.

3.3.3. Results

The CBA-DK model provides the deterministic point results as illustrated in Table 2 including a stochastic calculation which enhances the point results into interval results allowing for the decision-makers to explore their risk aversion towards the appraised scheme. The latter is performed through a Monte Carlo simulation with the Optimism Bias based input. The results of the focal reference scenario 5 are presented as an accumulated descending graph (ADG), see Figure 2. The shown ADG delivers information with regard to the probability of achieving a BCR higher than or equal to the x-axis value. Hence, the ADG is important as a means to involve decision-makers and support strategic decision-making based upon their revealed risk aversion.

The ADG pictured in Figure 2 shows that for approximately 80% of the cases the reference scenario 5 gives a feasible result with the BCR > 1.0. However, decision-makers with risk aversion would probably take into account that in 20% of the simulation runs scenario 5 gives an infeasible result.

The remaining 8 scenarios take the basis from the focal reference scenario 5. By using the two different types of regimes, the input parameters for the two probability distributions are set according to an assessment of the uncertainties as they are perceived under the specific scenario. This is carried out by using the principles of Reference Scenario Forecasting (RSF) as outlined below.

4. REFERENCE SCENARIO FORECASTING

In order to operationalise the use of scenarios in CBA-DK the previous technique of Reference Class Forecasting based on the Optimism Bias has been combined with Monte Carlo simulation and scenario analysis.

The reference scenario 5 will form the basis (focal scenario) for RSF and the related 8 scenarios will be set by assessing the development in expected travel time related benefits. It has been assumed that in the actual case the construction cost effect is independent of the regimes, for which reason the input parameters to the Erlang distribution remain as presented in section 3.2.1.

The travel time savings, however, will no doubt change as a consequence of the economic development. Clearly, deregulation and high economic growth will mean more people that travel both as tourists, residents

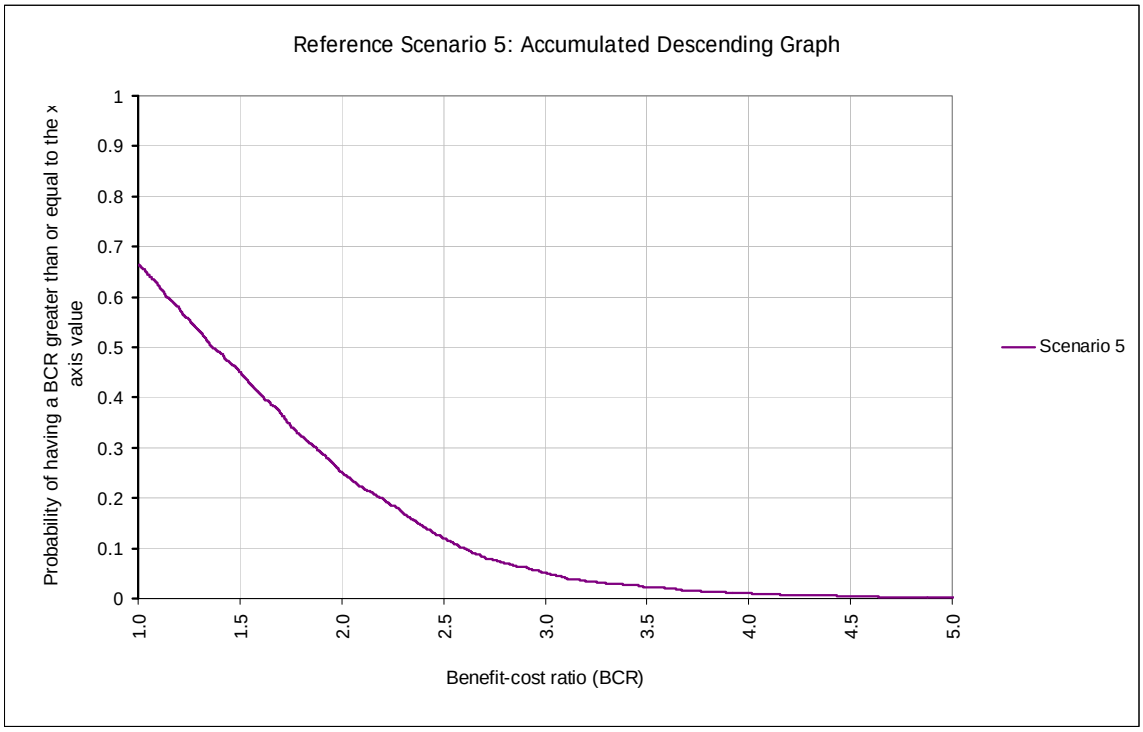


Figure 2: Resulting accumulated descending graph (ADG) for the focal scenario 5: the y-axis values for BCR = 1.0 indicate the certainty levels of the scenario

and business travellers. The opposite tendency will turn out in the case of stagnation or financial crisis. All trips will then be at a minimum and the travel time savings effect will decrease due to the lower passenger number.

The variation between scenarios is systematically explored and related to the scenario-grid (Figure 1). The specific scenario input concerning the Beta-PERT distribution is assessed by making use of the triple estimation technique in a "backward way" compared to its intended use (Lichtenberg 2000) and by anchoring its initial parameter-setting with the values for the focal scenario 5.

4.1. Triple values for the focal scenario 5

The main idea of Reference Scenario Forecasting is based on assessing the most likely (ML), the maximum (MAX) and the minimum (MIN) values under the various scenario conditions. The assessment is carried out based on knowledge of these values under the focal scenario 5, where the triple set values have been determined as follows with all values in mio DKK (1):

$$(MIN5, ML5, MAX5) = (10, 170, 300) \quad (1)$$

The assessment is based on this anchoring information being available and interpreting how the values will change under the changed scenario conditions. The importance of anchoring information has been treated by Goodwin and Wright (2004, pp. 309-325), while the value of using triple estimates for exploring uncertainty has been examined by Lichtenberg (2000, pp. 119-132) and Vose (2002, pp. 272-278). In the following we will exemplify some of the deliberations that have been used of to set the values shown in Table 3.

4.2. Triple values for scenario 2, 4, 6 and 8

In scenario 2 optimism in the global economy and deregulation lift the ML5 value to ML2 = 220 mio DKK. At the same time uncertainty is perceived to be decreasing, as indicated in Figure 1, which gives a higher MIN-value and a higher MAX-value. Hereby we obtain the following triple set for scenario 2 in mio DKK (2):

$$(MIN2, ML2, MAX2) = (50, 220, 330) \quad (2)$$

More or less the same tendency occurs with respect to Scenario 4 where the importance of Nuuk as a centre is growing. However, the uncertainty is increasing compared to the focal scenario 5, leaving the MIN5 more or less unchanged but giving a clearly higher MAX-value. Hereby, we obtain the following triple set values for scenario 4 in mio DKK (3):

$$(MIN4, ML4, MAX4) = (25, 200, 350) \quad (3)$$

The triple values for Scenario 8 are derived by taking into account that the global economy is stagnating, which leads to increasing uncertainty and a lower assessment of ML8 to 145 mio DKK. It has been assumed that the benefits from the travel time savings cannot be lower than 0 (lower boundary). In this way the following triple set has been arrived at for scenario 8 in mio DKK (4):

$$(MIN8, ML8, MAX8) = (0, 145, 300) \quad (4)$$

Finally, the triple set for scenario 6 is assessed based on the uncertainty being lower than under the conditions in the focal scenario. Nuuk as centre and growth pole is the same as today with a regulated regime. Thus, ML6 is lowered to 150 mio DKK with the following set of MIN and MAX values in mio DKK:

$$(MIN6, ML6, MAX6) = (10, 150, 285) \quad (5)$$

4.3. Triple values for the remaining scenarios

The remaining four scenario values have been found by using the triple sets assessed for scenario 2, 4, 6 and 8. As depicted in Figure 1, the highest uncertainty relates to scenario 7, while the most certain scenario is scenario 3. Table 3 shows the outcomes of the assessment of the nine scenarios from Figure 1 with the triple values in absolute terms (mio DKK).

Table 3: Summary of the triple values applied for the Reference Scenario Forecasting (mio DKK)

Scenario	MIN	ML	MAX
Scenario 1	50	250	400
Scenario 2	50	220	330
Scenario 3	25	175	325
Scenario 4	25	200	350
Scenario 5	**10**	**170**	**300**
Scenario 6	10	150	285
Scenario 7	0	170	315
Scenario 8	0	145	300
Scenario 9	0	100	250

In this context the triple values for the different scenarios have been set in accordance with discussion amongst the authors and mainly for the purpose to illustrate the approach of RSF. In a real-world application the values should be set by people with a thorough knowledge (stakeholders) of the project examined based on their assessment of the conditions that may influence them.

Consequently, a future task in this respect is to implement the use of a decision conference (DC) as part of the RSF approach. Essentially, a DC brings together decision analysis, group processes and information technology over an intensive two or three day session (Goodwin and Wright 2004, pp. 323-325). The DC makes it possible for the various stakeholders relating to the specific decision task to affect the course of action. Principally, a decision conference involves a set of stakeholders with all different perspectives towards the problem represented. For this demo-case, stakeholders could be representatives from the Home Rule authorities in Greenland, people from the Municipality of Nuuk, aviation experts, economists, etc. Their main challenge is to produce the triple set values under the different scenarios based on their knowledge and their assessment of the scenario conditions. Hereby the set of triple estimates in Table 3 may be changed into more realistic values.

4.4. RSF results for the Greenland case

Model runs in CBA-DK making use of the values in Table 3 produce 8 additional accumulated descending graphs (ADGs), see Figure 3.

The main output from the RSF is that none of the scenarios produces an ADG with 100% probability of achieving a BCR above 1.00. Scenario 1 returns a 95% certainty level that the Nuuk 2200 meter alternative is feasible whereas scenario 9 returns a 45% certainty

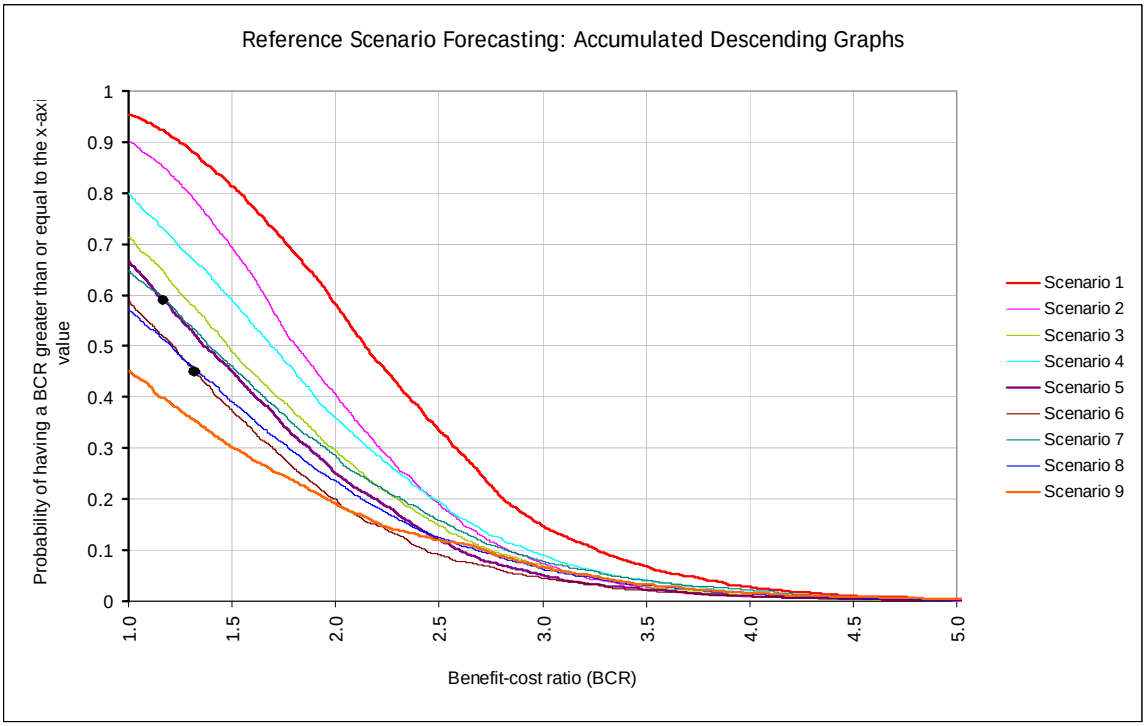


Figure 3: Resulting accumulated descending graphs (ADGs) from Reference Scenario Forecasting: the y-axis values for BCR = 1.0 indicate the certainty levels of the scenarios

level. Attention could be paid to scenario ADGs that intersect each other. Thus, scenario 8 crosses scenario 6 at a 46% threshold whereas scenario 7 crosses the reference scenario at 60%. Furthermore it can be noted that the flatness of the ADG corresponds to the degree of uncertainty assigned each scenario, i.e. a flatter ADG depicts a higher uncertainty.

Risk aversive decision-makers would probably accept the project under scenario 1, 2 and 3, where less risk aversive decision-makers would also include scenarios 4, 5 and 7. Under the condition of scenario 9 the project will probably not be accepted whereas scenarios 6 and 8 are more difficult to interpret. A next step towards a decision could be to estimate the probability of each scenario to get closer to a final decision. We foresee that making use of a decision conference will help qualifying the deliberations of the involved decision-makers.

5. CONCLUSION AND PERSPECTIVES

A characteristic feature of CBA is that it communicates its result by an economic index value, for example the benefit-cost ratio (BCR), which has been made use of in this paper to represent the calculation result of CBA. This index, BCR, can be seen as a point result as it communicates one value to represent the result of the assessment. Including risk considerations in transport project appraisal in general replaces the point result of the CBA with an interval result stemming from a wider analysis which combines CBA and risk analysis techniques.

By combining Optimism Bias and Monte Carlo simulation, the CBA-DK model makes a more explicit consideration of risk possible as concerns the probability of implementing a non-feasible project or for that sake of not implementing a feasible one. The concept of Reference Scenario Forecasting (RSF) has been introduced as a possible way of making operational use of scenarios, and its principles have been demonstrated by applying a case study from Greenland.

Altogether nine scenarios have been set out and assessed resulting in a set of graphs illustrating the influence on the appraisal result. These graphs allow the decision-makers to debate and decide on the basis of a risk-oriented feasibility approach within transport infrastructure appraisal. Currently, this new RSF approach uses two main types of regimes leading to the robustness valuation of the appraisal result. Further research will explore the application of more refined scenario descriptions with additional scenario information and the formulation of a decision conference set-up with the purpose of estimating the triple set values under the different scenario conditions.

In an ongoing research project about “Uncertainties in transport project evaluation” 2009-2012, funded by the Danish Strategic Research Council, the presented methodology will be further developed.

REFERENCES

Danish Ministry of Transport, 2003. *Manual for socio-economic analysis – applied methodology in the transport sector*. Copenhagen, Denmark (in Danish).

Flyvbjerg, B., Bruzelius, N. and Rothengatter, W., 2003. *Megaprojects and Risk – An anatomy of Ambition*. Cambridge University Press, United Kingdom.

Flyvbjerg, B. and COWI Consult, 2004. Procedures for Dealing with Optimism Bias in Transport Planning. *Guidance Document*, prepared for the British Department for Transport, United Kingdom.

Goodwin, P. and Wright, G., 2004. *Decision Analysis for Management Judgment*. 3rd edition, John Wiley & Sons, Ltd, United Kingdom.

Kahneman, D. and Tversky, A. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47 (2), 263-292.

Kristensen, K., 2007. *Atlantic airport in Nuuk saluted*. Home Rule Authority, Greenland. Available from: http://sermitsiaq.gl/indland/article19110.ece [accessed 14 November 2008]

Leleur, S., 2000. *Road Infrastructure Planning – a decision-oriented approach*. 2nd ed., Polyteknisk Press, Denmark.

Leleur, S., Holvad, T., Salling, K.B. and Jensen, A.V., 2004. Development of the CLG-DSS Evaluation Model, *CLG-Report task 9*, Centre for Traffic and Transport, Technical University of Denmark.

Leleur, S., Salling, K. B. and Jensen, A. V., 2007. Documentation and Validation of the TGB Evaluation model in Greenland. *Technical Report prepared for the Home Rule Authorities in Greenland*, Centre for Traffic and Transport, Technical University of Denmark (in Danish).

Lichtenberg, S., 2000. *Proactive Management of uncertainty using the Successive Principle*. Polyteknisk Press, Denmark.

Palisade Corporation, 2007. *@RISK, Risk Analysis and Simulation Add-In for Microsoft Excel - Version 5.0*. User Guide, NY, USA.

Priemus, H., Flyvbjerg, B. and Wee, B. van, 2008. *Decision-making on Mega-Projects: Cost-Benefit Analysis, Planning and Innovation.* Transport Economics, Management and Policy, General Editor: Kenneth Button, George Mason University, USA.

Salling, K.B., 2008. *Assessment of Transport Projects: Risk Analysis and Decision Support*. Thesis (PhD), Department of Transport, Technical University of Denmark.

Salling, K.B. and Banister, D., 2008. Assessment of large transport infrastructure projects: the CBA-DK model. *Proceedings of the International Conference on Infrastructure Systems*. November 10-12, Rotterdam, the Netherlands.

Vose, D., 2002. *Risk Analysis – a quantitative guide*. 2nd ed., John Wiley and Sons, United Kingdom.

BIBLIOGRAPHY

KIM BANG SALLING is currently employed as an assistant professor at the Department of Transport at the Technical University of Denmark. He defended his PhD thesis entitled: *Assessment of Transport Projects: Risk Analysis and Decision Support*, November 2008. He furthermore holds a Master's degree in Engineering within socio-economic evaluation methodologies and decision support systems with special emphasis on cost-benefit analysis and Risk Analysis. The Ph.D. project concerned a decision support model for assessing transport projects. By use of the developed CBA-DK software a new risk-oriented methodology for feasibility risk assessment is developed. Recently, he has been part of the Centre for Logistics and Freight (CLG) project where he co-developed a new Decision Support System applied for large scale infrastructure projects, under the CLG task 9 project. He has also been working on a software program for a composite evaluation model called COSIMA-VEJ for the Danish Road Directory. Furthermore, he has been involved in a large transport appraisal study in Greenland evaluating the overall transport system and a customized decision support model for the Rail Net Denmark in order for them to optimize asset management and project ranking. Currently, he is co-managing a large-scale research project for the Danish Strategic Research Council over the next four years entitled: *Uncertainties in Transport Project Evaluation.*

STEEN LELEUR is professor of decision support systems and planning at the Department of Transport at the Technical University of Denmark and the head of the Decision Modelling Group. Recently, he was, among other things, involved in research on topics about systems analysis and evaluation methodology in the Centre for Logistics and Goods Transport. His most recent book "Systemic Planning – Principles and Methodology for Planning in a Complex World" examines the issues of uncertainty and complexity as relating to strategic planning and evaluation problems. Previously, he has published textbooks in Danish and English about traffic planning and highway engineering. He received his M.Sc. in engineering in 1972 at the Technical University of Denmark and defended his doctoral dissertation at the same university about investment planning in 1984. From 1972 until 1986 he worked in the Danish Highway Agency. From 1986 until present Steen Leleur has been employed by the Technical University of Denmark, where he is responsible for the specialisation in traffic planning and highway engineering within the Master Programme "Transport and Logistics". For several years he was associated with the Copenhagen Business School, where he taught courses in systems science and planning theory. Over the years Steen Leleur has been involved in many international transport planning and evaluation research projects, among others several within the European Commission's strategic transport research programmes.

PROPOSAL OF A NEW MODEL FOR THE OPTIMIZATION OF THE ORGANIZATIONAL PROCESS IN INDUSTRIAL COMPANY THROUGH THE APPLICATION OF THE ANALYTIC NETWORK PROCESS

Fabio De Felice[(a)], Antonella Petrillo[(b)]

[(a), (b)]Department of Industrial Engineering - University of Cassino, Italy

[(a)]defelice@unicas.it, [(b)]a.petrillo@unicas.it

ABSTRACT

The evaluation process of a company presents many difficulties, especially with regards to the attribution of value to elements which are intangible by nature, such as the value of good management.

Thus, it is essential for companies to have tools that permit the evaluation and quantification of the elements of organizational and managerial nature.

The aim of this work is to analyze these issues by identifying and studying new indicators that take into account the characteristics and complexity of intangible assets.

Many models have been proposed, which focused on the enhancement of one or more aspects of the organization and business management.

In this work we have carried out through the multi-criteria decision making technique known as ANP - Analytic Network Process, a decision support system or a model dedicated to the assessment and quantification of the elements of organizational and managerial nature typical of SMEs in the engineering industry.

Keywords: ANP, Decision Support System, MCDA, Managerial skills.

1. INTRODUCTION

The model is a useful support to managers in the process of cultural revolution of the company management.

In fact, it is essential to have tools to highlight the strengths of the company which should be properly defended and supported as well as the weak points on which to intervene (Black and Gregersen, 2002).

In particular, we analyzed two different approaches, the first which shifts the attention to the monitoring elements of the so-called intangible assets, which are knowledge and skills that the company has (intangible resources), essentially establishing a set of indicators capable of measuring these elements.

The second approach, on the contrary, realizes the measurement of qualitative factors, not only the business activity in line with company-wide implementation of the concept of total quality.

Therefore it was, then necessary to identify variables which articulate the qualitative survey. These variables have been properly organized in a checklist divided into different hierarchical levels.

In particular in this work we have applied Multicriteria decision-making techniques that are suitable for such studies, in fact, they have been developed specifically to tackle problems where we must choose between a number of alternatives based on multiple attributes of various nature (Marakas, 1999).

The aim of our work is the realization of an instrument to measure the adequacy of the company with respect to qualitative parameters.

In particular, we have developed a model based on the Analytic Network Process to assess the managerial skills, in detail:

- **Business and strategic skills.**
- **Organizational and managerial skills.**

The method used allows us to evaluate different solutions and gives us the opportunity to choose the best one (Finan and Hurley, 2002).

The implementation of ANP allows us:

- To build a model that helps to measure and synthesize a large number of factors in complex decisions in an industrial plant.
- To take the best decision in relation to a multitude of targets allowing the decision maker the measure and the summary of the different factors / criteria or sub-criteria.

The assessment of qualitative factors, not necessarily subject to a numerical quantification is extremely delicate and in some ways too complex (De Felice, Falcone and Duraccio, 2000).

Thus, after having built the model to decrease the subjectivity and partiality of the evaluations made, we introduced the new indicators that take into account the characteristics and complexity of the main intangible assets:

- **GP index** - Global Productivity index (ie, system performance).
- **EA index** - Effective company Actions index (ie, make the right decisions).
- **CP index** – Company Profitability index (or ability to generate resources).

2. METHODOLOGY STRUCTURE

The accurate analysis of the quality of a company requires an analysis of the problem divided into elements which gradually decrease and are more easily measurable (Hult, Ketchen and Reus, 2001).

In order to create a discrimination between the elements, one of the most important problems that had to be overcome during the construction of the model was the identification of specific weights for the elements.

We concentrated our attention on the monitoring of the elements of a qualitative nature. We developed a model for the recognition of aspects of organizational management within the company. The steps we developed for the realization of the final model are:

1. Obtain the data, formulation and analysis of the problem.
2. Identify critical variables for the assessment (organized in a specific check-list).
3. Build multiattribute models. In particular, a model is proposed that provides for the allocation of various resources needed to quantify the organizational and managerial elements typical of a business.
4. Solve the problem using the ANP technique.
5. Construction of indicators summarizing the criteria, which are then combined to quantitatively assess each alternative. We performed a preliminary statistical analysis of these indicators and then carried out separate ranking for each criterion to study the behavior of areas in respect to individual aspects considered.

The goal was to develop a model able to describe how it should be organized according to a holistic concept, a company. The set of factors represented in the diagram, determine how the company should operate **(**Kanungo, Sharma and Jain, 2001).

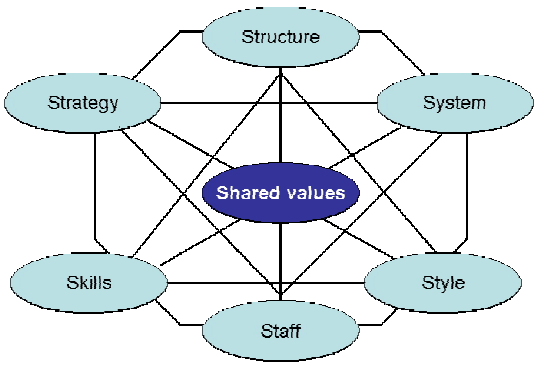


Figure 1: 7S model - holistic concepts

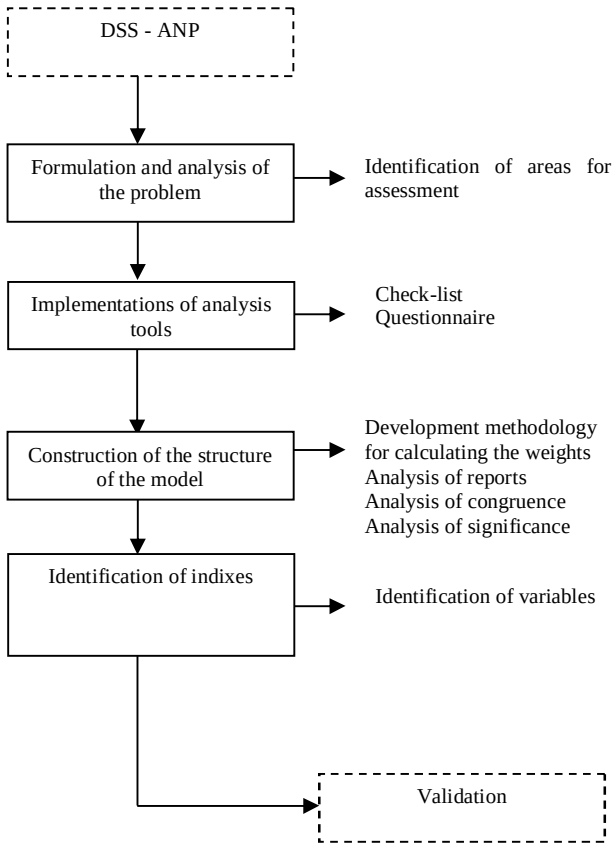


Figure 2: Phases of methodology

3. DEVELOPMENT OF METHODOLOGY

It is obvious that any analysis tool is to take note of the state of affairs and to measure the severity when it occurs (Olson and Courtney, 1997).

In fact, by knowing the problem you can find the solution.

At this point, we identified the areas of investigation relating to the organization and management of companies on which to focus our attention (Biggiero and Laise, 2003).

In this way we were able to identify a set of variables representative of the phenomenon in question worthy of consideration.

These same variables were organized into distinct and homogeneous groups according to the scope assigned to them.

Parallel to this analysis, we tried to identify an important discrimination between the variables in question, developing a methodology which would give different weights, and thus the priorities, the same variables.

The problem was addressed by involving experts in the field, explored their views and used them to assign weights to different variables.

3.1. Formulation and analysis of the problem

The analysis of the particular sector (engineering) led to identify that one of the main criticism is wide and low propensity to programming, a general approximation,

inadequate management and a company that remains more theoretical than an abstraction effective in practice, under which powers and responsibilities are focused, not always positively, in the figure of the entrepreneur (Humphreys, Ayestaren, McCosh and Mayon-White, 1997).

3.2. Implementation of analysis tools

To meet the challenge of managing an enterprise management means having tools that allow not only to have a precise view of the current situation, but mainly put in place the necessary measures that the situation requires (Lofti and Pegels, 1996).

In particularly, to establish a formal structure for the model we split the problem in different areas of inquiry and discover and emphasize the critical elements .

The problem, therefore, is twofold: firstly, it is necessary to identify the various elements of the phenomenon under study (in our case the essential elements of the organization and management of a metal) but at the same time we also need to try to quantify the presence of such elements within a business (Menkes, 2005).

Hence the creation of two different tools, **checklists and semi structured questionnaires.**

Check List

In particular, we used a check-list, formed from the variables identified and appropriately articulated on different levels, each incorporating progressively fewer elements.

At this point we prepared a questionnaire designed to enable us to detect those same variables in the company.

For each answer of the questionnaire there was a match score.

In practice, to obtain a rating of adequate quality standards the organization must have a proper management and an appropriate range of skills available to the entrepreneur.

This finding made sure that our investigation went along in two main strands: first we needed to deepen the theme of entrepreneurship, on the other hand it was necessary to clarify what factors could make the appropriate organization and management.

Specifically, then, it was necessary to create a list of variables to monitor.

Table 1: Check List

VARIABLES	
Business and strategic skills	**Organizational and managerial skills**
1 KNOW-HOW **Basic Training** Degrees Other securities Foreign languages **Further knowledge** Computer skills Knowledge of financial instruments Knowledge of bank risk Knowledge rules **Experience in the field** Activities in technical roles Activities in administrative roles Activities in managerial roles	**1. ORGANIZATIONAL CAPACITY** **Awareness of tasks and Individual responsibility** Presence organigramme Analysis of job profiles **Coordination capacity** Use of specific procedures Coordination meetings
2. CAPACITY 'OF BUSINESS DEVELOPMENT **Capacity analysis of market dynamics** **Analysis of external factors** Proper positioning of the company Identifying the strengths of the company **Market positioning and prospects business** Volume production the last three years Volume production next two years Quality Certifications Ability to diversify market Strategies of specialization Diversification strategies / action to the Global service Internationalization	**2. CAPACITY MANAGEMENT** **Planning and management control** System programming of activities System management control Checking the progress of work **Management of supplies** Criteria for the selection of Suppliers Quality control procedures **Human resource management** Staff training Adoption of an incentive system **Information System** Use of software for the management Using software to manage and control ICT deployment
3. STRATEGIC CAPACITY **Ability to establish financial strategies** Propensity to collaborate with banks Financing Project financing **Networking capability** Participation in consortia Propensity to cooperate Relationship for Innovation Universities	

In the check list we identified two main sections:

- **Business and strategic skills.** The fate of the small enterprise is inextricably linked to the ability of the entrepreneur: the emphasis has been on issues such as basic education.
- **Organizational and managerial capacity.** A key step consisted in taking as reference those elements that usually are considered as characteristic of a proper organization and proper management not just a building company but, more generally, any type of company.

We divided each of these sections into smaller elements, the expression of different areas which could, at least ideally, represent the two aspects of the subject.

The result was the splitting of the two sections, and therefore the check-list, on different hierarchical levels so as to determine the weights to be attributed to individual variables, each incorporating a number of elements, all pertaining to the same area of investigation and gradually becoming narrower.

These variables were then subject to verification of **congruence** (ie capacity to represent the object of analysis), **significance** (to verify the actual usefulness of these variables in determining the rating), and **relational checks** (in order to group similar variables and avoid repeatedly analyzing the same variable).

Questionnaire

The other essential tool of investigation used in our model is the questionnaire, the implementation which could only take place after completing the checklist.

The choice of the questionnaire is obvious: only the entrepreneur responding to the questions could provide the information necessary to draw a general picture of its modus operating within the company (Senge, 2006).

Among the various alternatives available, the choice finally fell upon a **semi-structured questionnaire** in which the questions are defined and not the sequence.

For the development of this instrument, however, we tried to allow those who had the task of providing a response to individual questions to be clear about the objectives of the questionnaire, trying to avoid any ambiguity that could undermine the effective validity of responses provided by the entrepreneur (Strebel, 2003).

The purpose of the questionnaire was to outline the profile of the manager in order to determine strengths and weaknesses.

We identified three distinct profiles: **blue, red and green.**

We interviewed the manager and we asked to choose 10 points in the questionnaire below that best represents them.

Table 2: Questionnaire

	PROFILE BLUE	**PROFILE RED**	**PROFILE GREEN**
HARD SKILLS	I comparisons I am practical I want evidence I arrive at a conclusion quickly	I am ordained and organized I love the details and facts I love order and categories I seek accuracy and precision	I make many suggestions I always find a solution I have different points of view I turn problems into opportunities
SOFT SKILLS	I am emotionally involved I am guided and I am motivated by personal values I am passionate I take commitments	I am sensitive and aware of other people I love to listen and observe I find posts and interpretations beyond words I relation with other people	I follow my instinct I imagine the future I explore possible scenarios I have new ideas
PROFILE DESCRIPTION	Not interested in the details. Very quick decisions. He lives in the past. **Characterizing words:** Action, Rating, Passion Comparison **Characterizing question:** WHAT IS RIGHT?	Interested in analyzing data to examine the situation. Not quick decisions. **Characterizing words:** Application, Facts, Information, Communication **Characterizing question:** WHAT IS TRUE?	He lives in the future. Innovator. **Characterizing words:** Ideas, improvement, change, vision **Characterizing question:** WHAT IS NEW?

Here below is an example of a profile to apply to our case study in which we marked the chosen options.

Table 3: Example of Answers to the Questionnaire

I comparisons (8) I am practical I want evidence I arrive at a conclusion quickly (2)	I am ordained and organized I love the details and facts I love order and categories (9) I seek accuracy and precision	I make many suggestions I always find a solution (10) I have different points of view (6) I turn problems into opportunities

I am emotionally involved (4) I am guided and I am motivated by personal values I am passionate I take commitments	I am sensitive and aware of other people I love to listen and observe (5) I find posts and interpretations beyond words (1)	I follow my instinct I imagine the future (3) I explore possible scenarios (7) I have new ideas

At this point we constructed the table of points that defines the details of the profile.

Table 4: Example of profile

	PROFILE BLUE	PROFILE RED	PROFILE GREEN	TOTAL
HARD SKILLS	2	1	2	5
SOFT SKILLS	1	2	3	6
TOTAL	3	3	5	

In the specific example the resulting "optimal" profile is a GREEN characterized by a balance of HARD SKILLS (with shades of both blue and green) and SOFT SKILLS (with shades of red profile).

Obviously the profile is just one example related to our case study.

There is no fair result and no wrong result (Igbavia and Baroudi, 1993). The validity of detection of the profile is closely linked to the specific needs (identified through the check list).

3.3. Construction of the structure of the model

The use of methods of multicriteria analysis is explained in the above-mentioned requirement to obtain a numerical assessment on the representative of the various components, highly qualitative.

Our aim is to construct a scale of priorities among the various actions that employers should put in at to improve the condition of its company (De Felice, Falcone, Silvestri and Petrillo, 2005).

We first define our problem (Initialization).

Made the checklist and the questionnaire corresponding to the retrieval of information, it was immediately noted that not all items were marked as important for achieving a better quality rating.

It was necessary to develop a methodology that allowed us to highlight the differences, especially in a non-numerical exaggerate disadvantage in the trial, companies which did not note abundant elements not considered of primary importance (Saaty and Peniwati, 2007).

In the end, we chose the technique ANP - Analytic Network Process.

The Analytic Network Process allows to outline a problem in network. This is one of the salient features of this technique. The methodology is particularly useful for the calculation of the weights to be attributed to the individual elements that define the problem.

The Analytic Network Process is a methodology which is a generalization of Analytic Hierarchy Process, (Saaty, 2001) a method to aid decision based on multiple selection criteria (MCDA, Multi-Criteria Decision Aid) developed by Thomas Lorie Saaty in the late 70s (Saaty, 2005) in fact, similar to the theory of AHP, the ANP is a multi-methodology, which is used to obtain scale of priorities by individual assessments.

Unlike the usual yes-no, by the logic or 0-1, the APN is a logical multi assessment.

The scale used in ANP allows different intensities and establishes priorities that indicate a range of possibilities for our preferences, compared to the classical zero (not preferred) or a (preferred) in traditional logic.

A single number is used to represent an evaluation of preference between two elements.

To date, there are many examples of applying the method to problems of evaluation in many different areas.

The first step of the methodology involves the construction of the network decision-making.

A network is a structure comprising multiple alternative decision-making and the objective assessment of general or goal.

All components of the network are compared in pairs with each other.

This comparison (Barzilai, 1998) is made in order to determine which is more important in relation to the overriding and to what extent the result of the comparison is the so-called dominance coefficient Aij, which represents an estimate of the dominance of the first element (s) compared to second (j).

To determine the values of the coefficients Aij it is necessary to express opinions of the elements compared (Leskinen, 2000). The ratings are expressed through the *Semantics scale of Saaty*, which brings together the first nine integers with as many opinions which express, in qualitative terms, the possible results of the comparison.

Table 2: Semantics scale of Saaty

INTENSITY OF IMPORTANCE a_{ij}	DEFINITION	EXPLANATION
1	Equal Importance	Two activities contribute equally to the objective
3	Moderate importance	Experience and judgment slightly favor one activity over another
5	Strong importance	Experience and judgment strongly favor one activity over another

INTENSITY OF IMPORTANCE a_{ij}	DEFINITION	EXPLANATION
7	Very strong or demonstrated importance	An activity is favored very strongly over another; its dominance demonstrated in practice
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation
2,4,6,8	For compromise between the above values	Sometimes one needs to interpolate a compromise judgment numerically because there is no good word to describe it

The analysis conducted by the two survey instruments allowed us to identify the network that best describes our goal, or the improvement of entrepreneurial skills.

Furthermore, the comparison in pairs between the various components allowed us to assign weights to the individual elements of the checklist.

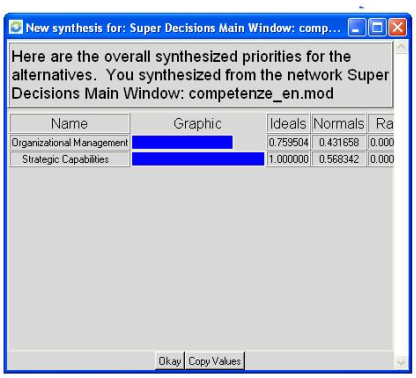


Figure 3: Priority vector - Synthesis Analysis

The analysis showed that the most critical parameter to monitor is the **strategic capabilities**, or a "good" manager should have an overall vision to be able to react to events and predict the future.

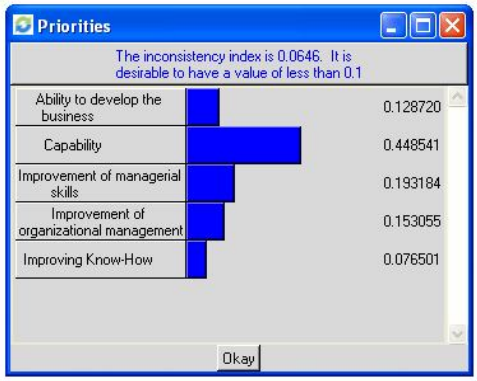


Figure 4: Index of inconsistency for the cluster Alternative

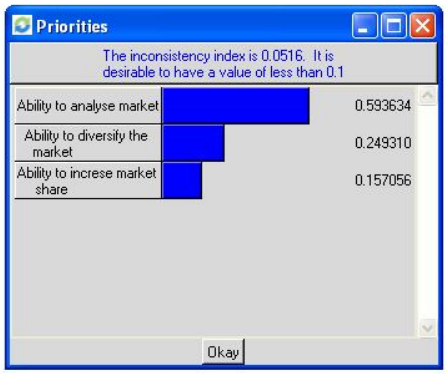


Figure 5: Index of inconsistency for the cluster Ability to develop business

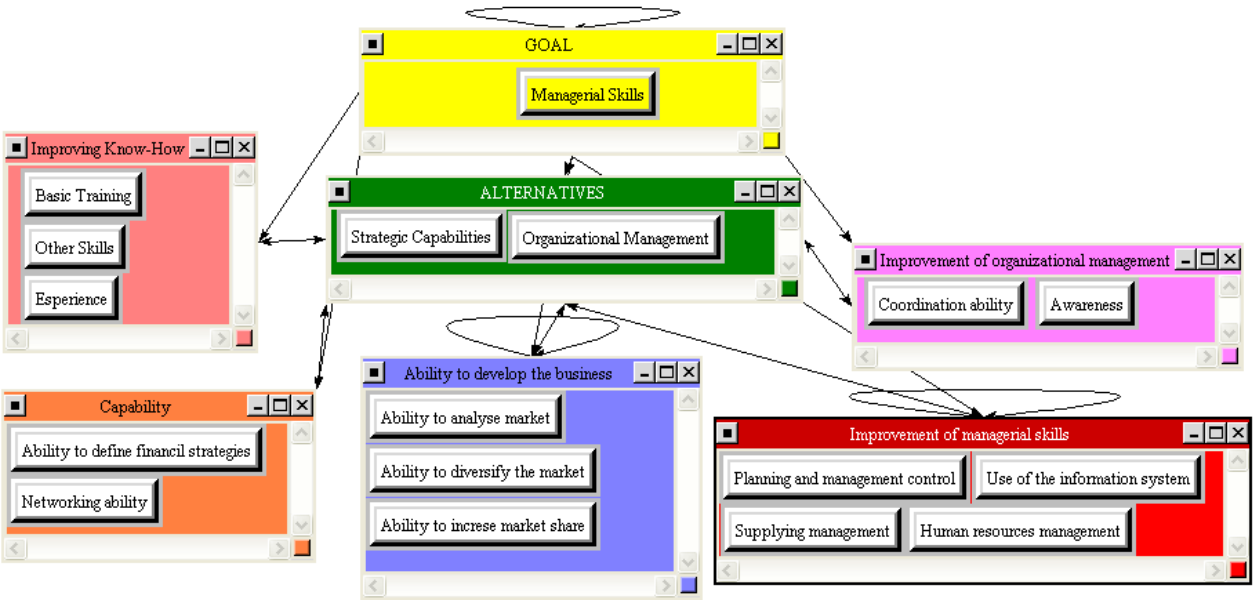


Figure 6: ANP Model

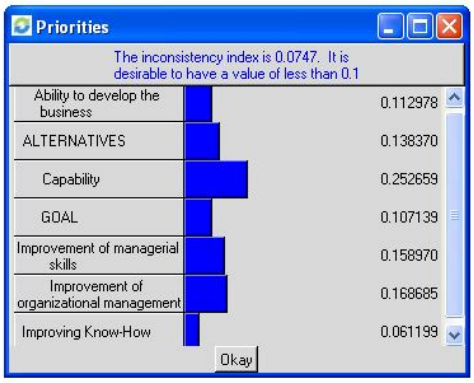


Figure 7: Index of inconsistency for the cluster Goal

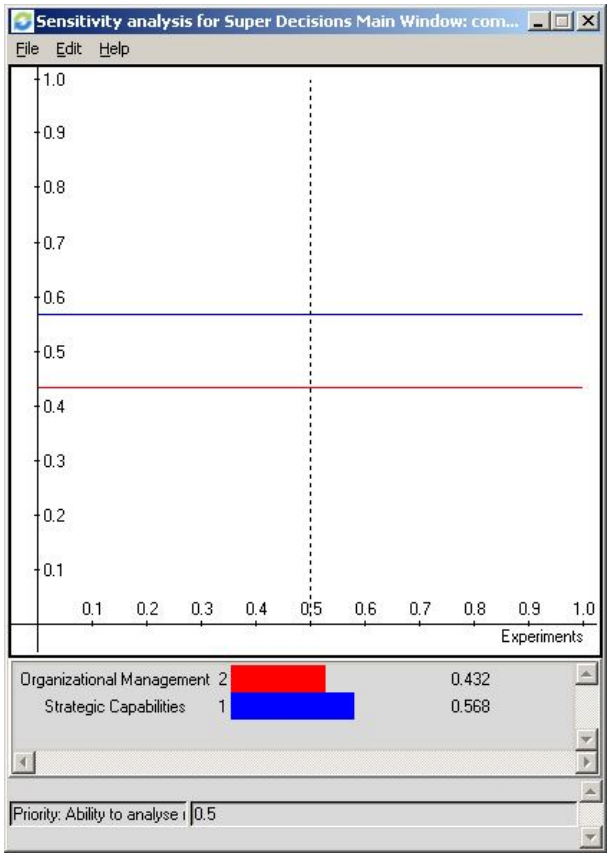


Figure 8: Sensitivity Analysis

The results show the prevalence of the human factor, and thus entrepreneurial skills, the organization formalized.

This is, in effect, an entirely predictable outcome considering that the scope of analysis is that SMEs in the engineering industry, rarely has formal organizational structures and its core is the figure of the entrepreneur.

The choice of weights, therefore, appears to be fully in line with the current situation of the sector, and of the Italian enterprise in general.

It is also clear that the amount of weight could certainly change if the size of the company changed.

Owing to the company size, the wise use of a distinction between different roles is a primary and, likewise, the means of planning and management become increasingly sophisticated.

Another element that certainly could influence the choice of weights is the story of the companies.

Certainly, in new companies, the spirit of new entrepreneur is the key driver and the main resource, in addition to the capital available for the development of the company.

Ultimately our aim was to construct a scale of priorities among the various actions that managers should put in act to improve the condition of its company.

We did not want, therefore, simply to compare different alternatives, but we highlighted the more or less importance of the variables we took into consideration when building the model.

All items in question were selected because they were considered significant in the determination of a trial on the qualitative aspects of the company.

3.4. Identification of indixes

To obtain a synthetic view on the qualitative aspects first reported and to allow benchmarking between businesses, similarly, the most effective seems to be to try to express that same opinion at issue in numerical form.

This is what we tried to do in our model, in the manner that is exposed below.

In summary, since the measure of performance is an issue for the company as a whole, at this point we completed the model by introducing some of the indices.

The indicators chosen are those considered significant in the determination of a trial on the qualitative aspects of the company, that is:

PG Index – Global Productivity (ie, system performance).

The PG index compares the change in production (index of quantity of production) with the variation of inputs (index of quantity of inputs).

The measure of "global productivity" is a partial measure of the performances of a production system, which must include additional parameters such as innovation, flexibility, quality, service.

In particular, we considered it appropriate to introduce the other two indexes.

EA Index – Effective Company Action (ie, make the right decisions).

The EA index is a measure of business investment in R & D than the market share obtained.

Finally, another index is introduced:

RA Index – Company Profitability (or ability to generate resources).

This index represents the net operating margin Report / Third party interest + Capital Equity.

It is also obvious, however, that not all these elements have the same importance for the determination of the proceedings in question.

Hence arose the need to create a system of weights that would make it possible to highlight and measure the global performances of a company.

In this respect we introduced evaluation criteria based on the scores tied to annual performance improvement that the company will record.

In the tables below you can examine the type of analysis chosen by us.

Table 3: Evaluation criteria

Parameter	% Annual increase	Score
PG Index	15-20	100
	25-35	150
	> 35	200
EA Index	15-20	70
	25-35	90
	> 35	110
RA Index	15-20	60
	25-35	105
	> 35	115

Table 4: Rating globale

Judgement	Range	Global Assessment
Low	0-230	**"Managing unsuitable"** Little ability to promote effective and efficient. Little aptitude for decision making
Medium	230-345	**"Suitable management"** Good management skills
High	345-425	**"Excellent management"** Synergistic vision of the various business aspects Ability to choose between the best alternatives

The control of corporate performance aims to bring the company to improve the ability to adjust its performance.

Establish objectives, budgets, operational plans, and then measure the performance.

We need a monitoring system to measure the degree of achievement of objectives.

This check will cover the overall result for a given period of time.

The monitoring of performance is used to influence policy making and implementation of the objectives only indirectly.

It sets general objectives that should be borne in mind when you take decisions.

There is an increasing need for systems of performance measurement able to give proper emphasis to the actual mix of inputs.

Often the common understanding of performance is compared to the efficiency of direct employment, which tends to take marginal importance compared to other inputs, its a "global productivity", as the productivity of materials, labour and indirect capital invested in stocks and resources.

Monitoring the performance allows us to consider:

1. The effectiveness of the company, that is doing the right thing.
2. Efficiency, that is optimize the ratio of resources consumed and results achieved;
3. The quality, the complex systems used by management.
4. Productivity or the ratio between input and output.
5. The welfare of those who work in the company.
6. Innovation.
7. Profitability.
8. Adaptability, the ability to address business changes.

4. CONCLUSIONS

This work is based upon the awareness that the success of an enterprise can not be the result of chance, it must rest on a solid foundation of sound management and organization on one hand, and on adequate knowledge and important skills.

Therefore we need an accurate and systemic interpretative model and measure performance to achieve the desired results.

The aim must be the development of measurement techniques and strong rating, able to grasp the "ranking" effective interest alternatives, and then measure the degree of proximity in respect to an ideal or satisfactory condition.

Indeed, while stressing the importance of good management, however adequate managerial tools and methodologies tailored to specific needs do not exist.

Therefore tools are needed to determine the skills gap (that is, training needs) that companies must overcome to improve their level of competitiveness.

The present work emphasizes the importance of an assessment that goes beyond the current methods usually applied.

We have therefore developed a model that integrates to qualitative as well as quantitative methodologies.

In particular the ANP made it possible to bind the rigorous quantitative and subjective sensitivity aspects.

The trial network will be assumed that any system is analyzed as a set of events or rather the ANP method is based on the recognition that a system is simplified, by reducing its complexity to a series of ever-smaller components, placed in the network .

In this way it was possible to establish a numerical relation (in the sense of allocating priorities and relative weights) between elements of the network.

The ANP leads to the development of alternative forecast scenarios by which the analyst can imagine the trend lines for developing the system of the subject, from the choices and the strategies selected by the relevant actors.

The ANP allows us to build a network between issue and areas of interest with respect to which the actor can " optimize" the process of allocation and planning.

Our intention was to create an assessment tool that could have a scope as wide as possible and then, at least in our expectations, independent of the type of company analyzed.

Ultimately we can conclude by saying that the model developed allows us to achieve excellence in various aspects of a business:

1. **Results orientation:** the Excellence is achieving results that satisfy all the stakeholders of the organization.
2. **Leadership and Constancy of Purpose:** The Excellence in leadership is visionary and inspiring, coupled with constancy of purpose.
3. **Process Management:** Excellence is to manage the organization through a number of systems, processes and interdependent and interrelated facts.

REFERENCES

Barzilai, J., 1998. Consistency Measures for Pairwaise Comparison Matrives. Journal of Multi-Criteria Decision Analysis.

Black, J.S., Gregersen H.B., 2002. Leading Strategic Change: Breaking Through the Brain Barrier. Financial Times Prentice-Hall.

Biggiero, L., Laise D., 2003. Comparing and choosing organization structures: a multicriteria methodology. *Human Sistem Management n° 22. IOS Press.*

De Felice, F., Falcone, D., Silvestri, A., Petrillo, L, 2005. Proposal of a New Decision Making multi criteria Methodology applied to the reengineering of layout in an electromechanical company: the Score Analytic Hierarchy Process. *Proceedings of 18th International Conference on Production research.*), July 31 – August 4, (Salerno ,Italy).

De Felice, F., Falcone, D., Duraccio, V., 2000. Proposal for the improvement of the design by means of an interfunctional team: The Project Quality Team. *Proceedings of ECEC2000 SCS-Conference*, April 17-19, (Leicester , United Kingdom).

Finan, J.S., Hurley, W.J., 2002. The Analytic Hierarchy Process: Can Criteria Be Ignored?. Computers and Operations Research.

Hult, G., Ketchen, D., Reus, T., 2001. Organizational learning capacity and internal customer orientation within strategic sourcing units. *Journal of Quality Management*.

Humphreys P., Ayestaren S., McCosh A., Mayon-White B., 1997. Decision support in organizational transformation. *Chapman & Hall*.

Igbavia M, Baroudi J. J., 1993. A short-form measure of career orientations: a psychometric evaluation. *Journal of Management Information Systems.*

Kanungo, S., Sharma, S., Jain, P.K., 2001. Evaluation of a decision support system for credit management decisions, *Journal of Decision Support Systems.*

Leskinen P., 2000. Measurement scales and scale independece in the Analytic Hierarchy Process. *Journal of Multi-criteria decision Analysis.*

Lofti V., Pegels C., 1996. Decision Support Systems for Operations Management and Management Science. *Irwin Publisher.*

Marakas G. M, 1999. Decision Support Systems in the 21st century. *Prentice Hall.*

Menkes, J., 2005. Executive Intelligence. Harper Collins World.

Olson D., Courtney J. F., 1997. Decision support models and expert systems, *Dame Publisher.*

Saaty, T.L., 2005. Theory and Applications of the Analytic network Process: decision making with benefits, opportunities, costs, and risks. *RWS Publications*, (Pittsburg ,USA).

Saaty T.L., 2001. Decision Making for Leaders: the Analytic Hierarchy Process for decision in a complex world. *RWS Publication*, Pittsburg (USA).

Saaty T. L., Peniwati K., 2007. Group Decision Making: Drawing out and Reconciling Differences. *RWS Publications*, (Pittsburg , USA).

Senge, P., 2006. The Fifth Discipline: The Art & Practice of the learning Organization. Doubleday.

Strebel B. J., 2003. The Manager's guide to effective meetings. *McGraw-Hill Publisher.*

AUTHORS BIOGRAPHY

Fabio De Felice, Researcher at the Faculty of Engineering of the University of Cassino, board member of several international organizations and responsible for scientific research and training in industrial plants.

Antonella Petrillo, degree in Mechanical Engineering, now is PhD Student at the Faculty of Engineering of University of Cassino where she conducts research activities at the Department of Industrial Engineering.

TWO HEURISTICS FOR IMPROVING THE EFFICIENCY OF A MARKOV CHAIN BASED DECISION MAKING METHOD

René Chelvier [(a)], Claudia Krull [(b)], Graham Horton [(c)]

[(a, b, c)] University of Magdeburg, Computer Science Department, Magdeburg, Germany

[(a)]rene@sim-md.de, [(b)]claudia@sim-md.de, [(c)]graham@sim-md.de

ABSTRACT

The paper describes two heuristics to reduce the number of comparisons necessary to reach a certain goal for a Markov model for multi-criteria and multi-person decision making. The motivation results from a demand observed in the early stages of an innovation process. Here, many alternatives need to be evaluated by several decision makers with respect to several criteria. With the implementation of the heuristics the number of comparisons necessary could be decreased significant. By reducing the evaluation effort necessary to reach a given goal, we will make the Markov-chain decision making method applicable to real world settings with a larger number of alternatives.

Keywords: multi-criteria decision making, reducing the number of necessary comparisons, heuristics

1. DESCRIPTION OF THE PROBLEM

We consider the problem of evaluating alternatives in the early stages of an innovation process. In this application area, alternatives need to be evaluated by several decision makers with respect to different criteria. There are possibly many alternatives that need to be considered; therefore it is necessary to make the evaluation process fast and simple. Consequently we address a multi-person and multi-criteria decision making problem (MPMCDM). The following example describes the intended application.

The early stages of a stage-gate process (Cooper 1988) often contain a large number of alternatives. In the very first stages all of them are described only with a title and a short characterisation. Because of limited resources an innovation team must identify only the top alternative to bring that forward to the next stage of the process.

Little or no quantifiable information is available about the alternatives in the first stages of a stage-gate process, therefore it is not possible to rank the alternatives based on objective criteria. Instead, only subjective impressions are available at this stage, enabling decisions of the form “A is better than B” with respect to a given criterion.

In Chelvier et al. (2008a) and Chelvier et al. (2008b) we described how to model the evaluation and decision process in the given application as well as how to build a complete ranking.

In this paper we address the large number of necessary comparisons. This drawback is a well known issue of pairwise comparison-based methods and sometimes referred to as the information overload problem (Dryer 1990). Our solution is suitable if only the top alternative instead of the complete ranking is wanted. The following questions need to be addressed: How to identify the top alternative without building a complete ranking? How to reduce the number of necessary pairwise comparisons to reach this goal?

2. BACKGROUND

2.1. Multi-criteria decision making

Multi-criteria decision making (MCDM) is a discipline aimed at supporting decision makers who are faced with the evaluation of many alternatives with respect to several criteria (Roy 2000; Hokkanen 1997; Belton and Stewart 2002). Depending on which type of result is needed, many different MCDM methods are available. Thirty available methods are discussed, for example, in Guitouni, Martel and Vincke (1998).

In multi-person decision making (MPDM), more than one person is involved in the decision making process. Because most MCDM methods assume only one decision maker, strategies for mapping several opinions onto a single result are needed (Meixner and Haas 2002; Vetschera 1991; Eisenführ and Weber 1994).

In the decision making method used in this paper, the decision makers perform independent partial evaluations which are subsequently combined to obtain an overall set of evaluations, which form the basis of the ranking computation.

In the field of MCDM many methods have been developed for specialised applications. Seven methods which can be used in the early stages of an innovation process are AHP (the Analytic Hierarchy Process) (Saaty 1980), WISDOM (Van Overveld 2003), IPC (Incomplete Pairwise Comparisons algorithm) (Harker 1987), Incomplete AHP (Caklovic and Piskac 2001), cost-benefit analysis (CBA) (Chakravarty 1987), ELECTRE (Benayoun, Roy and Sussman 1966) and MPMCDM with Markov Chains (Chelvier et al. 2008a).

However, two methods are not directly applicable to the intended task. ELECTRE deals with soft evaluation criteria, but this method needs all pairwise comparisons to compute the results. The CBA method needs measurable and quantifiable criteria to compute a valid result.

Accordingly, ELECTRE and CBA cannot be the preferred methods to evaluate alternatives under the given constraints.

2.2. AHP and WISDOM methods

The Analytic Hierarchy Process is based on a hierarchy of evaluation criteria, and uses paired comparisons of alternatives with respect to these criteria. Gradations in the comparisons are expressed using numerical values. The ranking of the alternatives is obtained from an eigenvalue computation on a suitably aggregated matrix.

Known drawbacks of the AHP method are the need to deal with inconsistent sets of evaluations, the large number of pairwise comparisons needed, a complex mathematical model which is intransparent to the user and questionable rankings resulting from innocuous individual comparisons. The incomplete AHP method tries to use a smaller number of pairwise comparisons without compromising the result. IPC and Incomplete AHP are based on AHP and inherit these drawbacks.

The Weightless Incremental Selection and Ordering Method (WISDOM) is also based on paired comparisons with gradations. However, in this case, the gradations are qualitative, rather than quantitative. The evaluation criteria are also weighted using pairwise comparisons. An algebra for computations on these qualitative relations is then developed for which an iterative method is used to compute a ranking.

WISDOM shares a lot of the positive attributes of AHP. However, it also has several advantages over AHP. The first is the use of qualitative expressions rather than numerical ones for the pairwise comparisons and the computations. It is also able to explicitly detect evaluation inconsistencies. Its main disadvantage – as with AHP – is the lack of a natural extension to multi-user applications. One such extension, named TeamWISDOM, was presented by Weber (2007).

Our method is similar to both AHP and WISDOM, in that it is also based on pairwise comparisons of the alternatives with respect to different criteria and that it subsequently computes a ranking vector. Like AHP, our method also performs an eigenvalue computation on a matrix representing the results of the individual comparisons. However, in contrast to both methods, our approach does not provide for differing weights in the pairwise comparisons (although this extension would be easy to accommodate), and its overall structure is simpler.

2.3. Markov chains

Discrete-time Markov chains (DTMCs) are well researched mathematical models with many applications in Science and Engineering. A DTMC is described by a stochastic matrix P and a probability vector π. The steady-state solution of the DTMC contains the probabilities of each of the system states and is given by the solution of the linear system of equations

$$\pi P = \pi \tag{1}$$

Markov chains are drawn as directed, annotated graphs, where the nodes represent the states and the arcs the possible state transitions. The weights associated with the arcs describe the one-step probabilities for each state transition. A state or set of states of a Markov chain is called absorbing, if it contains only incoming arcs.

We use the Markov chain-based decision making method, which was described in (Chelvier et al. 2008a and 2008b). Here, pairwise comparisons can be combined by weighting these according to the importance of the criteria and the decision makers, resulting in a discrete-time Markov chain. A random walk on this DTMC (Stewart 1994) models the decision process, where a longer state sojourn time implies a better alternative. The solution of the DTMC resulting from all pairwise comparisons then yields a complete ranking of all alternatives.

3. NEW IN OUR APPROACH

The Markov chain-based decision making method allows us to compute intermediate rankings during the evaluation process. Based on the intermediate rankings, we can use heuristics to estimate the effect of another comparison on the computed ranking. Comparing the effects of the possible comparisons, a heuristic suggests, which comparison should be done next, because it is expected to lead faster to the desired goal.

The goal is to evaluate the comparisons with biggest impact on the ranking first and the comparisons with lower impact on the ranking later. We assume that comparisons with high impact assign the right ranking position to the alternatives and comparisons with lower impact adjust the ranking position value but have little effect on the top ranking positions. Consequently, fewer comparisons are necessary to determine the top alternative.

4. HEURISTICS

We present two heuristics to estimate the next comparison with the highest impact on the intermediate ranking: *Maximum weight* and *Euclidean distance*.

4.1. Maximum Weight

The heuristic *Maximum weight* uses the information given by the initialisation of the evaluation. Therefore we denote the participants in the decision process by p_k with $k = 1 \ldots K$ and the decision criteria by d_l, with $l = 1 \ldots L$. We define a matrix A of dimension $K \times L$ whose coefficients α_{kl} satisfy $0 \leq \alpha_{kl} \leq 1$ and

$$\sum_{k=1}^{k=K}\sum_{l=1}^{l=L}\alpha_{kl} = 1 \qquad (2)$$

Thus the coefficient α_{kl} contains information about the level of expertise of Participant p_k with respect to Criterion d_l as well as the degree of importance of Criterion d_l, where larger values imply greater importance. From all open comparisons the heuristic chooses the one with the largest combined weight of Criterion d_l and Evaluator p_k.

Consequently, the comparison order for each participant is known by the facilitator at the beginning of the evaluation process.

4.2. Euclidean distance

The *Euclidean distance* heuristic compares the possible effect of all open comparisons, choosing the one that can produce the ranking with the largest distance to the current ranking. It is, however, still a heuristic, since the actual value of the comparison is not known; if alternatives A and B are selected to be compared next, the evaluator still has the choice to select either one of them as superior. We assume that the comparison, which results in the maximum *Euclidean distance*, implies the biggest impact on the ranking.

Compared to the heuristic *Maximum weight* the comparison order is not known at the beginning of the evaluation process. Here, after each evaluation we estimate the next comparison with a "brute force" algorithm for all open comparisons.

With these heuristics the evaluation process can be stopped if the top alternative is identified and is not challenged for a number of comparisons. The better the heuristic the less pairwise comparisons are necessary.

5. ADVANTAGES IN OUR APPROACH

The heuristics reduce the number of necessary pairwise comparisons to obtain the desired evaluation result in the intended application. In that case the named information overload problem of the Markov chain-based decision making method can be reduced. We notice that the simpler the evaluation goal the higher the possible savings of pairwise comparisons.

The benefit of this approach is most significant if soft evaluation criteria are involved. The method works from two alternatives to more than one hundred alternatives. The limiting factor is the necessary number of pairwise comparisons to reach a certain goal.

6. EXPERIMENTS

In our intended application it is sufficient to identify the top alternative. To study the effect of the heuristics we choose a real MCDM problem from industry: In an ideation workshop many product ideas for an automotive supplier are generated. After the idea generation the participants of the workshop identified those alternatives with the highest expected benefit for the company. After the workshop the best alternative should be identified. Therefore the innovation process provides evaluation criteria and roles of the decision maker. The initial situation is as follows:

- There are five decision makers as participants in the decision process: the General Manager, the Product Manager, the Marketing Manager, the Production Manager and the Sales Manager.
- There are nine new product alternatives as a result from the ideation workshop. Each alternative is described with a title, a short characterization and a list of pros and benefits.
- We have ten evaluation criteria: fitting to a megatrend, market potential, competitive situation, degree of innovation, strength of the unique selling proposition, potential sales volume, research and development costs, profit margin, proportion of value-added, proportion of strengths and weaknesses. In the following process we assume that all criteria are independent and we consider no feedback between the criteria.

Even though the criteria are not quantifiable with the given information, subjective impressions are available, enabling decisions of the form "A is better than B" with respect to a given criterion.

In preparation of our experiments we collected all comparisons from the decision makers in a personalised questionnaire. We asked each participant "You see the Alternatives m_1 and m_2. Which of them is better with respect to the given criterion"? Each participant had to carry out 360 pairwise comparisons. The number T of all pairwise comparisons is given by

$$T = \frac{K*L*M*(M-1)}{2} \qquad (3)$$

where M is the number of alternatives and every participant makes every possible pairwise comparison with respect to every criterion.

In the next step we use a program to simulate four evaluation workshops, each with a different strategy for the order of the pairwise comparisons. The four workshops are:

1. Random order (reference 1)
2. Facilitator order (reference 2)
3. Maximum weight (heuristic)
4. Euclidean distance (heuristic)

In workshop 1 we use a randomised order of the evaluations. We assume that randomisation is representative for all workshops with no order strategy. Additionally, the program simulates a simultaneous treatment of the participants (all at the same time). We use the random order strategy as reference for the experiments with the heuristics.

In workshop 2 we use an order strategy with minimum cognitive load for the participants. This means that the comparisons are ordered first by criteria,

second by alternative 1 and third by alternative 2. Consequently the evaluation criterion and one alternative are changing as few as possible in the evaluation process. We assume that this strategy is representative for a good facilitated real workshop. As in workshop 1 the program simulates a simultaneous treatment of all participants.

In workshop 3 and 4 we use the presented heuristics to order the pairwise comparisons. We assume that the heuristics estimate the next comparison with the highest impact on the intermediate ranking. The evaluation simulated a simultaneous treatment of the participants as well.

7. RESULTS

The four diagrams show the probability values of the DTMC solution vector over 1800 pairwise comparisons, the lines also correspond to the ranking vector of the alternatives.

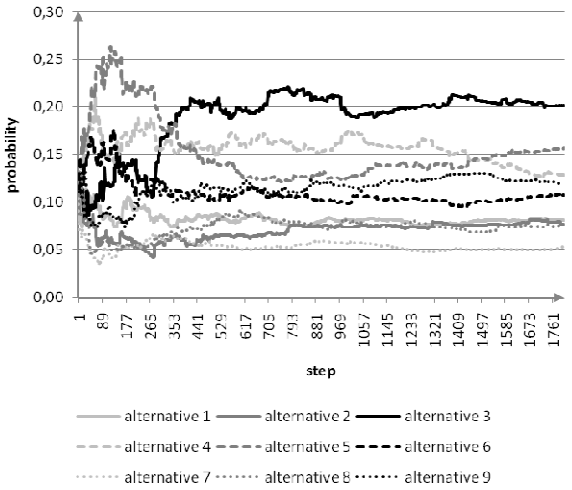


Figure 1: Order of Pairwise Comparisons is Randomised.

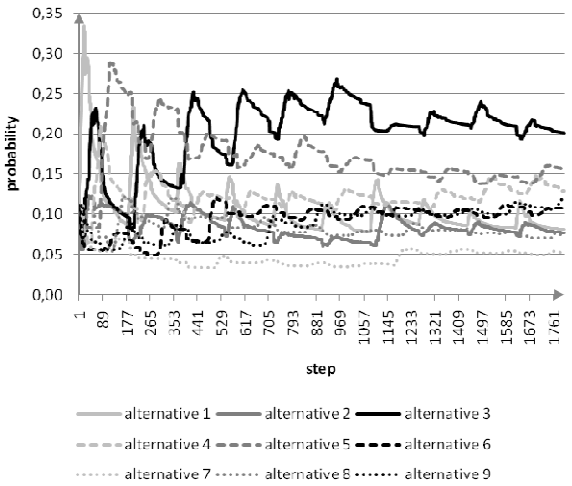


Figure 2: Order of Pairwise Comparisons for Minimum Cognitive Load (Facilitators order).

In Figure 1 the order of the pairwise comparisons is randomised. The best alternative is identified after approx. 350 comparisons. The second best alternative is identified after approx. 1530 comparisons. The late identification of the second best alternative shows the disadvantage of an arbitrary (in this case randomised) order of pairwise comparisons: Comparisons with a high impact might be done too late. One cannot be sure, when the evaluation process can be terminated even if only the top alternative is of interest.

In Figure 2 the best alternative is identified after approx. 560 comparisons. The second best alternative is identified after approx. 1700 comparisons. The late identification of the second best alternative is a disadvantage in that strategy, too. Additionally, the discontinuous progress of the values in the ranking vector makes the termination of the evaluation process very difficult even if only the top alternative is of interest.

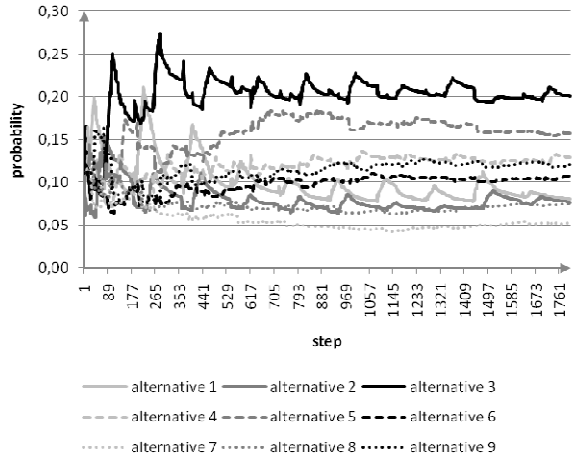


Figure 3: Order of Pairwise Comparisons is Computed with the Heuristic *Max Weight*.

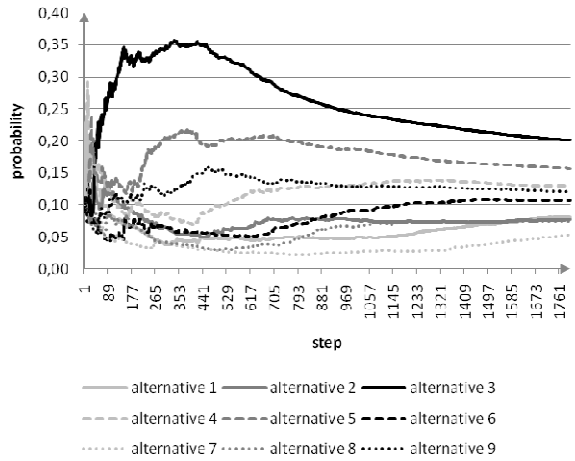


Figure 4: Order of Pairwise Comparisons is Computed with the Heuristic *Euclidean Distance*.

In Figure 3 the best alternative is identified after approx. 250 comparisons. The second best alternative is identified after approx. 450 comparisons. Later on,

ranking position swaps only occur in the lower ranking positions. Consequently the heuristic *Max weight* causes a positive effect towards the reduction of necessary comparisons in this example.

In Figure 4 the best alternative is identified after approx. 40 comparisons, the second best after approx. 200 comparisons. The heuristic *Euclidean distance* decreases the necessary comparisons significantly in this example. Another advantage of this heuristic is the smooth progress of the values in the ranking vector. Compared to the heuristic *Max weight*, trends might be more easily detectable.

In comparison to the reference strategies random order and facilitator order we can observe many improvements obtained from the heuristics:

1. The heuristic *Euclidean distance* avoids a ranking swap between the best and second best alternatives.
2. The heuristic *Euclidean distance* avoids the late ranking swap between the third best alternative and the fourth best alternative.
3. With the heuristics *Max weigth* and *Euclidean distance* the best alternative and the second best alternative turned out very early.
4. The heuristic *Max weight* clearly identifies the best alternative and the second best alternative. Unfortunately, the progress of the values in the ranking vector is slightly discontinuous.
5. The heuristic *Euclidean distance* identifies the best alternative and the second best alternative earlier than heuristic *Max weight* and much earlier than both reference strategies.
6. The progress of the values with heuristic *Euclidean distance* is very smooth so that trends are appreciable.
7. In contrast to the strategies random order, facilitator order and heuristic *Max weight* the values in the ranking vector of heuristic *Euclidean distance* converge continuously to the final values.
8. The strategy random order identifies the best idea with less pairwise evaluations than the facilitator order. Nevertheless, the strategy random order cannot be more than a reference because one cannot be sure when the evaluation process can be terminated.
9. With the reference strategy facilitator order participants need less rethinking between pairwise comparisons (less cognitive load) than in the other strategies. Instead they need more comparisons to reach a certain evaluation goal. The promised benefit is annihilated.

Fortunately, in this example, the brute force algorithm to compute intermediate rankings in the heuristic *Euclidean distance* takes not more than one second between each comparison.

8. CONCLUSION

The goal of the paper was to reduce the necessary number of pairwise comparisons if only the best alternative is needed. We used two heuristics to choose the next comparison to be made, in order to identify the top alternative as early as possible in the evaluation process.

As the examples show, the presented heuristics could reduce the number of necessary comparisons significantly. In detail, the *Euclidean distance* heuristic shows better results than the *Max Weight* heuristic. We assume that the use of intermediate results in the *Euclidean distance* heuristic is one reason for the better performance.

In our future work we want to identify more heuristics and implement the algorithm in a group decision support system. By reducing the evaluation effort necessary to reach a given goal, we will make the Markov-chain decision making method applicable to real world settings with a larger number of alternatives.

Further work will also include developing a method for enforcing irreducibility which retains sparsity, the implementation of comparisons of the form "much better than", comparing the results with those obtained from other methods, determining the intersubjectivity of the computed ranking among the decision makers and studying the behaviour of the method using data from more real-life problems.

Another question we would answer in further work is the definition of a definite stop criterion whereby the evaluation process can be terminated.

REFERENCES

Belton, V., Stewart, T. J., 2002. Multiple Criteria Decision Analysis, *Kluwer Academic Publishers*, Boston, Dordrecht, London.

Benayoun, R., Roy, B., & Sussman, B., 1966. ELECTRE: Une methode pour guider le choix en präsence de points de vue multiples. Note de Travail, 49. *Paris: SEMA-METRA International*, Direction Scientifique.

Caklovic, L., Piskac, R., Sego, V., 2001. Improvement of AHP method, *Mathematical Communications*, Supplement 1; 13-21.

Chakravarty, S., 1987. Cost-benefit analysis. *The New Palgrave: A Dictionary of Economics*. Volume 1, pp. 687-90.

Chelvier, R., Dammasch, K., Horton, G., Knoll, S.-W., Krull, C., Rauch-Gebbensleben, B., 2008a. A Markov Model for Multi-Criteria Multi-Person Decision Making. *Proceedings of the 5th International Conference on Innovations in Information Technology*, Al-Ain, United Arab Emirates.

Chelvier, R., Horton, G., Krull, C., Rauch-Gebbensleben, B., 2008b. Decision Making With a Random Walk in a Discrete Time Markov Chain, *Proceedings of the 6th Vienna International Conference on Mathematical Modeling*, Vienna, Austria.

Cooper, R., 1988. The New Product Process: A Decision Guide for Managers. *Journal of Marketing Management*, Volume 3.

Dryer, J. S., 1990. Remarks on the analytic hierarchy process, *Management science* 36(3), 249-258.

Eisenführ, F., Weber, M., 1994. Rationales Entscheiden, 2th Edition, *Springer-Verlag*, Berlin, Heidelberg

Guitouni, A., Martel, J.-M., Vincke, P., 1998. A Framework to Choose a Discrete Multicriterion Aggregation Precedure, *Defence Research Establishment Valcatier (DREV)*.

Harker, P. T., 1987. The incomplete pairwise comparisons in the analytic hierarchy process, *Mathematical and Computer Modelling*, 9(11):837-848.

Hokkanen, J., 1997. Aiding Public Environmental Decision Making by Multicriteria Analysis, *Report 72*, University of Jyväskylä.

Meixner, O., Haas, R., 2002. Computergestützte Entscheidungsfindung: Expert Choice und AHP: innovative Werkzeuge zur Lösung komplexer Probleme, *Wirtschaftsverlag Carl Ueberreuter*.

Roy, B., 2000. A French-English Decision Aiding Glossary, *European Working Group "Multicriteria Aid for Decision"*, Series 3, Number 1.

Saaty, T., 1980. The Analytic Hierarchy Process. *McGraw-Hill*, New York.

Stewart W.J., 1994. Introduction to the Numerical Solution of Markov Chains, *Princeton University Press*, Princeton, NJ.

Van Overveld, K., 2003. *WISDOM: A Weightless Incremental Selection, Decision and Ordering Method*, Lecture document from the Eindhoven University of Technology. Available from: http://www.idemployee.id.tue.nl/g.w.m.rauterberg/lecturenotes/ID-tools/wisdom_manual.pdf [accessed 25 May 2008]

Vetschera, R., 1991. Entscheidungsunterstützende Systeme für Gruppen: Ein rückkopplungs-orientierter Ansatz, *Physica-Verlag*, Heidelberg.

Weber A., 2007. Mehrkriterielle Verfahren zur Selektion und Bewertung von Ideen in Group Support Systemen, *Masters Thesis, Computer Science Department, University of Magdeburg*, 2007.

AUTHORS BIOGRAPHY

René Chelvier studied Computational Visualistics at the Otto-von-Guericke-University Magdeburg, obtaining his Diploma in 2005. He is currently working on his PhD in Computer Science at the same university.

Claudia Krull studied Computer Science at the Otto-von-Guericke-University Magdeburg, obtaining her Diploma in 2003. She spent an exchange year at the University of Wisconsin, Stevens Point, where she graduated in 2002. In April 2008 she successfully defended her PhD thesis at the Otto-von-Guericke-University Magdeburg and is currently working there as a research and teaching assistant.

Graham Horton studied Computer Science at the University of Erlangen, obtaining his Diploma in 1989. He obtained his PhD in Computer Science in 1991 and his "Habilitation" in 1998 at the same university, in the field of simulation. Since 2001, he is Professor for Simulation and Modeling at the Computer Science department of the University of Magdeburg.

EARLY TESTING PROCEDURES FOR SUPPORTING VALIDATION OF INTELLIGENT AGENTS FOR SIMULATING HUMAN BEHAVIOR IN URBAN RIOTS

Alberto Tremori[(a)], Enrico Bocca[(b)], Federico Tarone[(b)], Francesco Longo[(c)], Simonluca Poggi[(c)]

[(a)]MAST
[(b)]MISS-Genoa, DIPTEM University of Genoa
[(c)]Simulation Team

[(a)]alberto.tremori@mastsrl.eu, [(b)]{enrico.bocca, federico.tarone}@liophant.org,
[(c)]{francesco.longo, simonluca.poggi}@simulationteam.com

ABSTRACT
Over the last decade the authors have been working on several researches concerning the use of intelligent agents for managing simulation of critical situation in security, such as managing civil disorders, riots or terrorist actions. The adoption of intelligent agents to reproduce human behavior is a key factor to implement realistic models for security and civil protection. Human behavior is very complex phenomena to be simulated and it is important to anticipate final results to get consistent models. In this paper will be described the use of early testing procedures to anticipate results before the VV&A phase. It is, in fact, useful to adopt this kind of early testing activities in application of intelligent agents with human behavioral models.

Keywords: HBR, Intelligent Agents

1. INTRODUCTION
The human factor is the key element in every analysis: industrial and business, services, security and defense. Non conventional and guerrilla warfare scenarios, urban riots and civil disorders, global and local terrorism are the real challenges that security forces have been facing at world wide level in the past, currently and probably for the future. In these scenarios human factor is crucial and it's impossible to apply traditional doctrines in command and control as well in training.

The authors are involved in different research projects dealing with this kind of issues. In the last years, in particular, they were involved in the PIOVRA project (Poly-functional Intelligent Operational Virtual Reality Agents), developed for the European Defense Agency, and the Italian and French DoDs. An important follow up to this project is the demonstrator named RATS (*Riots, Agitators & Terrorists by Simulation)*.

The authors are currently developing other study in this field, exploring new areas such as ethnic and social impact on civil disorders management or integration of simulators with Command and Control (C2) systems.

In the present paper the early testing activities managed by the research team to consolidate models before the final VV&A phase will be described.

They provide an overview of the fine tuning of algorithms and models related to PIOVRA Project, consequently to test phase. The movement algorithms of agents, case studies and examples are herein presented.

The description of the movements is detailed and examples are given based on the early tests made on PIOVRA Model, describing the rules and strategies which produce different effects.

A description of the case study scenario and consequent examples of movements is detailed in the following of the present paper. As for PIOVRA, the algorithms and models have been tested and refined during VV&A.

2. PIOVRA ALGORITHMS: MOVEMENTS
The Movement Algorithms affect the Action Object; in effect PIOVRA units have specific positions in term of current zone and in term of Geographical Position X,Y: terrestrial curvature has been left out of consideration and the PIOVRA surface is completely plane. Also, Position attributes and movements are two dimensional.

The Action Objects move in PIOVRA zone based on their own characteristics, speed included, and they are able to find a target following a specific path made of a single final point and/or a sequence of points.

Such kind of moving take into account the object operative status: in this specific case, the effective speed is defined on the basis of Zone Object and Action Object characteristics and weather condition; the speeds allowed by the environmental factors are compared with the object proposed speed, due to the fact they can affect it. The Movement algorithms also compute time to destination and related path: each event related to turn-point arrival is determined.

The algorithms works by the following approach: the position of the moving object is continuously updated and the travelling time and path recalculated where needed (i.e. in case of alert)

Two different movement algorithms devoted to regulate Action Objects PIOVRA agents are used: HM-

Holonomic Movement (it is based on Zone Object Characteristics and devoted to move entities on the map identifying the Zone where the unit is operating)

VM-Vectorial Movement (it is based on Zone Object Characteristics and considers as reference for each movement the connections among the Zone Objects and road networks)

As for vectorial movements, the target function devoted to optimize the path is defined on the basis of distance, travelling time, risk factor (presence of dangers along the path).

Vectorial movement have been defined on two different strategies: Road Oriented and Zones Oriented. Each one applies the same algorithms, therefore the search set in Road Oriented is made up by roads and intersections, while in Zones Oriented the zones objects are used.

In such an approach, the path is identified thanks to available links between zones considering the zone average diameter as parameter for distance evaluation and the barycentric coordinates as reference target point in intermediate movements.

Multiple criteria for shifting from one movement type to another one can be defined, in the proposed study the following movement strategies were introduced:

- When starting and end points within the same zone Object the holonomic movement is preferred
- When the final target is out of the current zone Movement over a threshold distance value use as preference the Roads Oriented Vectorial Movement
- Movement under a threshold distance value use as preference the Zone Oriented Vectorial Movement
- Action Objects are characterized with preference for the use each of the movement strategies
- A general simulation parameter is used to define the opportunity to use each of the movements
- Action objects representing few people prefer to use vectorial movement
- Action objects representing large groups use as preference the holonomic movement

The preference levels and the opportunity levels are combined by fuzzy rules dynamically during simulation for deciding what algorithm to use.

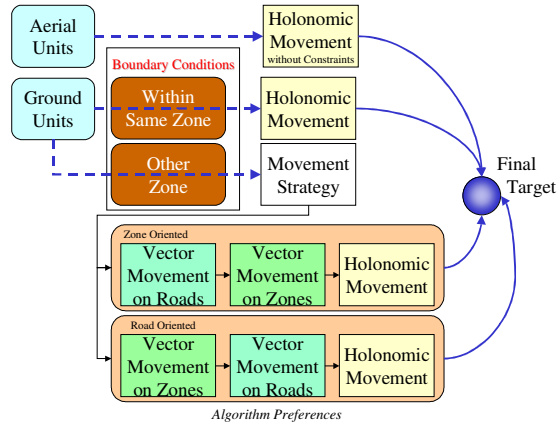


Figure 1-Algoritms Preferences representation

3. CASE STUDY

In this research the movement case study is related to a configuration of an Environment involving a town and the surrounding country side areas. In the described research, the movement algorithms are devoted to control action object moving inside PIOVRA Zone and each Zone Object is an entity that includes movement links and ground characteristics. Figure 2 is an example for the Movement case study and proposes a framework of links and zones.

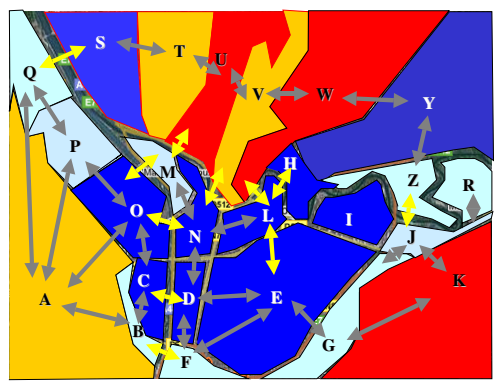


Figure 2: Zone Objects and Links

4. EARLY TESTING ACTIVITIES

The Regular Movement with Holonomic Approach in this research is summarized by the following figure:

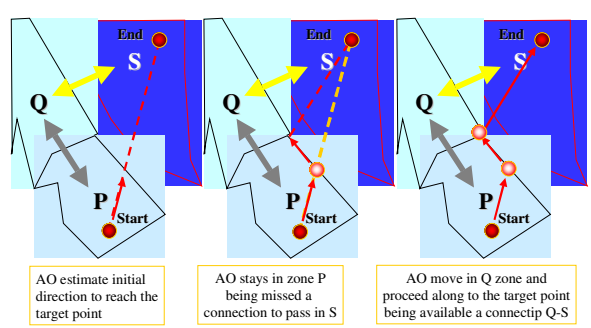


Figure 3 - Regular Movement with Holonomic Approach

The arrow between zones Q and S represent a bridge.

The Zone Object is related to the definition of the speed (and its limitations) based on the previously mentioned X,Y positions; different grounds provide different speed, for example urban speed for most vehicles is set to 10 km/h, while pedestrian speed in a forest is usually 2 km/h; therefore each Action Object have its speed profile that is related to the ground type and weather conditions; these movement characteristics are related to the coordinates of the Action Object and time respect the Set of zones.

The units move from a point to another one, based on an straight line changing the speed limitation due to current zone.

The passing from a zone to another one is enabled by the availability of an active link connecting the two entities (i.e. a bridge, or a free connecting border); the general procedure is summarized in figure 2 in a regular case.

It exist the possibility of special cases; for instance passing from zone x to zone y, corresponding to arriving in front of a river where the bridge represent the link between the two zone; a special event happen if at the arrival in front of the bridge, expected to be operative, it results to be closed/destroyed.

In this case, it is applied the following procedure: the unit checks the link (it is active or not) to proceed, then in case of the negative feedback, the unit try to continue to the target point along the border of the zone until it reaches a critical minimum distance point,

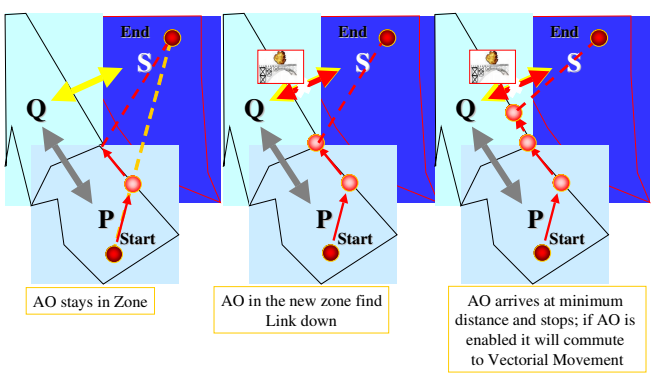


Figure 4 - Holonomic Movement in case of missed link

In correspondence of this phase the unit stops its movement and, if possible, it shifts to vectorial movement computing the path to reach the final destination.

For instance figure 4 present the same configuration of figure 3, but in this case the bridge connecting Q - S is down; in this case shifting to VM and choosing Zone Oriented approach the path is going trough P-O-M-U-T-S as summarized in figure5.

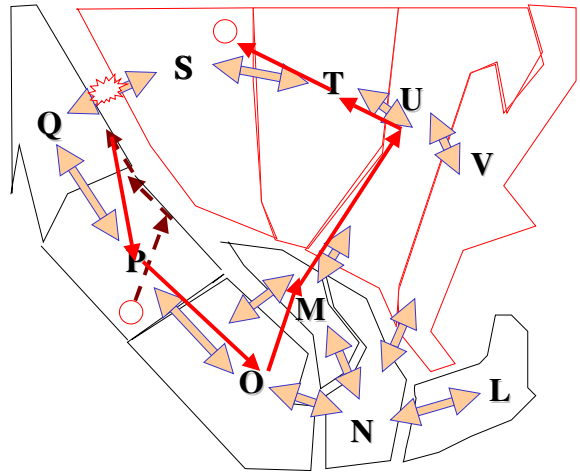


Figure 5 - Shifting from HM to Zone Oriented VM due to Zone/Road Environment Dynamic Evolution

In the figure 5 the directional arrows (□) represent the Links among Zone while the dashed line (brown) is the initial movement of the Action Object.

Considering the same example it is possible to evaluate the situation of figure 4 based on the hypothesis that a Road Oriented VM is used instead of a Zone Based.

In this case the path is proposed in figure 6

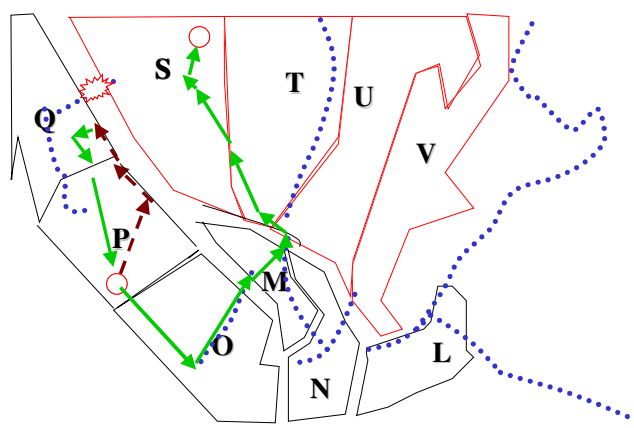


Figure 6- Shifting from HM to Road Oriented VM due to Zone/Roads Environment Dynamic Evolution

In the figure 6 the dotted lines (blue) represent the Road Network, while the dashed line (brown) is still the initial movement of the Action Object and the plain line (green) is the path obtained by applying a Road Oriented VM.

A second case is proposed about simulation evolution in PIOVRA OMF based on the hypothesis of combining different strategies on the scenario; in this case the target point is generated randomly inside PIOVRA zone.

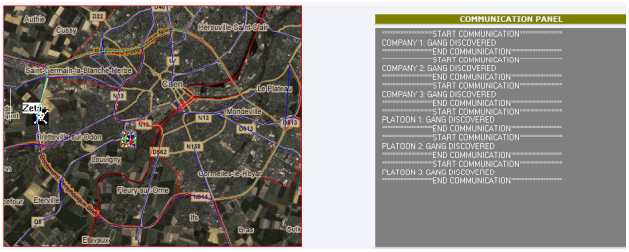


Figure 7 - Action Object Detection along Movement OMF

The reporting about detection of other entities is summarized in Figure 7 showing as its is represented the application area for PIOVRA Demonstration in the OMF; as action object is used a Gang moving with the profile of a Vehicle based unit with four people on board.

A set of runs using as preferable the Zone Oriented approach is proposed in figure 8; each time that the action object reach the target its mission is appointed to a new destination randomly generated.

A similar set of test is conducted using as preferable the Road Oriented VM is proposed in figure 9; also in this case each time that the action object reach the target its mission is appointed to a new destination randomly generated; based on this approach it was possible to compare the two method by long term simulation in term of travelling times respect original distances.

Figure 8 - Tracking Movement of a Gang driven by Zone Oriented VM in O

Figure 9 - Tracking Movement of a Gang driven by Road Oriented VM in OMF

5. PIOVRA ALGORITMS & PSYCOLOGICAL MODIFIERS

The psychological modifiers are applied to Action Objects and Comportment Objects; in this case the modifiers considered included the following set of parameters:

- Stress
- Harmony

The authors adapted a stochastic discrete event simulator devoted to simulate a Vessel Operation along its entire Life Cycle for testing the PIOVRA Algorithms related to stress and coordination capabilities.

As usual, it is necessary to applied proper procedures for human behavior modeling in order to guarantee the best ratio between cost/benefits by a ranking list; in effect data availability, fidelity improvement, impact on target functions are the key factors for identify the aspect where to focus; in this case the goal is to estimate the stress and coordination ability of the crew along operative time in relation to ship operations.

5.1 Case Study

In order to evaluate these aspects it was decided to adapt an available simulation model in reference to specific mission profiles and operation planning based on the Simulator ACASO (*Advanced Carrier Acquisition and Operation Simulation & Optimization*)

In this context vessel components are modeled by hierarchical objects; by this approach it is possible to explode each single ship plant (i.e. engine system) in its elements by creating multi-layers corresponding to different detail level; obviously among the ship component the crew represent a critical element.

As stochastic factors affecting the model was identified:

- changes in operative profile
- failures and breakdown both in term of:
- time intervals
- impact on the overall efficiency
- duration of maintenance operation
- cost of maintenance operation
- crew behavior

Due to the fact that most of these parameters need to be estimated by experts, beta distribution are extensively used for modeling stochastic entities; in this context the characteristic of the ship used correspond to a medium size aircraft carrier, while the detailed data are related to public domain data and hypotheses from PIOVRA team. It is interesting to focus on the psychological modifiers applied to the crew within this model; in fact it represents a very critical component of the ship; in this case the authors were interested in modeling this entity in order to estimate the impact of different operative profiles on the overall ship performance. The psychological modifiers affect the efficiency of the ship in term of operation and maintenance with different impact on sea operation respect port.

In figure 9 it is proposed the dynamic model introduced in the stochastic discrete event simulator for crew psychological modifiers (including the effect of crew turnover); the SMP (Small Maintenance Plans) and GM (General Maintenance) represent the condition of the ship for middle-life planned maintenance stops; while the signs inside diamonds represent the setting of the factors included in the algorithms for Stress and Cooperation Capabilities. In addition the model consider also the turnover of the crew (substitution of resources) in term of changing the status of the personnel (i.e. newcomers, experts etc.) based on a detailed composition of the crew mix.

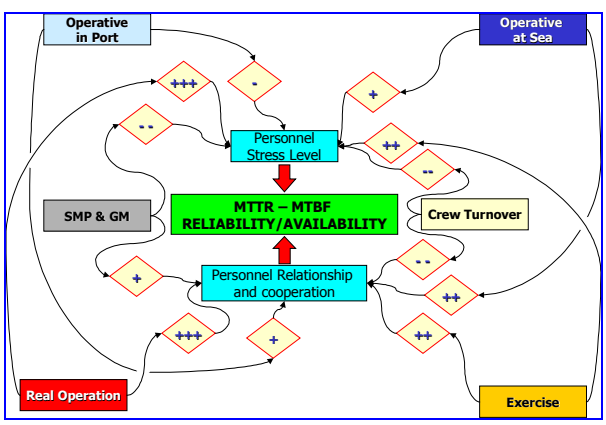
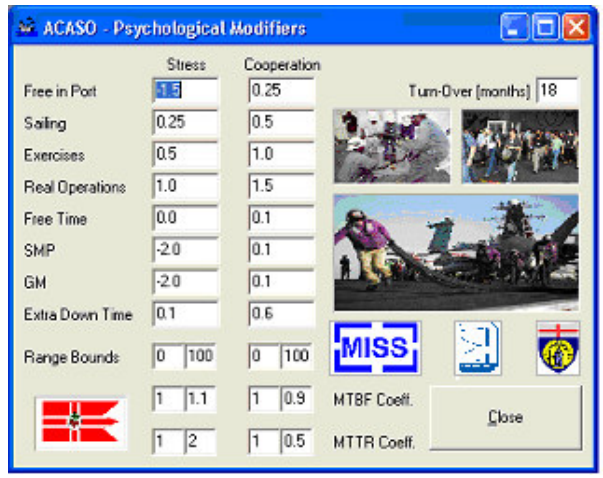


Figure 10 - Dynamic Model involving Human Factors affecting Ship Operations

5.2 Early Tests and Examples

The impact of these two phenomena on the crew has been defined in order to represent the evolution in its capabilities in relation to ship operative status and its history: for instance, as much as the crew is acting in real operations as much its stress level increases; therefore the relation for this increase correspond to a special trend that properly model psychological aspects such habits, limit points, etc.

In this case the setting of parameters are related to Dsc(x,y) Unitary Continuous Change in Stress Level due to x Vessel Status in y dynamic scenario environments:



These trends corresponds to the curves represented in figure 10 including the hysteresis due to the different behavior corresponding to stressing and relaxing the crew defined in PIOVRA Algorithms.

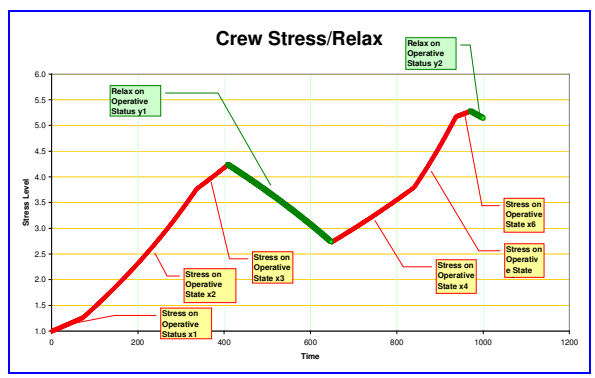


Figure 11 - Psychological Modifier Evolves based on their status, trend and ship operative Condition

In case of long term operations it possible to measure influence of crew coordination/stress in term of overall results; by running simulations it is highlighted the influence of the stress on the MTBF (Mean Time Between Failures) and MTTR (Mean Time To Repair); so by enabling the psychological modifiers it is modified the estimation of the overall

vessel efficiency in term of availability. It is interesting to underline the results obtained applying the PIOVRA Human Behavior Modifier models to a particular case, in order to obtain a pre-check of their correctness; in particular the HBM are applied to CALIPSO© Simulator, a software developed in order to analyze parameters of a new carrier generation. The results obtained focused on the comparison between the availability temporary evolutions, one without HBM and one with HBM.

This comparison, represented in figure 11, shows how the availability results obtained without Human Behavior Modifier is better than the one with HBM but after a period of time the situation is upside-down due to the improvement of crew harmony. Please note that, even if punctual values have been modified for security reason, the trends are clearly reproduced. This means that the application of PIOVRA HBM models on CALIPSO© Simulator have allowed to test them on a different scenario and to analyze them in order to make the final tuning on PIOVRA easier and faster.

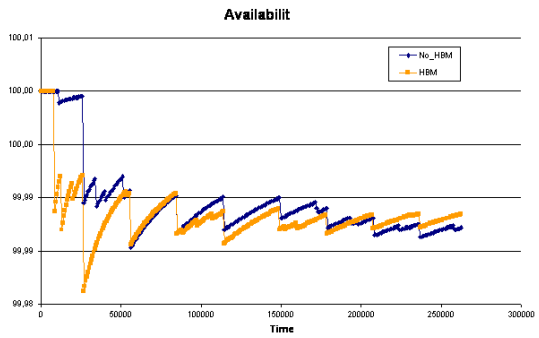


Figure 12 – Comparison between Availability without HBM and with HBM

6. CONCLUSIONS

In this paper have been presented the description of the early testing procedures for guarantee that during the definition of advanced algorithms and complex conceptual models it could be quickly redirect the development due to the feedback from dynamical verification and validation; in fact the dynamic VV&T is a critical issue, however often the implementation delays reduce the possibility to benefit from this approach; by the proposed methodology it becomes possible to anticipate these activities and provide immediately a direct feedback to model experts and developments; in fact the paper propose early test for movement algorithms used by PIOVRA Agents and basic case studies and examples devoted to complete early dynamic VV&T. The description of these case studies, tests and examples has supported the dynamic demonstration the behavior of agents during early development phases. It was possible to see how they react properly to the unforeseen (i.e. the crossing of the bridge) and to the differences among movements.

ACKNOWLEGEMENTS

The authors would like to thank Italian MoD, French MoD and EDA for sponsoring the project; the authors thanks also Prof.Agostino Bruzzone for its unique contribution to this initiative, as well as Prof.Claudia Frydman, Norbert Giambiasi and Roberto Mosca.

REFERENCES

Amico Vince, Guha R., Bruzzone A.G. (2000) "Critical Issues in Simulation", Proc.of SCSC, Vancouver, July

Banks J., Carson J., Nelson B. (1996), "Discrete-Event System Simulation", Prentice Hall

Bocca E., Bruzzone A.G., Rocca A. (2006) "Algorithms devoted to Reproduce Human Modifiers in Polyfunctional Agents", Proceedings of SCSC, Calgary, August

Bonabeau E., Dorigo M., Theraulaz G. (1999) "Swarm Intelligence: From Natural to Artificial Systems", Oxford Univ Press, Oxford

Bruzzone A.G., B. M., B. C., Procacci V. (2001) "FLODAF: Fuzzy Logic Applied to a Multi-Sensor Data Fusion Model", Proceedings of FLODAF2001, Montreal, August 7-10

Bruzzone A.G., B.C., Simeoni S., B. M., Orsoni S. (2003) "Simulation as Enabling Technology for Supporting Swarm Intelligence Applications", Proceedings of ASTC2003, Orlando FL USA

Bruzzone A.G., Cantice G. (2008) "Software Agents Italian Army Developments and Researches", Technical Report GESI, Wien

Bruzzone A.G., E.Page, A.Uhrmacher (1999) "Web-based Modelling & Simulation", SCS International, San Francisco, ISBN 1-56555-156-7

Bruzzone A.G., Giribone P. (1998) "Decision-Support Systems and Simulation for Logistics: Moving Forward for a Distributed, Real-Time, Interactive Simulation Environment", Proc. of the Annual Simulation Symposium IEEE, Boston

Bruzzone A.G., Kerckhoffs (1996) "Simulation in Industry", Vol. I & II, ISBN 1-56555-099-4

Bruzzone A.G., Massei M., B.M. (2007) "Piovra Report: Case Studies, Preliminary Test, Examples", PIOVRA Technical Report, October

Bruzzone A.G., Massei M., B.M. (2007) "Demonstration for Human Behavior Modeling Within Civil Disorder Scenarios", Proceedings of SCSC2007, San Diego CA, July

Bruzzone A.G., Mosca R., Massei M., Bocca E. (2008) "Critical Issues in HLA Integration: PIOVRA CGF vs. Wargaming System", Proceedings of Summer Computer Simulation Conference, Edinburgh, June

Bruzzone A.G., Petrova P., B.M., B.C. (2004) "Poly-functional Intelligent Agents for Computer Generated Forces", Proceedings of Wintersim2004, Washington DC, December

Bruzzone A.G., Roberto Mosca (2002) "Poly-functional Intelligent Operational Virtual Reality Agents", EUROFINDER PROPOSAL, CEPA 11, DIPTEM Technical Report, Genoa

Bruzzone A.G., Rocca A. B.M. (2006) "Movements and Attrition within Areas affected by Civil Disorders", Proceedings of EMSS, Barcelona, October

Bruzzone A.G., Tremori A., Massei M. (2008) "HLA and Human Behavior Models", Proceedings of MMS2008, Ottawa, Canada, April

Bruzzone A.G., Williams E.(2005) "Summer Computer Simulation Conference", SCS International, San Diego

Center for Technology and National Security Policy (2006) "Army science and Technology analysis for Stabilization and Reconstruction operations". October

Fishwick, P. (1995) "Simulation Model Design and Execution", Prentice Hall, New Jersey

Garrido De Ceita A.L.Torres, Frydman C. (2003) "Specification and Simulation of Reactive Knowldege Based Systems", Proceedings of SCSC2004, Montreal, July

Gozel, R. (2000) "Firepower Score Attrition Algorithms in Highly Aggregated Combat Models", Rand Technical Report

Hamilton John A. Jr., Nash David A., Pooch Udo W. (1996) "Distributed Simulation", CRC Press

Henscheid Z., Middleton D., Bitinas E. (2007), "Pythagoras: An Agent-Based Simulation Environment", Proc. and Bulletin of the Int. Data Farming Community, Issue 1 Vol. 1

Honing intercultural engagement skills for stability operations DARWARS Ambush! NK (non-kinetic). PEOSTRI.

Kennedy J., Eberhart R.C., Shi Y. (2001) "Swarm Intelligence", Morgan Kaufmann Publishers, NYC

Kuhl F., Weatherly R., Dahmann J., Kuhl F., Jones Anita, (1999) "HLA: Creating Computer Simulation Systems", Prentice Hall; ISBN: 0130225118

Kuhn H.W. (1997) "Classics in Game Theory", Princeton University Press

Lanchester F.W (1995) "Aircraft in Warfare the Dawn of the Fourth Arm", Lanchester Press

Lorenz M. (1993), "Object-oriented software development: a practical guide", Prentice Hall

Massei M., B.M., B.C. (2007) "Human Behavior Model for an HLA Federation", Proceedings of EUROSIW2007, Santa Margherita, Italy, June

Montgomery Douglas C. (2000) "Design and Analysis of Experiments", John Wiley & Sons

MORS - Agent-Based Models and Other Analytic Tools in Support of Stability Operations SAIC, McLean, VA 25-27 October 2005

Mosca R., Bruzzone A.G. & Costa S. (1996) "Simulation as a Support for Training Personnel in Security Procedures", Proc. of Military Government and Aerospace Simulation, New Orleans LA, April 8-11

Mosca R., Giribone P. & A.G.Bruzzone (1995) "AI Techniques to Optimize the distribution of Fire-Fighting Squads during an Accident", Proc. of Simulators International XII, Phoenix, April 9-13

NPS (Naval Postgraduates School) - SEED Center (Simulation Experiment Efficient Design). Harvest of proceedings and bulletin of the IDFC (International Data Farming Community): IDFW 15 & 16 Team 2 and Team 9. (http://harvest.nps.edu)

One SAF Objective System: Modeling the three block war. PEOSTRI.

Pingitore A. (2004) "Military Operations Research and Operation Iraqi Freedom", Thesis Report, Wells College

Zadeh, L.A., Kacprzyk (1992), "Fuzzy Logic for the Management of Uncertainty" Wiley & Sons, N

Zeigler B.P., Ball G., Cho H.J., Lee J.S. (1999) "Implementation of the DEVS Formalism over the HLA/RTI: Problems and Solutions", SIW Orlando, June

AUTHORS BIOGRAPHY

Alberto Tremori is member of the Simulation Team at DIPTEM University of Genoa and Liophant Simulation.

His expertise is simulators development for supporting process re-engineering, logistics, transportation systems and warehouse automation. From 1997 he was involved in several Simulation Projects (i.e joint project involving National Simulation Center, Istitute for Simulation and Training, Kennedy Space Center, McLeod Institute of Simulation Science). He currently serves as President of MAST.

Enrico Bocca achieved the Logistics and Production Engineering degree in 1999 with a thesis on "Automated Reorder System in Supermarket Chains", summa cum laude. During 2000-2001 he was involved in Customer Satisfaction analysis for Retail Business Sector. He completed during 2003 the thesis in Management Engineering working on "Innovative Technology Management in Retail Sector", summa cum laude. He participated to IEPAL experience coordinating an International Team of Engineering and MBA Students (from France, USA, Turkey) working in a Prefeasibility study related to Crowd Control in

Shopping Centers; this initiative was sponsored by FIPSE US Dept.of Education, DGEAC EC, LSC. He had experiences as consultant on different initiatives: Business Plans related to R&D Projects; Logistics and Project Management for CFLI. He was visiting researcher in Kennedy Space Center working on a joint project DIPTEM/NASA for Modelling Test Lab Crane Operations, the related demonstrator was presented officially in I/ITSEC2005 Orlando. He worked as researcher in Simulation Team of DIPTEM with special attention to Logistics. Currently he is charge as Responsible of M&S R&D for MAST.

Simonluca Poggi during June 2004 he participated to thre International Cultural and Academic Meeting of Engineering Students in Istanbul achieving 2nd Place with his Project on "Company Model Analyze Dynamic Reengineering of Joint Activity" in the International Competition with representatives of different Universities; while during 2005 and 2006 he participated as Team Advisor to ICAMES Conference with Student Teams.
He achieved the Management Engineering degree with a thesis on "Model for Life Cycle Analysis of A New Aircraft Carrier" from Genoa University
During 2006 he was involved in Customer Satisfaction analysis for Retail Business Sector.
He worked as consultant on different initiatives on ERP Solution for Retail; Inventory Management, Logistics and Project Management for DIPTEM.
He participated to several International conferences in Europe (i.e. HMS, I3M, EMSS, MAS etc.) Currently he is completing is PhD in Simulation Team of DIPTEM with special attention to Logistics Applications.

Francesco Longo was born in Crotone on February the 8th 1979. He took his degree in Mechanical Engineering, summa cum Laude, in October 2002 from the University of Calabria. He received his Ph.D. in Mechanical Engineering from the Department of Mechanical Engineering at University of Calabria in January 2006. He is currently researcher in the same Department. His research interests include Modeling & Simulation for production systems design and supply chain management. He was General Chairman of the 20th European Modeling & Simulation Symposium He was Program Co-Chair of the 4th International Modelling Mediterranean Multiconference and Program Co-Chair of the 19th European Modeling & Simulation Symposium. He belongs to the International Program Committee (IPC) of the Summer Computer Simulation Conference and Track Chair of the Emergency Simulation session in the same conference. He belongs to the IPC of the European Conference on Modeling & Simulation. He works as reviewer for the most important conferences on Modeling & Simulation such as the Summer Computer Simulation Conference, the European Conference on Modeling & Simulation, the Winter Simulation Conference, the Workshop on Applied Simulation, etc. He works also as reviewer for the international journal Computer & Industrial Engineering. He is Responsible of the Modeling & Simulation Center – Laboratory of Enterprise Solutions (MSC-LES). He teaches “Design of Production Systems” and “Laboratory of Industrial Plants” respectively for students of the 2nd and 5th year in Management Engineering. He is also tutor on the courses “Industrial Plants” and “Production Systems Management” respectively for students of the 3rd and 5th year in Mechanical Engineering.

Federico Tarone is Phd student at DIPTEM, University of Genoa. He is currently involved in some Simulation projects in the field of Logistics, Inventory Management and Web Applications. He is member of the Liophant Simulation. In 2008 he attended the I3M International Conference as speaker.

PANDEMIC DYNAMIC OBJECTS & REACTIVE AGENTS

Terry Bossomaier[(a)], Agostino G. Bruzzone[(b)], Marina Massei[(c)], David Newth[(d)], Joseph Rosen[(e)]

[(a)] CRiCS, Charles Sturt University, Australia
[(b)] DIPTEM University of Genoa, Italy
[(c)] Simulation Team
[(d)] CSIRO, Australia
[(e)] Dartmouth College, USA

[(a)]tbossomaier@csu.edu.au , [(b)]agostino@itim.unige.it, [(c)] marina.massei@simulationteam.com
[(d)] David.Newth@csiro.au , [(e)]Joseph.M.Rosen@dartmouth.edu

ABSTRACT

This paper proposes an innovative approach to pandemic models introducing agents for considering human behavior impact. The authors propose a methodology for developing these models as well as the related support features in relation to the open-issues generated by pandemic threats. Special attention it is proposed on providing a protocol for quick development and fast result achievement and concurrent strong attention to validation and verification .

Keywords: Pandemic, Health Care Modeling, Intelligent Agents, Human Beha

1. INTRODUCTION

Today there is a growing importance in quick model development protocols; in fact simulation it is often the most promising technique to investigate complex system and to face issues such as pandemics, however the development time required by this approach can result in critical and void its potential benefits.

Due to this aspects the authors propose a protocol, based on a lean simulation concept, for effectively and quickly modelling and analyzing pandemics

By the way in this sector, special attention is required for considering boundary conditions that heavily affect these phenomena such as human factors strongly affecting population reaction to possible countermeasures

So this paper proposes a protocol for exploring the dynamics of the spread of a pandemic by computer simulation, using a H1N1 influenza A virus as case study.

The authors start from analyzing an existing model such as EpiCast a computational model developed by the Los Alamos National Laboratory (LANL), used as part of the US National Strategy for Pandemic Influenza, and currently being used by the Department of Homeland Security to evaluate the current influenza situation in the US.

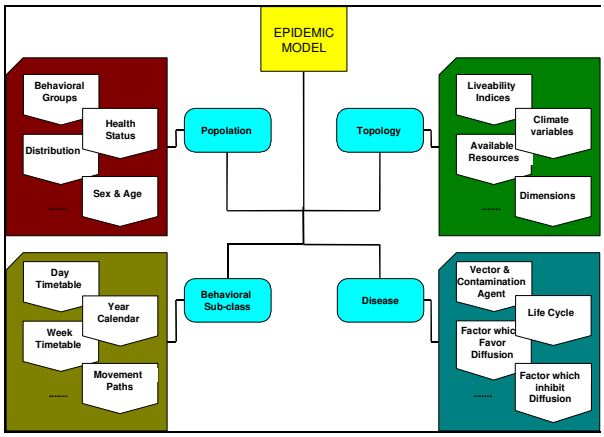


Figure 1 Example Epidemic Model Components

A critical aspect of models relates to the fact that basically they need really good data to be effective. In tbhis case this imlies:

modeling the virus itself, hopefully with additional information respect agent model and

modeling population (ethnographic) - knowing positively that someone is infected - it is necessary to go into detail on defining what data are available for defining the scenario

Therefore is important to model new elements in order to face current scenarios and to support decision makers in Government, Health Care and Institutions. In fact considering the Australian it is useful to extend existing models by including more realistic human behavior, which adapts as the pandemic spreads. It is also critical to develop advanced decision-support tools with enhanced data assimilation and visualization to support policy makers in formulating effective mitigation and intervention strategies.

Therefore the Australian case, given the population size of around 18 million represents an interesting and complex, but feasible example, for a nation size model.

Another very critical aspect it is the of supporting traii ning and accreditation in the use of the model by decision makers; today many models are providing pretty different

figures on these events, due to the fact that data are strongly changing the whole model behavior, so emphasis is needed on data certification and model development, consistently and concurrently, with the real system.

2. MODELING PANDEMIA

In a pandemic, there is a complex interplay between these factors but there are also significant variations introduced by the changing nature of the virus: in 1918 for instance, the virus had three waves

- mild
- moderate virulence
- severe virulence

but in the last phase it varied in its virulence from a low value of 2% kill to up to 20% kill of those that contracted the virus; this aspect it is similar in later pandemics 1957 and 1968 and more recently in the SARS; epidemic; recent varieties of H5N1 have had sometimes very high virulence - up to 50% killing and poor response to anti virals, in the case of H1N1 Australia, not only the virus can increase its virulence in waves, but it can also mutate in a way reudce the effectiveness of a vaccine an important considerations to be included in the epidemic model.

In an epidemic crisis - the starting point , the waves of the events - usually at least two, the end point are harder to determine. Therefore the epidemic evolution effects include, health, transportation, economy governance: national versus global.

The overall strength of the defense strategies at this time are depend on the resilience of vaccines and anti-virals. In the next wave with adverse mutations the virus will get more virulent maybe 5 to 10% mortality and could become more mutable meaning that a vaccine would not work and the virus could adapt to the anti virals so their effectiveness would go down. A virus that is causes no symptoms the several days, but is infectious - is hard to treat with current antivirals which are most effective in the first 48 hours.

Thus key is to track the population, put in place personal isolation measures (isolation, restricted movement, masks, and when needed quarantine) - this requires advanced modeling and simulation, that includes physical response to the virus, behavior of people in populations exposed (who goes to work who stays home) and, in the future, genomics to predict who will get sick , who will die, and who in the population will not get sick.

Specifically, the proposed approach focuses on countermeasure effectiveness with realistic boundary conditions and people, government, institution, companies reactions; for instance modeling the progression of disease respect quarantining and the impact of intervention measures considering human reactions and factors by introducing intelligent agents. In order to proceed in this direction it is necessary to address the following epidemiological questions: (1) efficacy and cost benefit of various mitigation strategies such as school closures; (2) expected level of absenteeism, and the impact of absenteeism upon commerce, industry, and the economy; (3) identify at-risk groups and mitigation strategies which will protect those groups; and (4) risks and returns (benefits) attached to implementing different mitigation strategies across state borders.

This research brings together a state-of-the-art high-resolution model of disease spread with cognitive science and advanced computational techniques to create decision support tools for policy development.

The idea to model every individual member of the population to create robust evidence-based policy is a very innovative concept proposed by the authors.

In fact the authors propose in order to speed up the development of new models to consider as first step forward the development of the following extensions to the basic EpiCast framework:

- Including more realistic, adaptive human behavior. Existing differential equation based models have used very simple, non-adaptive behavior when modeling individuals. Capturing individual human behavior has been highlighted as an important modeling task . The authors' aim is focused on capturing people's attitudes and likely behavior during an epidemic through an epi-ethnographic methodology. This information then needs to be introduced into the model by introducing agents generating the diversity of human behaviors.
- Decision-support tools. Another important aspect is related to the necessity to significantly improve existing decision-support tools; the analysis highlighted the following actions: (1) development of a data assimilation framework to create real-time forecasts of disease spread; and (2) application of the latest visualization techniques to generate easy-to-grasp movies and visual statistics of epidemic dynamics for better communication with and between health professionals, policy makers and the community.

Some of the authors (Newth) acquired experience focusing on the first use of EpiCast in Australia with Australian data; therefore it will be a significant step forward in term of innovation to create a simulation at the individual person level at a nation level including adaptive human behavior and extensive ethnographic work on pandemic perception over time.

So the proposed model offers an integrated wide approach to managing the emerging pandemic; as compared to smaller localized agent based simulations.

In addition, this model captures a higher level of detail than what is offered by differential equation based approaches .

3. RESEARCH INITIAL APPROACH

Early epidemic models were based around systems of ordinary differential equations, with unrealistic assumptions of large populations and perfect mixing. A breakthrough in modeling occurred with the work by Germann et al. in which the entire US population of 281 million people was treated individually as the agents within an agent-based model (ABM). EpiCast combines detailed census data with travel data and disease transmission dynamics to quantify the dynamics of epidemic spread. The model now forms an operational test-bed for the evaluation of quarantine, vaccination and anti-viral strategies.

The analysis suggest to focus on four major directions in order to develop new simulator:

3.1 Creating Reference Scenarios

For instant based on proper configuration of existing simulators (i.e.EpiCast). This includes data collection; estimation of key social contact parameters from time usage data; sensitivity analysis; and baseline scenario analysis. This phase provides an opportunity for developing detailed scenario analysis, as per Germann et al., on the Australian context. In this phase the authors are also addressing the impact of absenteeism on small-to-medium businesses, regional areas and smaller communities; social reliance and sustainability of regional areas with limited access to support services; and economic impacts per-sector. In order to achieve this result it is necessary construct a scenario development tool (SDT), Pandemic Simulation Language, to aid policy makers in formulating scenarios. The SDT creates scenarios to be read by the EpiCast engine, so this subsequently could becomes the interface used by government and health professionals.

3.2 Ethnographic surveys.

It is necessary to define procedures and protocol for data collection; in particular it is proposed to proceed with data collection of the behavioral model in two phases with two foci. The first consists of interviews, the second, a survey to refine and validate the beliefs and intentions gleaned from the first phase. There is a focus on general attitudes and specifically workplace issues, including surveys within companies.

a) Interviews have to be conducted across the following dimensions: urban versus rural; income category; family status; professional category. Interviews have also to be conduced across the business sector to ascertain attitudes and responses to absenteeism. From these interviews behavioral rules are formulated in the form: if infection level is high, I would not go to work.
b) The set of rules in (a) is assembled into a web questionnaire, collecting a significant sample, with the quantifier as a response, yielding a probability distribution for behavioral choice. Considering Australia as a possible case study, the comparison between Sydney and Melbourne could help to ascertain attitude changes as the pandemic spreads. A key objective here is to capture the change in attitudes and behavior as the pandemic spreads, through people's direct experience, social influences and the media.
c) Interviews with the general public have to be carried out during new model development to identify specific issues related to human behavior and for validation and verification purposes. Following on from this, interviews with business about absenteeism could be conducted. The set of rules inferred from the interviews have to be embedded in the new models. Sensitivity analysis of the full model is the proposed methodology to check consistencies and to support verification and validation. The human behavior model needs to be extended in order to consider the survey data.

3.3 New decision-support and visualization tools.

It is critical to develop two key proof-of-concept decision-support tools, exploiting:

a) Data assimilation, an ensemble of statistical and computational techniques for the fusion of model and real-world observations in order to create forecasts — in our case, of disease spread. Data assimilation techniques are extensively used by weather forecasting systems, but can be adapted to any observable system regardless of its complexity or non-linearity.
b) Effective visualization tools taking the complex detailed information created by the simulation and make it digestible and communicable to the public. For example, visualization of the EpiCast simulation results has been used by Ira Longini to communicate the likely future course of flu spread in the US to the general public.

3.4 Comparative Analysis:

The current H1N1 pandemic has already shown unexpected dynamics; Australia results to be a strongly affected country as it is confirmed by the high per person infection rate and the concentration of cases in Victoria. Thus insights into the spread dynamics can be obtained from simulating different social structures. Preliminary simulation results of Melbourne, Sydney and Brisbane suggest that this is the case. Public transport usage, family and community involvement differ between country by country and it is critical for future pandemic model to consider comparative analysis over different geographical regions and nations. Tthis research involves investigators active in different sites: Australia, USA and Italy. For instance, in Italy social mixing in the home features much stronger linking across generations than is common in Australia or USA. These aspects provide interesting suggestion in modeling the pandemic behavior in term of

identify open issues, factors to be considered and data to be verified and validated.

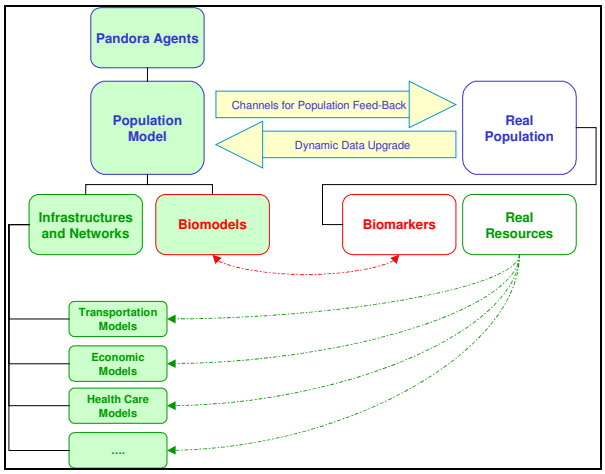


Figure 2. PANDORA General Architecture

4. DEVELOPMENT PROCESS

Basically the procedure is based on the concept to apply the following steps:

- Collecting data from on-going pandemic
- Add real data to the model
- Upgrade of the model
- Play into the behavior agents
- Comparing alternative choices simulating impact of resources costs, timer required for being activated/changes
- Evaluating impact of parameters that are estimated by SME such as casualties, deaths, severity of the disease, economic impact, impact on governance

In fact it is proposed the roadmap for facing a complex case study such as that one related to modeling H1N1 pandemic over Australia; in the following the authors provide a list of fundamental steps in order to achieve consistent results in this case in short time:

- Reconfigure Existing Models for the Case Study Australian context
- Design interviews and obtain ethics clearance
- Run existing models for a variety of contexts
- Develop extended model for interacting with decision agents reproducing human factors
- Carry out first phase of interviews
- Develop human factor agents
- Conduct second interview phase on company/workplace issues
- Run the new model including human behavior
- Analysis of socio-economic impacts
- Integrate workplace data and run full model

Each one of the above activities generates an outcome. The new configuration of existing models makes possible to simulate various scenarios from the basic original models.

Then the investigators, in close collaboration with officials, could run simulations and visualizations of spread and impact of outbreak under various scenarios. The reporting activity bring together the results of scenario analysis, ethnographic data collection and analysis. At this stage it is possible to carry out an evaluation of mitigation strategies with external stakeholders; following on from this, a detailed scenario analysis is conducted.

This models are expected to support several areas such as:

- Policy Definition and Regulation Revision
- Evaluation of Countermeasure Strategies
- Analysis of Quarantine Strategies
- Vaccine Management, Production and Distribution Policies
- Evaluation of Alerts and Tests Cost/Effectiveness
- Evaluation of new policies
- Effectiveness of Mitigation Strategies for reducing Virus Mutation Risk
- Definition Health Care Standing Operation Procedures
- Health Care System Reorganization
- Decision Maker Training
- Social and Economic Analysis

5. PANDORA

The main In addition to the proposed quick development process, it is suggest to develop a new simulation approach, defined PANDORA (PANdemic Dynamic Objects Reactive Agents), to incorporate the conceptual models and proposed approach within a common framework.

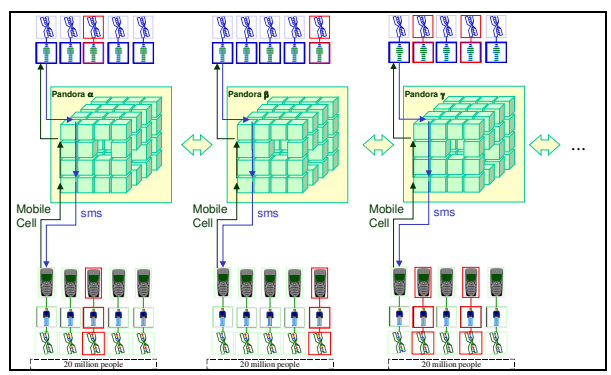


Figure 3a: Tracking each Individual by Clustering Regions

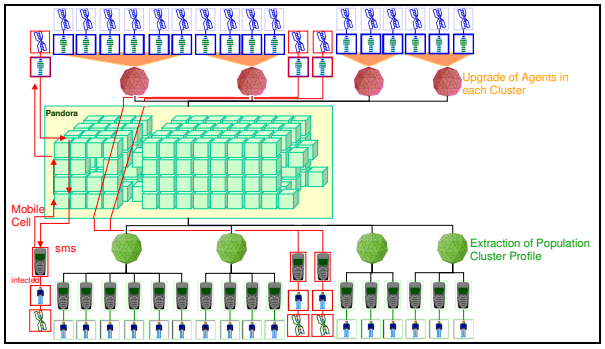


Figure 3b: Mapping Region Situation and Infected Tracking

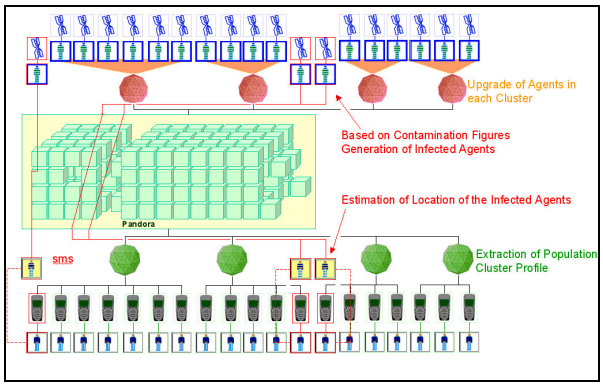


Figure 3c: Mapping Region Situation & Estimating Infection Source

PANDORA is envisaged as a decision support system based on simulation;to facilitate optimal choices before, during and after the pandemic, among the others these includes a number of key factors such as:

- security
- transportation
- food and water
- vaccinations
- antivirals
- Development of stand off detection systems
- personal distancing and protection
- masks
- quarantine
- isolation

For instance it is interesting to use these models to check if quick test with low cost and low reliability are more effective on the general behavior than more precise tests, considering people reactions.

The distribution of antivirals is itself a critical issue, currently being faced by the UK. There a web site and helpline provide immediate diagnosis and issue of antiviral from a collection depot. The effectiveness of this strategy will become clear in the coming months, but it is amenable to simulation.

In fact this represents a strong innovation directed to develop new simulators interoperating with different components in order to properly reproduce pandemic and to support countermeasures.

In this case the solution it is based on on-line stochastic simulation in a closed loop with live systems as proposed in the general architecture in figure 2:

- Population Models: Models of the individuals, groups and entities that correspond to the population all their characteristics
- Infrastructures and Networks: Modeling the environment, transportation, facilities and their characteristics
- Agents directing People Behavior and Human Factor
- Dynamic Data Upgrades: Channels for collecting on-line the situation of the people and their situation (i.e. mobile phones)
- Channels for Population Feed-back: Solution to provide directions to people in individual and group term (i.e. media, mobile phone)
- BioModels: Devoted to model genomics and diseases for reproducing their correlation for and supporting the estimation of probability to develop the infection as well as the relative evolution
- BioMarkers: Information allowing to identify individuals in term of their configuration, attitude and probability for being affected by specific infections

In fact by this approach it becomes possible to track the population; future experiments should direct to identify the proper granularity respect alternative solutions summarized in figure 3: in this case it is possible to have individual tracking within regions organized hierarchically (fig.3a), extraction of general parameters from the population and tracking of critical identified elements (fig.3b), or just group areas with injections of agents for estimating the most-likelihood situation of threats (fig.3c). It is necessary to point out that the three schemes are not alternative, in fact it is possible to think on combined solutions; for instance the above mentioned figures represent just tracking and communicating hypotheses through mobiles phones, however other communication channels could be characterized by different resolutions, in addition non cooperative behavior from individuals could necessiate a mixed configuration.

Therefore even with a so effective tracking of the population, if it will result possible by technical, legal and operational point of view, it is evident that the pandemic simulation needs to be able to look forward; these phenomena evolves, as already mentioned, quickly with strong inertia and influence of stochastic aspects (i.e. mutation, human decisions); for instance even instructions directed to single individuals or groups could be easily not respected: i.e. request to stay at home could be ignored by subjects due to underestimation of the related risks or due to personal interests in going to some relevant event (work or social activities), or even for fear to be framed in critical situation (in case of high mortality levels). So while the

proposed approach results very reactive and allows to keep on-line the simulator with the real situation, intelligent agents represent the critical aspect for being able to project over the future scenarios evaluating risks and possible impact of decision, events and stochastic factors on pandemic evolution.

6. HUMAN BEHAVIOR AGENT

Modeling humans represent a challenge; in pandemic this issue introduce the necessity to include at least influences such as: perception of the threats.

Among the others these agents need to demonstrate both rational and emotional capabilities; the intelligence of these agents is limited to operative decisions, and it should be simple enough to face the challenge of the problem size related to a national wide simulation. Specific features and aspects need to be defined, modeled and tested in order to verify the capabilities of this intelligent agents respect the need of the epidemic simulation.

In fact among the others the following measurable characteristics are required to PANDORA Human Behavior Agent combining intelligence and emotions:

- Capability to move and interact with other ones based on the current situation and individual profile
- Capability to be configured based on specific parameters: i.e. physical, social, ethnic, religious regional aspects
- Capability to change individual profile based on the situation evolution
- Capability to demonstrate awareness of the current situation
- Capability to demonstrate effect of phenomena as:
- Social Network Attractors (i.e. family, friends)
- Stress, Fear, Fatigue, Agressiveness Evolution
- Effect of the group/social contents on the behavior
- Effect of media and communications
- Effect of economic situation
- Effect of Physical situation
- Effect of Emotional Situation
- History affecting future behavior
- Hysteresis of human behavior modifiers
- Psychological State Evolution:
    - denial, anger, bargaining, depression, Acceptance
    - underestimation, fear, terror, habit
- Working Elements
    - Attitudes of employers towards absenteeism

It is important to note that in term of human behavior the media plays a big part and so it does the training; so evaluating not only the equipment and technical solutions but even the readiness and specific skills of the personnel and resource involved.

Modeling humans represent a challenge; in pandemic this issue introduce the necessity to include at least influences such as: perception of the threats.

Among the others these agents need to demonstrate both rational and emotional capabilities; the intelligence of these agents is limited to operative decisions, and it should be simple enough to face the challenge of the problem size related to a national wide simulation. The authors decided to define features and aspects that should be model and tested in order to verify the capabilities of this intelligent agents respect the need of the epidemic simulation.

In fact among the others the following measurable characteristics are required to PANDORA Human Behavior Agent combining intelligence and emotions:

- Capability to move and interact with other ones based on the current situation and individual profile
- Capability to be configured based on specific parameters: i.e. physical, social, ethnic, religious regional aspects
- Capability to change individual profile based on the situation evolution
- Capability to demonstrate awareness of the current situation
- Capability to demonstrate effect of phenomena as:
    - Social Network Attractors (i.e. family, friends)
    - Stress, Fear, Fatigue, Agressiveness Evolution
    - Effect of the group/social contents on the behavior
    - Effect of media and communications
    - Effect of economic situation
    - Effect of Physical situation
    - Effect of Emotional Situation
    - History affecting future behavior
    - Hysteresis of human behavior modifiers
    - Psychological State Evolution:
        - denial, anger, bargaining, depression, Acceptance
        - underestimation, fear, terror, habit
    - Working Elements
        - Attitudes of employers towards absenteeism

It is important to note that in term of human behavior the media plays a big part and so it does the training; so evaluating not only the equipment and technical solutions but even the readiness and specific skills of the personnel and resource involved.

7. CONCLUSIONS

This papers describes a framework for proceeding with fast prototyping of new models for facing pandemics. In fact currently the existing H1N1 pandemic provides a unique opportunity, due to the low level of lethality and the pretty good data collected to improve existing epidemic models and to bring together international scientists to tackle a pressing international problem. The approach outlined here provides a national scale integrated approach to evaluating mitigation strategies. Finally the proposed

models aims to deliver proof-of-concept modeling and decision support tools that can be taken to production and implemented within the public health sector.

ACKNOWLEGMENT
The authors are very glad to thank .Pascal Perez at the Australian National University, Tim Germann at the Los Alamos National Lab and Don Gunasekera at the Australian Bureau for Agricultural and Regional Economics for their support in development of these concepts .

REFERENCES

Avalle L., Bruzzone A.G., Copello F., Guerci A. (1996) "Determination and Quantification of Functional Parameters Relative to Contamination Vector Logic in an Epidemic Simulation", Proceedings of ESS'96, Genoa, October 24-26, Italy

Avalle L, Bruzzone A.G. , F. Copello, A. Guerci, P.Bartoletti (1999) "Epidemic Diffusion Simulation Relative to Movements of a Population that Acts on the Territory: Bio-Dynamic Comments and Evaluations", Proceedings of WMC99, San Francisco, January

Bossomaier, T., and Snyder, A. (2004) Absolute Pitch Accessible to Everybody by Turning Off Part of the Brain. Organised Sound, vol. 9. p181–190.

Bossomaier, T., Standish, R., and Harré, M. (2009) Simulation of Trust in Client –Wealth Management Advisor Relationships. International Journal of Modelling and Simulation. CRiCS Centre for Research in Complex Systems, Charles Sturt University. School of Mathematics and Statistics, University of New South Wales.

Bradbury, R H. (2006) Towards a new ontology of complexity science. In: Pascal Perez and David Batten (eds) Complex science for a complex world: Exploring human ecosystems with agents, pp 21-26. Canberra: Australian National University E Press.

Bruzzone, A.G., Mosca R., Revetria R., Bocca E.B.E. (2005) Agent Directed HLA Simulation for Complex Supply Chain Modeling. Simulation, vol. 81(9). 647–655.

Bruzzone, A.G., Bocca E.R.A. (2006) Algorithms devoted to Reproduce Human Modifiers in Polyfunctional Agents. Proceedings of SCSC2006, Calgary, Canada, July 30-August 3

Bruzzone, A.G. (2007) Challenges and Opportunities for Human Behaviour Modelling in Applied Simulation. Keynote Speech at Applied Simulation and Modelling 2007, Palma de Mallorca.

Cornforth, D., Green, D. G., and Newth, D. (2005) Ordered asynchronous processes in multi-agent systems. Physica D– Nonlinear Phenomena, vol. 204 (1–2). 70–82.

Halloran, M. E., Ferguson, N. M., Eubank, S., Longini, I. M. Jr., Cummings, D. A. T., Lewis, B., Xu, S., Fraser, C., Vullikanti, A., Germann, T. C., Wagener, D., Beckman, R., Kadau, K., Barrett, C., Macken, C. A., Burke, D. S., and Cooley, P. (2008) Modeling targeted layered containment of an influenza pandemic in the United States. Proceedings of the National Academy of Science of the United States of America. vol. 105(12). 4639–4644. doi:10.1073/pnas.0706849105.

Germann, T.C. Kadau, K. Longini, I. R. Jr., and Macken, C. A. (2006) Mitigation strategies for pandemic influenza in the United States. Proceedings of the National Academy of Science of the United States of America. Vol. 103(15):5935–5940. doi: 10.1073/pnas.0601266103.

Buetre, B., Kim, Y., Thomson, J., and Gunasekera, D. (2006) Avian influenza: Potential economic impact of a pandemic on Australia. Australian Commodities, Vol. 13(2). 351–59.

Grigg E. Rosen, F., Koop C.E. (2006) "The Biological Disaster Challenge: Why we are least prepared for the most devastating threat and what we need to do about it", Journal of Emergency Management 4 (1), 23-35

Gunasekera, D. Ford, M., Heyhoe, E., Gurney, A., Phipps, S.J., Harman, I., Finnigan, J.J, and Brede, M. (2008) Global Integrated Assessment Model: a new analytical tool for assessing climate change risks and policies. Australian Commodities, Vol. 15(1). 197–217

Harman, I., Ford, M. Jakeman, G., Phipps, S.J., Brede, M., Finnigan, J.J, Gunasekera, D. and Ahammad, H. 2008 Assessment of future global scenarios for the Garnaut Climate Change Review: an application of the GIAM framework, CSIRO and ABARE, Canberra.

Hoegh-Guldberg O, Mumby PJ, Hooten AJ, Steneck RS, Greenfield P, Gomez E, Harvell CD, Sale PF, Edwards AJ, Caldeira K, Knowlton N, Eakin CM, Iglesias-Prieto R, Muthiga N, Bradbury RH, Dubi A, Hatziolos ME (2007) Coral reefs under rapid climate change and ocean acidification. Science 318:1737-1742

Lotze, Heike K., Hunter S. Lenihan, Bruce J. Bourque, Roger H. Bradbury, Richard G. Cooke, Matthew C. Kay, Susan M. Kidwell, Michael X. Kirby, Charles H. Peterson, and Jeremy B. C. Jackson (2006) Depletion, degradation, and recovery potential of estuaries and coastal seas. Science 312, 1806-1809.

Newth, D. (2009) Asynchronous Iterated Prisoner's Dilemma. Adaptive Behavior. vol. 17(2). 157–183. doi:10.1177/1059712309104313.

Newth, D. and Cornforth, D. (2009) Asynchronous Spatial Evolutionary Games BioSystems. vol. 95(2). 120–129. doi:10.1016/j.biosystems.2008.09.003.

Perez P. and Batten D. (eds) (2006) Complex Science for a Complex World. Exploring Human Ecosystems with

Agents. ANU E Press, Canberra Australia, ISBN 1920942386.

Perez P. (2008) Embracing social uncertainties with complex systems science. In: M. Smithson and G. Bammer (eds), Challenges of Uncertainty. EarthScan Publisher, London, UK , pp 147-155.

Ratna N. N., Dray A., Perez P., Grafton R. Q., Newth D. and Kompas T. (2008) Diffusion and Social Networks: Revisiting Medical Innovation with Agents. In: Qudrat-Ullah, Spector and Davidsen (eds.). Complex Decision Making: Theory and Practice. Springer-Verlag, 247–268.

Rosen J.M., Grigg E.B., McKinght F., Koop C.E., Lillibridge S., Kindberg B.L., Hettinger L., Hutchinson R. (2004) "Transforming Medicine for Biodefense and Healthcare Delivery", IEEE Engineering in Medicine and Biology Magazine, January/February 89-101

Snyder, A. W., Bossomaier, T., and Mitchell, D., J. (2004) Concept formation: object attributes dynamically inhibited from conscious awareness. Journal of Integrative Neuroscience. vol. 3 (1). 31–46.

AUTHORS BIOGRAPHY

Agostino G. Bruzzone since 1991, he has taught "Theories and Techniques of Automatic Control" and in 1992 he has become a member of the industrial simulation work group at the ITIM University of Genoa; currently he is Full Professor in DIPTEM. He has utilized extensively simulation techniques in harbor terminals, maritime trading and sailboat racing sectors. He has been actively involved in the scientific community from several years and served as Director of the McLeod Institute of Simulation Science (MISS), Associate Vice-President and Member of the Board of the SCS (Society for Modelling & Simulation international), President of the Liophant Simulation, VicePresident of MIMOS (Movimento Italiano di Simulazione) and Italian Point of Contact for the ISAG (International Simulation Advisory Grouyp) and Sim-Serv. e has written more than 150 scientific papers in addition to technical and professional reports in partnerships with major companies (i.e. IBM, Fiat Group, Contship, Solvay) and agencies (i.e. Italian Navy, NASA, National Center for Simulation, US Army). He teaches “Project Management” and "Industrial Logistics” at the University for students in the Mechanical Engineering (4th year), Management Engineering (4th year) and Logistics & Production Engineering (3rd year) degree courses.

Marina Massei is currently enrolled in DIPTEM University of Genoa as member of the Simulation Team of Prof.Agostino Bruzzone. She is involved in the organization of International Scientific Events (i.e. Summer Computer Simulation Conference 2003-Montreal, 2004-San Jose', 2005-Philadelphia) She served as Chair for Applications in Management, Planning & Forecasting Workshop in SummerSim 2004, 2005, 2006. She is Associate Director of the MISS Center at the University of Perugia She was Finance Control & Administration Director of MAST.

Terry Bossomaier has worked in complex systems for over a decade. He co-edited the first advanced text for Cambridge University Press. He has 1 Science paper, had 3 ARC Discovery and 5 ARC Linkage grants. Papers: 114; Citations: 576; h-index: 10; g-index: 21. Time allocation 20% 2009 and 25% 2010. He has completed the Australian Institute of Company Directors training in Governance and Financial Management.

Davis Newth was the first PhD graduate in complex systems from Charles Sturt University and went on to win the Vice Chancellors award for research excellence. He was awarded the John Philip Science Medal for CSIRO young scientist of the year in 2008. Papers: 32; Citations: 139; h-index: 7; g-index: 10.

Joseph M. Rosen, M.D.: Dr. Rosen is Professor of Surgery (Plastic and Reconstructive) at Dartmouth-Hitchcock Medical Center in Lebanon, New Hampshire. He also teaches two courses at the Thayer School of Engineering, Dartmouth College - Healthcare and Biotechnology in the 21st Century and Virtual Medicine and Cybercare. Dr. Rosen is a member of the Melanoma Tumor Board and Affiliate Member of the Norris Cotton Cancer Center at Hartmouth-Hitchcock Hospital. He received his B.A. from Cornell University and an MD from Stanford University School of Medicine. Major research interests are: Nerve Repair and Human Machine Interfaces, Microsurgery and Transplantation of Limbs, Computer-Aided Surgery, Virtual Reality Simulators and methods of Education, Telemedicine and informatics, Healthcare Delivery for Medical Disasters and Counter Measures for Bio-Terrorism. He is a former Director, Section, Emerging Technologies Assessment at the Institute of Security Technology Studies at Dartmouth College, Hanover, New Hampshire. He is a committee member on the Bio-terrorism Working Group at the IEEE - USA Medical Technology Policy Committee and a consultant at the Institute of Defense Analysis in Washington, DC. Dr. Rosen is a member of the Plastic Surgery Research Council, American Society of Plastic Surgeons and the Plastic Surgery Educational Foundation, and is a member of the Strategic Committee of the Association of Plastic Surgery. Since 1981, Dr. Rosen has done volunteer surgery in developing countries for Interplast, Inc.and other groups.

SIMULATION AND OPTIMIZATION OF EVOLUTION IN BIOSYSTEMS BY INFORMATION-THEORETIC ENTROPY

Stanislaw Sieniutycz

Faculty of Chemical Engineering, Warsaw University of Technology
00-645 Warsaw, 1 Warynskiego Street, Poland
sieniutycz@ichip.pw.edu.pl

ABSTRACT
Evolution of animals with multiple organs or limbs is considered. We report computer-aided modeling and simulation of evolution in biological systems with living organisms as the effect of extremum properties of classical statistical entropy of Gibbs-Boltzmann type or its associates, e.g. Tsallis *q*-entropy. A variational problem searches for the maximum entropy subject to the geometric constraint of constant thermodynamic distance in a non-Euclidean space of independent probabilities p_i, plus possibly other constraints. Tensor form of dynamics is obtained. Some processes may terminate rapidly due to instabilities. A gradient dynamics can be predicted from variational principles for shortest paths and suitable transversality conditions.

Keywords: evolution, entropy, gradient dynamics.

1. INTRODUCTION

Systems capable of increasing their size and/or number of states (growing systems) can exhibit critical behavior when their size increases beyond a certain value. In effect, some developmental processes, i.e. biological evolution, may progress in a relatively undisturbed way, whereas others may terminate rapidly due to strong instabilities. To describe such phenomena quantitatively, we present here an extremum principle for entropy S of a physical or a biological system with variable number of states, thus making it possible to investigate processes of biological development and evolution. The extremum principle is of the variational nature and may be formulated as the problem of maximum S subject to the geometric constraint of the constant thermodynamic distance in a (generally non-Euclidean) space of independent probabilities p_i, plus possibly other constraints. The dynamics found are presented in the tensor form. An essential result implies that various dynamics, in particular those of growth processes (characterized by the increase in number of states), are governed by the gradient of the entropy in a Riemannian space.

2. ROLE OF COMPLEXITY AND ENTROPY

In the thermodynamic theory of evolution extrema of complexity Γ with respect to entropy S provide important information. There is a multitude of complexity measures in the literature, all capturing some aspects of what we mean when we say a process is complex. According to Saunders and Ho (1976, 1981) the complexity growth is the most probable effect in evolving systems. Complexity Γ is a function of disorder D and order Ω, which, in turn, are functions of the information-theoretic entropy S and number of possible states n. When speaking about the complexity and related entropy one issue is particularly important: a nonequililibrium entropy has to be necessarily applied because the difference between the maximum entropy and the actual system's entropy governs the organization in a complex system. Schrödinger (1967) has defined the disorder D and order Ω as expotential functions of non-dimensional quantities S and $-S$ respectively, in units of k_B. Yet, as pointed out by Landsberg (1984) these notions are inappropriate for growing systems. For such systems, Landsberg's definitions of disorder D and order Ω apply

$$D \equiv S / S_{\max} \tag{1}$$

$$\Omega \equiv 1 - D = 1 - S / S_{\max} . \tag{2}$$

As S_{max} depends on the number of states, n, in Landsberg's definition both disorder D and order Ω are functions of the information entropy S and number of states, n, i.e. $D = D\,(S, n)$ and $\Omega = \Omega\,(S, n)$. One especially simple form is the complexity

$$\Gamma_n(S) \equiv 4D\Omega = 4(S / S_{\max(n)})(1 - S / S_{\max(n)}) \tag{3}$$

The coefficient 4 in Eq. (3) is introduced to normalize the quantity Γ_n. The subscript n refers to a complexity sequence in a system with growing number of states, n. The solution to the following equation

$$d\Gamma_n(S) / dS = 0 \tag{4}$$

allows one to determine the extremum value of the information-theoretic entropy, $\hat{S}$, which maximizes complexity Γ_n (Szwast, Sieniutycz and Shiner 2002). The maximum attained by the function $\Gamma_n = \Gamma(S, n)$ equals the unity. This maximum appears for D = 0.5.

Hence the complexity-maximizing S equals to one half of the maximum entropy (the total randomness)

$$\hat{S} = \frac{1}{2} S_{\max}(n) \tag{5}$$

Evolution occurs on submanifolds that are surfaces of constant entropy (Szwast, Sieniutycz and Shiner 2002). Here the generalized q-entropy or Tsallis entropy S is used in probabilities $\mathbf{p}= p_0, p_1, ...p_n$

$$S = (q-1)^{-1}\left(1 - \sum_{k=1}^{n} p_k^q\right), \tag{6}$$

where p_k is the probability of finding an element in the state i among n states possible and $\Sigma p_k=1$. Equation (6) refers to Tsallis statistics (Kaniadakis, Lissia, Rapisarda, 2002) which generalizes the Boltzmann-Gibbs statistics by introducing an additional parameter q, that is called non-extensivity parameter. The additive Boltzmann entropy is regained from (6) in the limit of q=1. Performing maximization of S one can easily show that for a sole constraint $\Sigma p_k=1$ maximum of entropy occurs for the total randomness. All probabilities satisfy then the equality $p_i=p_k=n^{-1}$ and the maximum entropy is

$$S_{\max} = (q-1)^{-1}\left(1 - n^{1-q}\right). \tag{7}$$

In the classical case ($q \to 1$) this formula yields a well-known result, $S_{\max}= \ln(n)$.

Yet, in an example considered, a multi-organ animal is a system with $2n+1$ probabilities that describe n pair of legs plus the remaining part of the body, hence

$$S_{\max} = (q-1)^{-1}\left(1 - (2n+1)^{1-q}\right). \tag{8}$$

In the classical case ($q \to 1$) this formula yields, $S_{\max}= \ln(2n+1)$. The complexity-maximizing entropy $\hat{S}_n$ equals to one half of these quantities. In the classical case of q=1 one obtains

$$\hat{S}_n = (1/2)\ln(1+2n)\,. \tag{9}$$

Generally, the complexity-maximizing entropy is

$$\hat{S}_n = \frac{1}{2} S_{\max} = \frac{1}{2}(q-1)^{-1}\left[1 - (2n+1)^{1-q}\right], \tag{10}$$

where $\hat{S}_n$ of Eq. (9) follows in the classical case of q=1. To work with independent p_i we eliminate the last probability from entropy (6) and normalization condition $\Sigma p_i=1$. We then obtain a *tilde entropy function*

$$S = \tilde{S}\ (p_0,\ p_1,...p_{n-1}) \tag{11}$$

and, from Eq. (3), a related complexity

$$\tilde{\Gamma}_n\ = F_n\ [\tilde{S}(p_0,\ p_1,...p_{n-1})]\,. \tag{12}$$

Using independent probabilities we work with the tilde entropy. For the evolutions satisfying the maximum complexity (Saunders and Ho 1976), the entropy $\tilde{S}(\mathbf{p})$ equals to the complexity-maximizing entropy $\hat{S}$. This is consistent with the statement of Saunders and Ho (1976) "The only completely reversible changes are those which are isocomplex". For $\hat{S} = \tilde{S}\ (p_0, p_1,..p_{n-1})$, a subset of probabilities $\mathbf{p}$ = $(p_0,p_1,..p_{n-1})$ is found describing the evolution submanifolds by an equality

$$\tilde{S}(p_0, p_1,...p_{n-1}) = \hat{S}_n \tag{13}$$

and assuring the value of $\hat{S}_n = (1/2)S_{\max}(n)$. Within this manifold, a reversible modification of states is possible. In other words, in the evolution examples, the solutions to equality $\tilde{S}\ (p_0,\ p_1,..p_{n-1}) = (1/2)S_{\max}(n)$ refer to the *submanifolds of evolution*, or surfaces on which modifications (mutations) of organs may occur.

3. EVOLUTIONS OF MULTIPLE ORGANS WITHOUT MUTATIONS

Following earlier works (Saunders and Ho 1976, 1981; Szwast, Sieniutycz and Shiner 2002; Szwast 1997) we analyse here the *evolution of a multiple-organ or a multi-limb organism*, e.g. trilobite, an animal with many pair of legs. Although trilobites died out millions years ago, their anatomical structure is known due to the excavations. For our purposes it is sufficient to distinguish one pair of legs, of probability $2p_i \equiv 2p_1$ from the remaining parts of the organism. The remaining part has probability p_0. For a multi-organ animal with n pairs of legs, the following holds

$$S = (q-1)^{-1}\left(1 - p_0^q - 2np_1^q\right) \tag{14}$$

and, in the classical case,

$$S = -\,p_0 \ln p_0 - 2np_1 \ln p_1\,. \tag{15}$$

They both hold subject to the condition of $\Sigma p_i=1$, or

$$p_1 = (1-p_0)/2n \tag{16}$$

Whence the entropy in terms of single independent p_0

$$\tilde{S}(p_0) = (q-1)^{-1}\left(1 - p_0^q - (2n)^{1-q}(1-p_0)^q\right) \tag{17}$$

and its classical limit of q=1

$$\tilde{S}(p_0) = -\,p_0 \ln p_0 - (1-p_0)\ln[(1-p_0)/2n]\,. \tag{18}$$

Each of these entropies must be equal to its complexity-maximizing counterpart as described by Eq. (10) and its classical limit (9). By comparing Eq. (10) and (17), for an arbitrary q, we obtain equations for independent probabilities in terms of n and sole p_0

$$(q-1)^{-1}\left(1-p_0^q-(2n)^{1-q}(1-p_0)^q\right) = \frac{1}{2}(q-1)^{-1}\left[1-(2n+1)^{1-q}\right] \tag{19}$$

In the case of classical entropy

$$-p_0 \ln p_0 - (1-p_0)\ln[(1-p_0)/2n] = (1/2)\ln(1+2n) \tag{20}$$

From these equations probabilities p_0 and (then) p_1 can be calculated in terms of n. Graphs of these results for q=1 are presented in the literature (Szwast, Sieniutycz and Shiner 2002; Szwast 1997), where appropriate results are restricted to points describing the evolution without modifications or specializations.

4. ORGANISMS WITH MUTATIONS OR SPECIALIZATIONS OF ORGANS OR LIMBS

Here we analyze the evolution of a multiple-organ animal with mutations or specializations. An example relevant to this case follows the scheme including the stage trilobite $\Rightarrow$ crab. With entropies S expressed in terms of independent probabilities p_i (functions $\tilde{S}$ (p_0,....p_{n-1}) one may consider effects of reversible modifications (mutations) of multiple organs (e.g. pair of legs), for a fixed value of $\hat{S}_n$. In the considered example, after modification of a pair of legs to claws a crab emerges from a trilobite. Considering the anatomical structure of the crab, one pair of claws is distinguished with probability $2p_2$. When the modification occurs without change in number of pairs of legs and claws, the following equality holds

$$2p_2 + 2(n-1)p_1 + p_0 = 1. \tag{21}$$

For an organism with one pair of organs modified (specialized) on the reversibility surface, Eq. (13) and (21) are applied in the space of independent probabilities to describe the equality of the generalized entropy $\tilde{S}$ and the complexity-maximizing $\hat{S}_n$. After comparing the two entropy expressions we obtain

$$(q-1)^{-1}\left(1-p_0^q-2(n-1)p_1^q-2p_2^q(p_0,p_1)\right) = \frac{1}{2}(q-1)^{-1}\left[1-(2n+1)^{1-q}\right] \tag{22}$$

in the generalized case, and

$$-p_0 \ln p_0 - 2(n-1)p_1 \ln p_1 - 2p_2(p_0,p_1)\ln p_2(p_0,p_1) = \frac{1}{2}\ln(1+2n) \tag{23}$$

in the classical case. The probability function $p_2(p_1, p_0)$ used in these equations has the form

$$p_2 = \frac{1}{2}(1-p_0-2(n-1)p_1) \tag{24}$$

that follows from the condition Σp_i=1 represented by Eq. (21). Note that $2p_2$ is the probability attributed to the modified organ or limb. The complexity-maximizing entropies ($\hat{S}_n = S_{max/2}$) are those used earlier, Eq (10). The evolution submanifolds are now the family of lines $p_0(p_1, n)$. They describe organisms possessing n-1 of identical organs (pairs of legs) and one organ being modified, specialized, or subjected mutations. A special subset of data refers to organisms without specialization (Sec.3).

In the evolution literature Williston's law is frequently quoted (Saunders and Ho 1976, 1981; Szwast, Sieniutycz and Shiner 2002), which subsumes the results of observation and comparative analysis. This law states that if an organism possesses many of the same or similar elements, a tendency appears to reduce the number of these elements along with the simultaneous modification (specialization) of these elements which are saved by the organism. In the example considered here, the evolution submanifolds describe organisms possessing n-1 of identical organs (pairs of legs) and one organ being modified, specialized, or subjected mutation. Spontaneous increase of complexity is here a basic feature of evolution. For biological systems a reasonable measure of complexity is the different number of components they contain, so as to be consistent with well-known Williston's law which predicts that very similar components will either merge or specialize. Staying on the ground of the idea of increasing complexity, the system's organization acts as a force which prevents loss of components and allows complexity increase by the integration of new ones. This leads to a principle stating that an organism with more of organs (e.g. pairs of legs) is more susceptible to evolution towards an increase in the number of these organs.

Yet, during reversible specialization of organs, the state of an organism can fall into the region of the catastrophic decrease of number of these organs. These catastrophes constitute the price of specialization. The likelihood of falling in the catastrophe region increases with the number of organs. This explains why organisms possessing large number of similar organs ultimately reduce this number, despite the fact that they are more susceptible to evolutionary increase in the organ number. This also agrees with the well-known formulation of Williston's law of evolution (Saunders and Ho 1976) that subsumes the results of observation and is confirmed by the excavation experiments. In the

dynamical description of this problem an extremum principle provides a quantitative picture of biological development. It shows that a discrete gradient dynamics (governed by the entropy potential) can be predicted from variational principles for shortest paths and suitable transversality conditions.

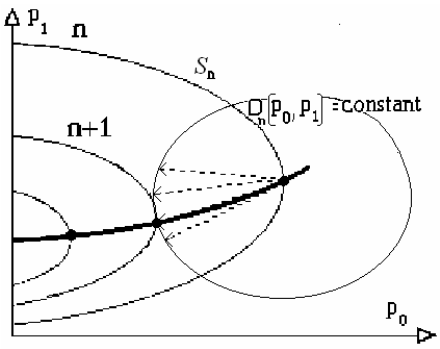


Figure 1. An original optimization problem for a system with *n* states is that of maximum change of entropy or complexity Γ between a point and the surface of constant distance from the point.

5. A VARIATIONAL APPROACH TO THE EVOLUTION DYNAMICS

Working in the dynamical context, (Szwast, Sieniutycz and Shiner 2002), we may analyze the evolution of living organisms as multi-organ or multi-limb systems by using the complexity criterion based on a potential, usually taken as the classical statistical entropy (of Shannon-Boltzmann) and the entropy-based complexity.

Here, however, in order to penetrate a vaster spectrum of stability (instability) properties, the generalized Tsallis entropy S is used (in k_B units) as a function of independent probabilities p_0, p_1, ...p_n. In the analysis of this sort, classical thermodynamic quantities do not appear, yet the model used satisfies an extremum principle that, similarly as in thermodynamics, implies the maximum of entropy subject to the geometric constraint of a given thermodynamic distance.

More specifically, an original optimization problem for a system with *n* states is that of maximum change of entropy or entropy-related complexity Γ between a point and the surface of constant distance from the point (Fig.1).

Dual forms of this principle can also be considered, where one mimimizes the thermodynamic length subject to a fixed change of the system's complexity or entropy (Fig.2). In the dual problem one searches for minimum length between a point (for a system with *n* states) and the entropy manifold of the system with *n*+1 states. In this formulation specific properties of the shortest lines variational problems can directly be applied, (Lyusternik 1983), if one doesn't want to use standard theory of variational calculus (Elsgolc 1960).

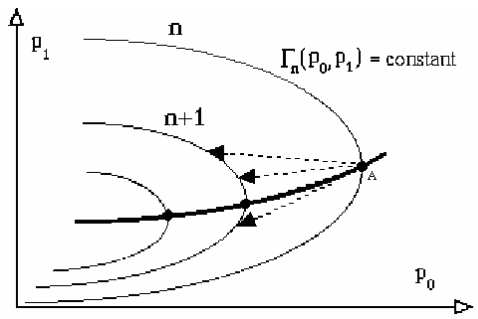


Figure 2. Dual optimization problem for a system with *n* states is that of minimum length between a point and the surface of constant entropy for the system with *n*+1states.

In the variational formulation, an s-dimensional evolution can be obtained via minimization of the length functional

$$J = \int_{p_0^n}^{p_0^{n+1}} a(p_0, p_1)\sqrt{1+(dp_1/dp_0)^2}\, d\, p_0$$

$$\equiv \int_{p_0^n}^{p_0^{n+1}} L(p_0, p_1, dp_1/dp_0)dp_0 \qquad (25)$$

where independent probabilities p_i are constrained to reside on the constant-entropy manifold satisfying the constraint $d\tilde{S}=0$. (As in our example above we restrict to the system of two independent p_i.) The "conformal coefficient" $a(p_0, p_1)$ takes into account the deviation from the Euclidean measure of length in the simplest way possible. Regarding the problem of geodesic lines, see, e.g., (Lyusternik 1983). To handle the constraint $d\tilde{S}=0$ one introduces the function $F_n(p_0, p_1)= \tilde{S} - \hat{S}_n$ whose numerical value equals zero for all states corresponding to the complexity-maximizing entropy. These are states residing on the evolution submanifolds or "reversible isolines". The direction coefficient of *tangent* to the submanifold $\tilde{S}_n$ $(p_0, p_1..)$ is the derivative $v_n \equiv (dp_1/dp_0)_n$. This derivative can be determined in terms of the partial derivatives functions $\tilde{S}_n$ or F_n as

$$v_n = (\frac{d\, p_1}{d\, p_0})_n = -\left(\frac{\partial \tilde{S}_n(p_0, p_1)}{\partial\, p_0}\right)\left(\frac{\partial \tilde{S}_n(p_0, p_1}{\partial\, p_1}\right)^{-1}. \qquad (26)$$

To predict the location of a point on the submanifold *n*+1 when a point on the submanifold *n* is given, we consider *the variational problem* of the shortest line between the two points located on the submanifolds $\tilde{S}_n$ and $\tilde{S}_{n+1}$.This corresponds with a minimum of the length functional (25).

The necessary extremum condition (the Euler equation) shows that any deviation from the Euclidean metric (measured in terms of *a*) influences, in general,

shapes of shortest-distance extremals which lead from the submanifold n to the submanifold n+1. Of course, upon restriction to the Euclidean metric and the Euclidean distance (a =1), extremals are the family of the straight lines $p_1=C_1p_0+C_2$. Generalization to non-flat spaces is given below, but we shall first consider *the transversality condition* for an arbitrary integrand L of (25) that may be associated with curvilinear spaces. An extremal which starts from a point on the submanifold n and terminates on the submanifold n+1 satisfies

$$\delta J = \{L - \frac{\partial L}{\partial(d\,p_1/d\,p_0)}(d\,p_1/d\,p_0)\}\delta\,p_0 + (\frac{\partial L}{\partial(d\,p_1/d\,p_0)})\delta\,p_1 = 0, \quad (27)$$

(c.f. Elsgolc 1960), where δp_0 and δp_1 are linked by condition (22) applied for n+1 rather than n.

For any length-type integral J, Eq. (27) defines the condition associated with the extremal which starts from a point on the submanifold n and terminates on the submanifold n+1. In view of arbitrary variations of p_0 the substitution of δp_1 from Eq. (26) into Eq. (27) yields the equality u_{n+1} = -1/v_{n+1}, where u_n is the slope coefficient of the normal to isoline $\tilde{S}_{n+1}$. This means the *orthogonality of the slopes* u_{n+1} *and* v_{n+1} in Euclidean spaces. In other words, in the case of an assumed or imposed Euclidean geometry (the transversality condition unaffected by the coefficient $a(p_0, p_1)$=1) the slope coefficient of the tangent to the extremal is

$$u_{n+1} = (\frac{d\,p_1}{d\,p_0})_{n+1} = \frac{\partial\,\tilde{S}_{n+1}(p_0, p_1)}{\partial\,p_1} \Big/ \frac{\partial\,\tilde{S}_{n+1}(p_0, p_1)}{\partial\,p_0}. \quad (28)$$

This condition implies the *gradient dynamics* in flat spaces, with probabilities changing with time in the form

$$(\frac{d\,p_i}{dt})_{n+1} = \omega\frac{\partial\,\tilde{S}_{n+1}(p_0, p_1)}{\partial\,p_i} \quad (29)$$

(i=0,1). The frequency-type coefficient ω has the interpretation of a kinetic constant. A related discrete dynamics of evolution contains the finite differences **p**(n+1)- **p**(n) instead of the time derivatives.

Next, it may be shown that the gradient dynamics also holds also in curvilinear spaces. In fact, admitting non-flat metrices (i.e. working with situation when a Lagrangian associated with a non-flat metric is effective) one may show (Gołąb 1956) that *the tensor generalization* of the continuous model (29) is

$$\frac{d\,p_i}{d\tau} = g_{ik}\frac{\partial\,\tilde{S}\;(p_0,\ldots p_s)}{\partial p_k}, \quad (30)$$

where $\tau = \omega t$ is a nondimesional time and ω is the frequency coefficient of Eq. (29) and g_{ik} is s-dimensional Riemannian tensor (Gołąb 1956). The consequence of this equation is the tensor form of the discrete evolution dynamics with the Onsager-like structure, where his symmetry matrix $L_{ik} = \Delta t\,\omega g_{ik}$ appears

$$p_s(n+1) - p_s(n) = L_{s1}\frac{\partial\,\tilde{S}_{n+1}(p_0,.\,p_s)}{\partial\,p_0} + \ldots + L_{ss}\frac{\partial\,\tilde{S}_{n+1}(p_0,.\,p_s)}{\partial\,p_s}, \quad (31)$$

as in the classical irreversible thermodynamics. Therefore, the evolution processes can be imbed quite naturally into a relatively large family of thermodynamic processes.

6. CONCLUDING REMARKS

By applying the tensor calculus, one can develop a discrete, nonlinear representation of evolution dynamics in metric spaces that may be curvilinear. Dynamic programming algorithms (Bellman's equations) can be derived and computer-aided simulations of their solutions can be performed. Systems governed by nonclassical q-entropies may exhibit irregular shape of entropy hill and show quantitatively distinct picture of instabilities than classical.

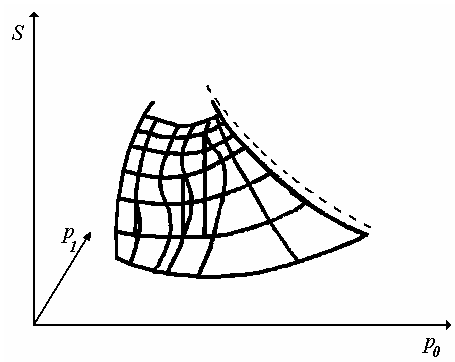


Figure 3. Szwast's (1997) computer simulation of instabilities in evolution systems. Trajectory of increase of number of pairs of legs (broken line), manifolds of reversible modification (specialization) of one pair of legs (lines of S=const) and vertical jumps describing a rapid decrease in number of pairs of legs (vertical lines).

Evolutions of living organisms can be described in terms of variational principles for maximum of generalized entropy along with suitable transversality conditions. General gradient dynamics (in curvilinear spaces), that governs the evolution problems, is of Onsager's structure and is consistent with the entropy principle of extremality as the driving factor in the discrete dynamics of complex and living systems, postulated recently (Szwast, Sieniutycz and Shiner 2002). We have shown that such a principle should be

integrated with the evolution theory of biological systems.

Calculations of gradients of entropy or disorder versus number of organs, *n*, performed by Szwast (1997) show that these gradients increase with the number of multiple organs. This allows one to formulate the principle which states that organisms with more of organs are more susceptible to evolution towards an increase in the number of these organs. Yet, during reversible specialization of organs, the state of an organism can fall into the region of the catastrophic decrease of number of these organs. These catastrophes constitute the price of specialization. The calculations show, Szwast (1997), that the likelihood of the system's falling in the catastrophe region increases with *n*. Some of the related results are presented in Fig.3. This discussion explains why organisms possessing large number of identical organs ultimately reduce this number, despite the fact that they are more susceptible to the evolutionary increase in the organ number, in agreement with Williston's law. This law states that if an organism possesses many of the same or similar elements, a tendency appears to reduce the number of these elements along with the simultaneous modification (specialization) of these elements which are saved by the organism (Saunders and Ho 1976; Szwast, Sieniutycz and Shiner 2002; Szwast 1997). Entropy-based models, quantifying these critical phenomena, are enriched in this paper by inclusion of non-classical statistical entropies, e.g. *q*-entropies of Tsallis or Renyi, that may modify magnitudes of unstable regions in the space of process probabilities.

ACKNOWLEDGMENTS

This research was supported in 2009 by the statute grant of the Faculty of Chemical and Process Engineering at the Warsaw University of Technology.

REFERENCES

Elsgolc, L.E., 1960. *Variational Calculus*. Warsaw: Państwowe Wydawnictwa Naukowe.

Gołąb, S., 1956. *Tensor Calculus.* Warsaw: Państwowe Wydawnictwa Naukowe.

Kaniadakis, G., Lissia, M., Rapisarda, A., eds, 2002. *Proceedings of the International School and Workschop on Non-extensive Thermodynamics and its Applications, NEXT* 2001, 23-30 May 2001, Villiasimius, Italy. *Physica A* 305, No. 1+2.

Landsberg,, P.T., 1984. Can entropy and order increase together. *Phys. Lett* 102 A: 171.

Lyusternik, L.A., 1983. *The Shortest Lines Variational Problems*. Moscow: Mir Publishers.

Saunders, P.T., Ho, M.W., 1976. On the increase in complexity in evolution. *J. Theor. Biol*. 63: 375-384.

Saunders, P.T., Ho, M.W., 1981. On the increase in complexity in evolution II. The relativity of complexity and principle of minimum increase. *J. Theor. Biol.* 90: 515-530.

Schrodinger, E., 1967. *What is life?*. Cambridge University Press. (First edn in 1944).

Szwast, Z., Sieniutycz, S., Shiner, J., 2002. Complexity principle of extremality in evolution of living organisms by information-theoretic entropy. *Chaos, Solitons and Fractals* 13: 1871-1888.

Szwast, Z., 1997. An approach to the evolution of selected living organisms by information-theoretic entropy (in Polish). *Reports of Faculty of Chemical and Process Engineering at Warsaw TU* 24: 123-143.

AUTHORS BIOGRAPHY

Prof. **Stanislaw Sieniutycz** (1940), PhD; ScD, has been since 1983 a Professor Chemical Engineerig at the Warsaw Technological University, Warsaw Poland. He has received his MsD in Chemistry in 1962, PhD in Chemical Engineering in 1968, and ScD (habilitation) in Chemical Engineering in 1973, all from Warsaw TU. He is recognized for his applications of methods of analytical mechanics and optimal control theory in thermodynamics and engineering. Within this framework his research embraces a range of topics: efficiency of energy conversion, optimization of separation systems and chemical reactors, energy savings, conservation laws, wave systems, heat and mass transfer and variational principles in continua.

His work with graduates has resulted in several PhD theses: Discrete Maximum Principle of Pontryagin's type (Z. Szwast 1979); Thermodynamic Liapounov Functions in Gas-Solid Systems (J. Komorowska-Kulik, 1979); Optimizations in Catalytic Cracking (A. Kebang 1980 and A. Dlugosz 1983); Modeling and Optimization in Separation of Coal Conversion Liquids (A. Dunalewicz 1989), Nonlinear Models of Mechanical Energy Production in Non-Ideal Generators Driven by Thermal or Solar Energy (M. Kubiak 2005 and P. Kuran 2006). Prof. **Sieniutycz** has contributed about 250 research papers in the field of chemical engineering and irreversible thermodynamics. He is also an author of several books: Optimization in Process Engineering (WNT Warsaw 1978 and 1991); Practice in Optimization Computations (with Z. Szwast; WNT, Warsaw 1982); Conservation Laws in Variational Thermo-Hydrodynamics (Kluwer Academic, Dordrecht 1995), and, recently, with Prof. J. Jeżowski, Energy Optimization in Process Systems, (Elsevier, Oxford and Radarweg, 2009) He has been a visiting professor in a number of schools: University of Budapest (Physics), University of Bern (Physiology), University of Trondheim (Chemical Physics), San Diego SU (Mathematics), University of Delaware (Chemical Engineering), and (several times at) The University of Chicago (Chemistry).

TOWARDS MULTI-LEVEL SIMULATION OF CONFLICT AND PEACEKEEPING

Mamadou Seck

Systems Engineering Group
Faculty of Technology, Policy and Management
Delft University of Technology
Jaffalaan, 5, 2628BX, Delft, THE NETHERLANDS

m.d.seck@tudelft.nl

ABSTRACT
This paper presents a multi-level simulation approach for analyzing a conflict opposing two groups (ethnic, political or religious groups) with a third actor, an armed policing force playing the peace keeping role.

The opposing groups are modeled at the macro-scale by a continuous logistic event-history model (Myers 2008) and at the lower scale by Discrete Event behavioral units representing the different roles at play within each group, interacting through a cell based terrain.

Keywords: conflict, peace keeping, DEVS

1. INTRODUCTION

Traditionally, simulation efforts designed for analysis, planning, or training in the military domain focus on operational and tactical issues. Such applications have been termed Computer Generated Forces (CGF) or Semi Automated Forces (SAF), combining Systems Engineering and Artificial Intelligence techniques to model aspects of military activity. These approaches generally model highly detailed human and technical systems in 2D or sometimes 3D Virtual Reality environments.

Strategically, it has been recognized that Military conflicts are drifting from pure force on force scenarios to situations in which societal dynamics are gaining more and more importance. This evolution is reported in numerous works *ex.* (Hammes 2006). The contemporary armed conflict is more likely intrastate, asymmetric, and urban, with civilian populations, activists, peacekeeping forces, and humanitarians interacting (Hobbs 2003).

As a consequence of this multiplication of roles and higher degree of complexity, the attrition models of the Lanchester-type (Schaffer 1967) or classical Computer Generated Forces (CGF) can no longer be used to gain full insight on the complex phenomena involved.

The social dynamics at play in present day armed conflicts call for more holistic simulation approaches more likely to be found in the social simulation research community, especially in agent based simulation studies. A number of authors have approached Low Intensity Conflicts, Peacekeeping, Civil Violence, or Counter Insurgency issues through multi-agent simulation.

This paper presents a multi-level simulation approach for analyzing a conflict opposing two groups (ethnic, political or religious groups) with a third actor, an armed policing force playing the peace keeping role. The opposing groups are modeled at the macro-scale by a continuous logistic event-history model (Myers 2008) and at the lower scale by Discrete Event behavioral units representing the different roles at play within each group, interacting through a cell based terrain.

2. GENERAL MODEL DESCRIPTION

This section presents the simulation model of a conflict involving two communities and an armed force playing a peace keeping role. The proposed model recognizes the necessity of multi-scale representations in complex social systems simulation, to this end, both macro-level (opinion dynamics) and lower level behaviors (individual activities of behavior units) are represented.

2.1. Macroscopic level

The macroscopic dynamics are captured by a logistic model adapted from Myers & Oliver's Opposing Forces Diffusion (OFD) model (Myers 2008).

Underlying the model is the idea that two competing *ideologies*, i.e. *Provocation* and *Repression* (P, R), shape the expression of collective action within a population. Collective action events are seen as the consequence of the two competing ideologies' diffusing within the population through imitation. The intensity of the provocation ideology can be interpreted as the extent to which contentious behavior is seen as an efficient strategy within the population, thus causing mobilization. Conversely, the repression force is interpreted as an ideology promoting demobilization. Event probability is obtained by P-R.

We consider two forms of collective actions in this model, namely, *Peaceful Demonstration* and *Violent Behavior,* each of these has an instance of the OFD model simulated in each community:

- The *Peaceful Demonstration* represents the tendency of the population to express its grievances through peaceful demonstrations, (although a demonstration could develop, given specific

circumstances and local interactions, into violent rioting behavior).

- The *Violent Behavior* represents activist groups engaging in provocation activities during peaceful demonstration or perpetrating terrorist acts.

OFD models for both behavior types are simulated to influence and be influenced by lower level agents.

2.2. Micro-level

In contrast to the higher level continuous behavior, lower level models interact discretely. They consist of agents that we term Behavioral Units (*BU* in what follows). These are aggregate autonomous agents belonging to a community or to the policing force, interacting dynamically through the terrain.

A *BU* models a collection of individuals pertaining to the same community and acting collectively at some instant in the conflict. Because major actions in armed conflicts are collective, we choose not to model individuals explicitly, but rather a collection of them inheriting their properties from their community of origin. In the current model, two behavior unit types exist in each community:

- A *Crowd* agent, representing a *mild* and numerous subset of the general population whose salient actions, depending on the environment, can be: '*inactive*', to '*demonstrate*', and to '*riot*'. This *BU* is activated by the *Peaceful Demonstration* OFD model.

- An *Activist* agent, representing a more *engaged* subset of the population whose salient actions can be: '*inactive*', '*provocation*', '*terror*'. This *agent* is activated by the *Violent Behavior* OFD model.

The outcomes of *BU* actions are compiled and synthesized in a model component also responsible for updating the OFD models imitation indexes, thus realizing the micro to macro link. The same model component activates the *BUs* with the event probabilities computed in the OFD models, thus realizing the macro to micro link. This component also communicates asynchronously with its counterpart in the adverse population model, thus representing the way one population's actions can be interpreted by the other population (useful for simulating concepts such as retaliation or intimidation).

The *agents* receive activations in a discrete-time fashion, i.e. with a predefined time step.

The Peace Keeping force is represented by several BU *agents*. In the current model, three *agent* types pertaining to the armed forces are modeled.

- A *Reconnaissance Patrol agent*, sequentially visiting locations in the terrain for information gathering purposes.
- A *Combat Patrol agent*, sequentially visiting locations in the terrain to harass and/or destroy *Activist agent.* This *agent* also collects information on current activities in the terrain.
- A *Crowd Control agent*, representing a subset of the force dedicated at quelling contentious collective actions on both sides, its possible actions can be: '*monitor*', '*block*', '*disperse*'. To become active, it must be called by a *Combat Patrol* or a *Reconnaissance Patrol agent.*

The terrain is a check board of *location* objects with the following attributes:

- symbolic value to population A (real [-1,1])
- symbolic value to population B (real [-1,1])
- proportion of population A (real [0,1])
- proportion of population B (real [0,1])
- accessibility (integer [0,1])
- units in presence (and their current activity)

Each location constitutes a cell in a board representing the territory.

Communication in the model is achieved in two different ways:

- Message passing through the DEVS transition and output functions,
- Terrain attributes read/write by *agent*s

3. MODELING FORMALISM

We adopt in this work the framework for modeling and simulation established by Zeigler and colleagues (Zeigler 2000). This framework has the benefit of separating concerns regarding modeling, simulation and experimentation. Conceptual models, whether continuous or discrete, are represented in a mathematical formalism, with predefined high-level modeling constructs (e.g. *events, states, transitions functions, output functions, coupling*). Regarding simulation, abstract simulators are proposed, completely specifying the operational semantics necessary to run model instructions. This approach facilitates verification and validation by separating conceptual and implementation issues, which is particularly useful in the case of discrete event simulation. As one would expect, these advantages come with a cost: for efficient communication, it is necessary to share a common understanding of the modeling formalism constructs. In this section, we briefly introduce the DEVS formalism, before providing the models' graphical specification.

DEVS is a modeling and simulation formalism for Discrete EVent Systems. First proposed by Zeigler in 1976, it consists of **sets** *(input values, output values, and states),* and **functions** applied to the latter sets *(internal transition, external transition, output, and time advance)* allowing complete and unambiguous specification of systems according to the discrete event abstraction and a simulation according to the event scheduling worldview. Atomic DEVS models are basic components for specifying behavior. Predefined atomic models can be composed hierarchically to represent complex networks called Coupled DEVS models.

An atomic DEVS model is a structure (Zeigler 1986):

$$M = < X,S,Y, \delta_{int}, \delta_{ext}, \lambda, ta >$$

Where:
X is the set of input ports and values
S is the a set of states
Y is the set of output ports and values
$\delta_{int}: S \rightarrow S$ is the internal transition function
$\delta_{ext}: Q \times X \rightarrow S$ is the external transition function,
where $Q = \{(s,e) | s \text{ in } S, 0 \leq e \leq ta(s)\}$ is the total state set
e is the time elapsed since the last transition (internal or external)
$\lambda : S \rightarrow Y$ is the output function
$ta: S \rightarrow R^{+}_{0,\infty}$ is the time advance function
where $R^{+}_{0,\infty}$ is the set of positive reals, including 0 and ∞

At any time, the system is in a given state s. State change can only occur through an event, either internal or external:

- An **internal event** takes place when ta(s)= e, meaning when the lifetime of the state is reached, in other words, when the accumulated time since the last transition has reached the value defined in the time advance function ta(s). As a consequence, the system **outputs** the value y in Y through the $\lambda(s)$ function and makes an **internal transition** to a new state, say s', defined as δ_{int}(s)=s'.
- An **external event** corresponds to the arrival of an **input** (x *in X*) on one of the models input ports before the lifetime of the state has expired, meaning that the elapsed time, e, is $0 \leq e \leq ta(s)$. As a consequence, the system changes to a new state, say *s''*, through an **external transition** that depends on the current state, the elapsed time, and the input, as defined by δ_{ext}(s,e,x)=s''.

Atomic components can be connected hierarchically to form coupled models. For brevity, the formal specification of coupled models will not be presented in this paper because it plays no significant part in basic behavior understanding.

Another advantage of this formalism is its readiness for interoperability with other classical modeling approaches, in particular the differential equations systems specification formalisms.

In this paper, the graphical representation of DEVS models is used. As shown in Figure 1, the atomic DEVS model is represented in a box with input and output ports. A phase defines an explicit subset of the state set. Phases are represented by nodes and transitions by arcs. Nodes are circles with a continuous line when the phase is passive and with a dotted line when the phase is active (A phase is active when an internal transition can fire to another phase and is passive otherwise.) In the case of an active phase, the lifetime function is represented. Labelled arcs represent transitions. External transitions are represented by continuous line arcs. Above the transition is noted the input port followed by a "?" symbol and the event value when the latter is defined. Internal transitions are represented by dotted line arcs. Above the transition is noted the output port followed by a "!" symbol and the event value. Under arcs, an expression defines the conditions of the transition.

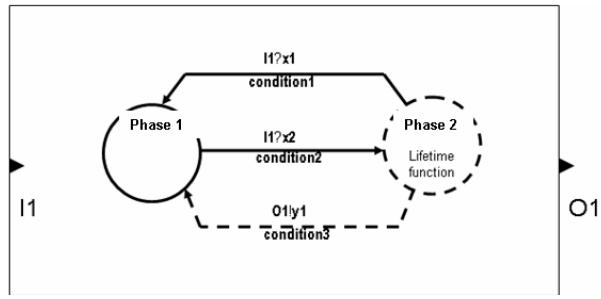


Figure 1: Graphical representation of a DEVS atomic model

4. MODEL SPECIFICATION

The general structure of the model is presented on Figure 2. Plain arrows represent event driven communication between sub-models; dashed arrows represent communication through common variables, such as the terrain. This section provides the specification of some model components.

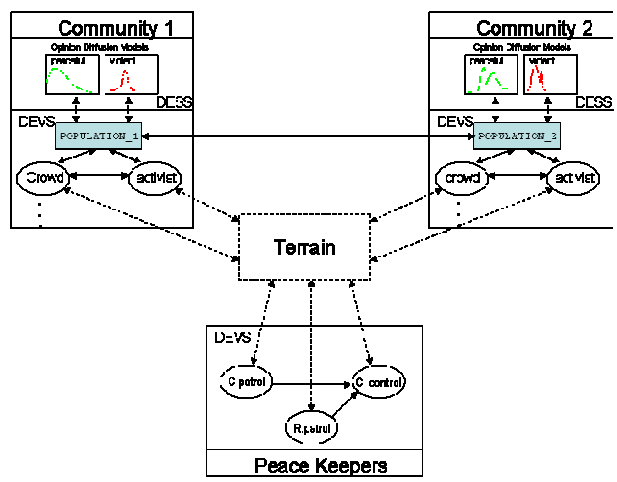


Figure 2: General model architecture

4.1. Opinion Diffusion Models

The Opposing Forces Diffusion model of collective behavior assumes the existence of two competing ideologies, *provocation* and *repression*, shaping the dynamics of any collective behaviors. Here,

> *"Ideology refer[s] to a system of beliefs about action and its consequences."* (Myers 2008).

In (Myers 1999), *provocation* (P) and *repression* (R) are formalized as two similar logistic functions representing the proportion of adopters of two competing ideologies. The intensity of collective

actions depends on the size of the difference between P and R, when P > R as in equations (1) and (2).

$$P^*(t) = \frac{1}{1 + \frac{1 - N^*_0}{N^*_0} \exp[-pt]}, \quad R^*(t) = \frac{1}{1 + \frac{1 - N^*_0}{N^*_0} \exp[-rt]}. \tag{1}$$

Where N^*_0 represents the initial proportion of adopters in the population, p and r represent, respectively, the provocation and repression ideologies' infectiousness.

Event probability at any instant is obtained by:

$$\frac{dV(t)}{dt} = P^*(t) - R^*(t). \tag{2}$$

The OFD model is a theoretical and explicative model. It fits well various datasets collected in the USA race riots of the 1960's (Myers 2008). Its theoretical grounding and straightforward parameter interpretation make it a good candidate for a simulation application.

The logistic model is specified as a DEVS model with an active phase making a transition back to itself indefinitely. To model the changes in the conflict state as a consequence of the interactions between the factions and the military force, we allow N^*, p, and r to change. Figure 3 shows an OFD model on which the repression ideology's infectiousness is increasing preiodically, and causing the event probability to decrease.

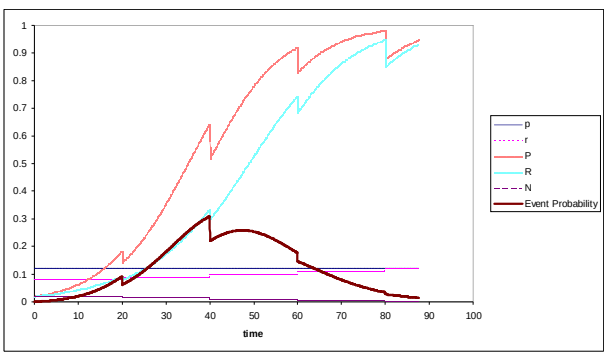


Figure 3 : Opposing Forces Diffusion Model with changing parameters

4.2. POPULATION sub model

The Population model has two roles, 1) transferring event probabilities from the OFD components to Behavioral Units at predefined time interval, and 2) synthesizing the outcomes of local interactions to adapt OFD models' diffusion indexes. Adapting diffusion indexes as a result of micro scale outcomes necessitates hypotheses regarding how for example, a terror act will have a future effect on the adoption of the violent *provocation* ideology in both the source population and the enemy population. Theories and possible hypotheses concerning influences between peaceful and violent ideologies within a community or between opposing communities can be diverse. From a modeler's perspective, it seems more useful to propose a generic scheme capable of representing any of such assumptions.

Behavioral Units generate outcomes after each action, depending on local interactions. These outcomes are the following:

- (o_1) Efficient crowd control : *a peaceful demonstration did not escalate to rioting*
- (o_2) Rioting : *a peaceful demonstration escalated to rioting*
- (o_3) Failed terrorist act : *the force was able to counter a planned terror act*
- (o_4) Successful terrorist act : *activists were able to perpetrate a terrorist act*

Each outcome in the model can be predefined to have a null, positive, or negative effect on peaceful and violent ideologies' diffusion indexes (p and r) in both communities. Table 1 shows a possible set of such modeling assumptions. Line 'o4' on the table defines the assumption that a successful terrorist act by an activist group affiliated to Community 1 has the following effects:

- strengthens *provocation ideology* for demonstrations in Community 1
- does not have any effect on the *repression ideology* for peaceful demonstration in Community 1
- strengthens *provocation ideology* for violent behavior in Community 1
- does not have any effect on the *repression ideology* for violent behavior in Community 1
- does not have any effect on the *provocation ideology* for peaceful demonstration in Community 2
- strengthens *repression ideology* for peaceful demonstrations in Community 2
- strengthens *provocation ideology* for violent behavior in Community 2
- does not have any effect on the *repression ideology* for violent behavior in Community 2

Table 1: Action outcome effects on repression and provocation ideologies

	SELF				***OTHER***			
Beh.	Peaceful Demonstration		Violent Behavior		Peaceful Demo.		Violent Behavior	
	Δp_P	Δr_P	Δp_V	Δr_V	Δp_P	Δr_P	Δp_V	Δr_V
O_1	0	+	-	0	+	0	+	0
O_2	-	+	+	-	+	-	+	-
O_3	+	0	0	+	+	0	0	+
O_4	+	0	+	0	0	+	+	0

The DEVS specification of the population sub-model is given in graphical form (Song 1994) on figure 4. External transitions are represented in continuous arrows, internal transitions in dashed arrows, states are represented by nodes. The symbols '?' and '!' respectively represent inputs and outputs.

The population sub model is represented on Figure 4. The other DEVS sub models will be directly presented in their graphical form.

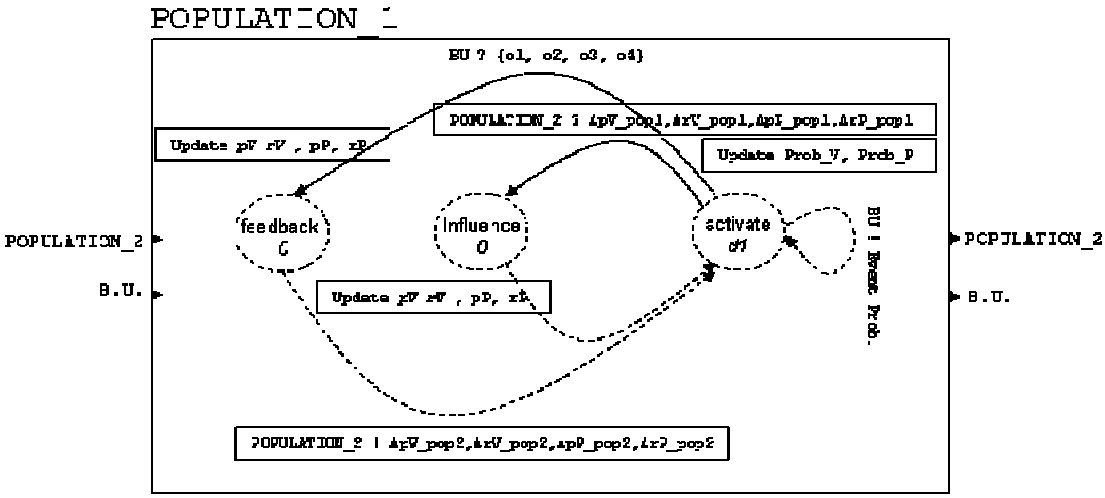


Figure 4: Graphical representation of the POPULATION model.

Every 'd1' time units, the POPULATION model updates the event probability calculation retrieved from ODF model and transfers the result as an activation message to BUs.

When BUs terminate their action execution, the action outcome is received by the population model which updates *provocation* and *repression* parameters of the OFD models according to the assumptions specified as on table 1.

4.3. CROWD model

The crowd model is one of the Behavioral Units considered. It represents a subset of the general population whose salient actions can be a demonstration or riot. This *BU* is activated by the *Peaceful Demonstration* OFD model.

When the activation message is received from the population atomic model, the crowd goes to the decision state named 'Decide' and either goes back to the inactive state or to the state 'demonstration'. The decision here is stochastic and is based on the probability generated by the OFD model as in Figure 3. When changing the state to 'demonstration', the activists in the same population are informed. 'demonstration' is an active phase that lasts for a predefined time period dn, after that, the terrain is read to obtain information about the other units in presence. Depending on elements like the presence of activist provocateurs, enemies or the peacekeeping force's perception and control policy, the demonstration might escalate to a riot or disperse. The outcome is sent as an event to the POPULATION atomic model and interpreted as in table 1 to alter provocation and repression indexes. Figure 5 depicts the graphical representation of the 'Demonstration' model.

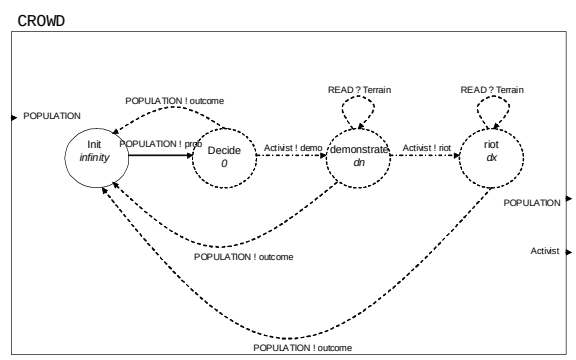


Figure 5: Graphical representation of the CROWD atomic model.

4.4. Activist model

Activist is a subset of the population whose salient actions can be: '*inactive*', '*provocation*', '*attack*. This *agent* is activated by the *Violent Behavior* OFD model. When the activation event is received by the activist model, it goes to a decision phase which either decides to stay inactive or to perpetrate an attack. The decision is based on the possibility to find a favorable location. The attack can succeed or fail, based on the presence of the force.

In case the activist model is inactive and a demonstration is organized by a CROWD of the same population, the activists can try to join the demonstration and turn it into a riot.

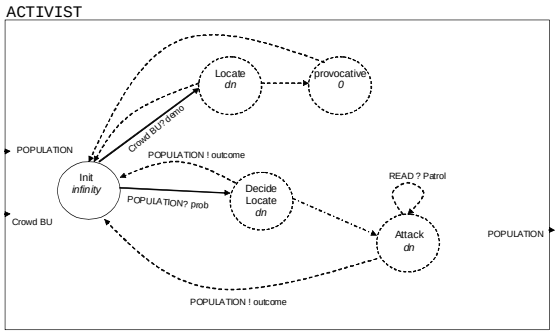


Figure 6: Graphical representation of the 'activist' atomic model.

4.5. Combat Patrol

The Combat Patrol agent sequentially visits locations in the terrain. If an Activist group is detected, the Combat patrol agent can intervene. This agent also collects information on current activities in the terrain and can send a message to a Crowd Control agent in case of a riot. This model component's behaviour is autonomous as it receives no incoming event.

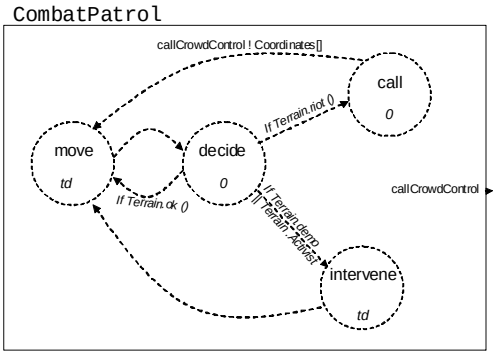


Figure 7: Graphical representation of the 'Combat Patrol' atomic model.

5. MODEL DYNAMICS AND DISCUSSION

This section looks at the model's dynamics over time. To understand the general dynamics that we wish to portray with the model, let us study the simplified scenario displayed on figure 7. The diffusion model for violent behavior in one population is initialized with the values (0.2, 0.12, and 0.08) for the parameters N^*_0, p, and r. As time elapses, P and R evolve. Because p is higher than r, provocation ideology quickly dominates, and the event probability starts growing, as a result, the likelihood of an attack by the activists gets stronger and stronger. At time 12, an activist agent manages to conduct a successful attack. Based on table 1, this results in an increase in the value of p, while r remains unchanged. The increase in p also causes a slight increase in N^*_0 which represents the current status of the conflict. The same happens at time step 23. At time 34, an attack by the activists is countered by a combat patrol. This leads to a decrease in the infectiousness of the provocation ideology and an increase in the repression ideology as specified on table 1. This event again transforms the stage of the conflict and makes the probability of activist actions lower.

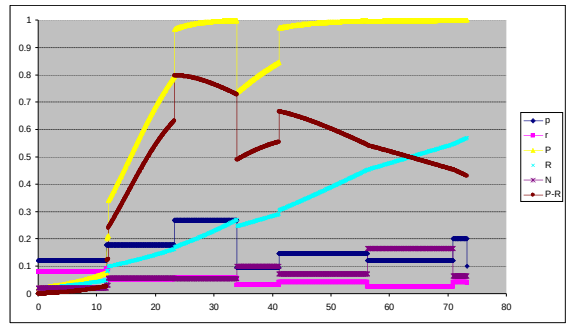


Figure 8: Model dynamics

The presented framework aims at capturing some of the complexity of conflicts through simulation.

The approach recognizes the necessity of multi-level representations with upward and downward influences. A design choice is also made to leave room for the implementation various assumptions of sociological relevance, making the model extensible.

To make full use of its potential, a validation and calibration with a well documented historical case is envisioned. This will allow making analyses on the effect of different peace keeping strategies with the simulation. The model would as a result be possible to use in analysis, planning and training, as well as serious gaming applications.

A number of additional improvements will be considered in future works, including a visualization of the conflict theatre, the inclusion of other sociological factors such as culture, economical status or religious determinants.

REFERENCES

Hammes, T.X., 2006. Four Generations of Warfare. In: *The Sling and the Stone, On War in the 21st Century*, St. Paul, MN. p. 293.

Hobbs, W.S.R, Dexter P., Egudo M., and Grieger D., 2005. *An Analytical Framework for the Study of Asymmetric Warfare.DSTO-TR-1670*. Edinburgh, Australia: Land Operations Division, Defence Science and Technology Organisation.

Myers, D.J., and Oliver, P.E., 2008. The opposing forces diffusion model: the initiation and repression of collective violence. *Dynamics of Asymmetric Conflict*,1 (2), 164-189

Schaffer, M.B., 1967. Lanchester Models of Guerrilla Engagements. *Operation Research* 16 (3), 457-488.

Zeigler, B., Praehofer, H., and Kim, T., 2000. *Theory of modeling and simulation*. Academic Press, San Diego: CA, USA, 2nd edition.

AUTHORS BIOGRAPHY

Mamadou Seck is an assistant professor in the Systems Engineering Group of the Faculty of Technology, Policy and Management of Delft University of Technology. He received his Ph.D. degree and his M.Sc. from the Paul Cezanne University of Marseille, and his M.Eng. from Ecole Polytechnique Universitaire de Marseille, France. His research interests include modeling and simulation formalisms, human behavior representation, social simulation, and agent directed simulation. His e-mail address is <m.d.seck@tudelft.nl> and his website is <http://www.tudelft.nl/mseck>.

A DECISION SUPPORT SYSTEM FOR INVENTORY MANAGEMENT OF INTERMITTENT DEMAND

Katrien Ramaekers[(a)], Gerrit K. Janssens[(b)]

[(a) (b)]Transportation Research Institute (IMOB) – Hasselt University
Universiteit Hasselt, campus Diepenbeek
Wetenschapspark – gebouw 5, 3590 Diepenbeek

[(a)]katrien.ramaekers@uhasselt.be, [(b)]gerrit.janssens@uhasselt.be

ABSTRACT
Demand forecasting is one of the most crucial aspects of inventory management. For intermittent demand, i.e. demand peaks follow several periods of zero or low demands, forecasting is difficult. Furthermore, the choice of the forecasting method can have an impact on the inventory management policy that is best used. A simulation model is used to study a single-product inventory system facing demand of the intermittent type. In this paper, a decision support system is presented to choose between several forecasting methods and inventory management policies for intermittent demand.

Keywords: simulation-optimization, intermittent demand, forecasting, inventory management

1. INTRODUCTION

Inventory systems have to cope with uncertainty in demand. The inventory control literature mostly makes use of the Normal or Gamma distribution for describing the demand in the lead-time. The Poisson distribution has been found to provide a reasonable fit when demand is very low (only a few pieces per year). Less attention has been paid to irregular demand. This type of demand is characterised by a high level of variability, but may be also of the intermittent type, i.e. demand peaks follow several periods of zero or low demands. In practice, items with intermittent demand include service or spare parts and high-priced capital goods. A common example of such goods are spare parts for airline fleets.

Demand forecasting is one of the most crucial aspects of inventory management (Willemain et al. 2004). However, for intermittent demand, forecasting is difficult, and errors in prediction may be costly in terms of obsolescent stock or unmet demand (Syntetos and Boylan 2005). The standard forecasting method for intermittent demand items is considered to be Croston's method (Croston 1972). This method builds estimates taking into account both demand size and the interval between demand occurrences. Despite the theoretical superiority of such an estimation procedure, empirical evidence suggests modest gains in performance when compared with simpler forecasting techniques (Syntetos and Boylan 2001). Furthermore, the choice of the forecasting method can have an impact on the inventory management policy that is best used.

In this paper, a decision support system is presented to choose between several forecasting methods and inventory management policies for intermittent demand. Because of the uncertainty present in the inventory system, often mathematical models cannot accurately describe the system. Therefore, a simulation model is used. The simulation model is optimised to find the best strategy in combining inventory decision making and demand forecasting for intermittent demand. However, the best strategy depends on uncontrollable factors, i.e. the costs of the inventory system and the distribution of demand during lead time.

A good decision support system is necessary because there is a considerable increase in the total costs of the inventory system when not using the best strategy. The decision support system is presented as a decision tree where levels of the uncontrollable factors indicate which strategy in combining inventory decision making and demand forecasting is best chosen.

The organisation of the paper is as follows: in section 2 the simulation model and research approach are described; the experimental environment is described in section 3; section 4 discusses the results of the simulation model and presents the decision support system and in section 5 conclusions are formulated.

2. SIMULATION MODEL AND RESEARCH APPROACH

2.1. Simulation Model

The study focuses on a single-product inventory system facing demand of the intermittent type. The simulation model is developed in Microsoft Excel spreadsheets and uses VBA. The simulation model starts by generating intermittent demand as described in the previous section. Next, the inventory system is simulated for 52 periods. At each review-time, a demand forecast and an

order decision are made. The total costs of the inventory system are determined. 10 replications are made for each simulation run.

To generate intermittent demand, demand occurrence and demand size are separately generated. The demand occurrence is generated according to a first-order Markov process with transition matrix

$$\mathbf{P}=\begin{pmatrix} p_{00} p_{01} \\ p_{10} p_{11} \end{pmatrix},$$

where p_{00} is the probability of no order in the next period when there has been no order in this period and p_{10} is the probability of no order in the next period when there has been an order in the current period. Individual order sizes are generated using a Gamma distribution with shape parameter γ and scale parameter β.

The standard forecasting method for intermittent demand items is considered to be Croston's method. However, in practice, single exponential smoothing and simple moving averages are often used to deal with intermittent demand. These three forecasting methods are compared.

In this research, two periodic review models are used. The first one is the (R, s, S) system. This means that every R units of time, the inventory level is checked. If it is at or below the reorder point s, a sufficient quantity is ordered to raise it to S. The second system (R, s, Q) is similar to the (R, s, S) system but uses a fixed order quantity Q instead of an order-up-to-level S.

A deterministic lead-time L is assumed. The following costs are considered: unit holding cost per period C_h, ordering cost C_o and unit shortage cost per period C_s. The simulation starts with an initial inventory level I_0.

2.2. Experimental Design

The parameters of the inventory system to optimise include both qualitative and quantitative factors. The experimental design includes two qualitative factors: the forecasting method and the inventory management policy. In addition, depending on the choice of the qualitative factors, a set of quantitative factors are part of the experimental design. If the (R, s, Q) inventory management policy is used, the safety stock SS and order quantity Q are the parameters to optimise. If the (R, s, S) inventory management policy is used, the safety stock SS and order-up-to-level S are the optimising parameters. For single exponential smoothing and Croston's method, the smoothing parameter α is optimised and for moving averages, the weights of the past values are optimised.

2.3. Research Approach

Because of the dependence of the quantitative factors on the choice of the qualitative factors, we use for the optimisation the research approach described in this section.

For every combination of forecasting method, inventory management policy and review period, the optimal values of the quantitative factors are determined. The total costs of the inventory system are optimised using tabu search. Tabu search is shortly described below. Once the optimal values are found, the best combination of forecasting method, inventory management policy and review period is chosen.

Tabu search uses a local or neighbourhood search procedure to iteratively move from one solution to the next in the neighbourhood of the first, until some stopping criterion has been satisfied. To explore regions in the search space that would be left unexplored by the local search procedure and escape local optimality, tabu search modifies the neighbourhood structure of each solution as the search progresses. The solutions admitted to the new neighbourhood are determined through the use of special memory structures. Tabu search uses both long-term and short-term memory, and each type of memory has its own special strategies (Dengiz and Alabas 2000, Glover 1989).

Tabu search is a heuristic optimisation technique developed specifically for combinatorial problems. Very few works deal with the application to the global minimization of functions depending on continuous variables. The method we propose in this paper is based on (Chelouah and Siarry 2000, Siarry and Berthiau 1997). The purpose in these papers is to keep as close as possible to original tabu search. Two issues must be examined: the generation of current solution neighbours and the elaboration of the tabu list.

To define a neighbourhood of the current solution, a set of hyperrectangles is used for the partition of the current solution neighbourhood. The k neighbours of the current solution are obtained by selecting one point at random inside each hyperrectangular zone.

Once a new current solution is determined, the immediate neighbourhood of the previous solution is added to the tabu list.

3. EXPERIMENTAL ENVIRONMENT

The experimental environment contains the uncontrollable factors of the inventory system: the costs of the inventory system and the parameters for generating intermittent demand. These factors can have an effect on the results that are obtained. The research approach described above, is executed using a single combination of the costs of the inventory system and demand. A fractional factorial design of 16 experimental points is set up for these factors and the optimisation phase is repeated for each experimental point.

Demand occurrence is generated using a first-order Markov process with transition matrices:

$$\mathbf{P_1}=\begin{pmatrix} 0.7875 & 0.2125 \\ 0.85 & 0.15 \end{pmatrix}$$

or

$$\mathbf{P_2}=\begin{pmatrix} 0.5667 & 0.4333 \\ 0.65 & 0.35 \end{pmatrix}.$$

They correspond with a probability of 20% to have demand in a certain period for the first matrix and a probability of 40% to have demand in a period for the second matrix. The *size of demand* is generated using a Gamma distribution with 4 different combinations of the scale parameter γ and the shape parameter β. These values are summarized in Table 1.

Table 1: Parameters of the Gamma Distribution

Level	γ	β
1	6	1
2	12	1
3	3	2
4	24	0.5

The levels of the *costs* of the inventory system are given in Table 2. The initial inventory level I_0 equals 5.

Table 2: Levels for the Costs of the System

Level	C_o	C_h	C_s
1	100	2	5
2	200	4	10

The fractional factorial design is shown in Table 3. This fractional factorial design makes it possible to determine the impact of uncontrollable factors as the cost structure and the demand during lead time.

4. RESULTS

Each run of a single experiment from the fractional factorial design leads to a best inventory policy, together with its set of optimal parameter values, and to a best forecasting method, together with its set of optimal parameter values (Table 4). This section aims to investigate which design factors have an influence on the choice of inventory policy and forecasting method. At first, a detailed study is made of the influence of each individual factor, and afterwards an attempt is made to simplify and structure these findings in a decision support system, which is generated using a classification tree.

Eight experimental points have the order-up-to-level (OUL) inventory management policy with S=1 as best strategy but with various best forecasting methods. For the other eight experimental points, the best strategy is an OUL-inventory management policy with S≥15 or a fixed order quantity (FOQ) policy with Q≥15. An FOQ-inventory management policy as best goes together with the moving averages (MA)-method as best forecasting

Table 3: Experimental Design

Exp	C_o	C_h	C_s	Freq	γ	β
1	200	4	10	0.4	12	1
2	100	4	5	0.4	12	1
3	200	2	5	0.4	24	0.5
4	100	2	10	0.4	24	0.5
5	200	2	5	0.4	3	2
6	100	2	10	0.4	3	2
7	200	4	10	0.4	6	1
8	100	4	5	0.4	6	1
9	200	2	10	0.2	12	1
10	100	2	5	0.2	12	1
11	200	4	5	0.2	24	0.5
12	100	4	10	0.2	24	0.5
13	200	4	5	0.2	3	2
14	100	4	10	0.2	3	2
15	200	2	10	0.2	6	1
16	100	2	5	0.2	6	1

method. Also in case the OUL-inventory management policy with S≥15 is best, MA shows to be the best forecasting method. In case the OUL-inventory management policy with S=1 is best, no specific forecasting method is preferred. The results also indicate that the parameters of the forecasting method have no significant impact on the results. In the next paragraphs, the influence of the uncontrollable factors on the results is examined in further detail.

When the *demand frequency* is generated using matrix P_1, corresponding to a probability of 20% of having demand in a certain period, an order-up-to-level S of 1 unit is optimal. When the demand frequency is generated using matrix P_2, which corresponds to a probability of 40% of having demand in a certain period, the order-up-to-level S or fixed order quantity Q is a value between 15 and 30. This can be explained because the intermittent character of demand is more distinct when the probability of demand is equal to 20%, leading to an optimal order-up-to-level S of 1 unit. When the intermittent character of demand is less distinct (40%), it is better to order a quantity of at least 15 units. The only exception to this order-up-to-level S of 1 unit for a demand probability of 20% can be found when both the ordering cost and the unit shortage cost are high and the unit holding cost is low. In these circumstances it is better to order a bigger quantity

because it is less costly to hold inventory than to have a stock-out or to order a small quantity every time.

Table 4: Optimal Results based on Tabu Search

Exp	Best strategy
1	MA/FOQ; ROP=0; Q=25
2	ES/OUL; ROP=0; S=1
3	MA/OUL; ROP=0; S=30
4	MA/OUL; ROP=0; S=25
5	MA/FOQ; ROP=0; Q=20
6	MA/OUL; ROP=0; S=15
7	MA/FOQ; ROP=0; Q=15
8	ES/OUL; ROP=0; S=1
9	MA/FOQ; ROP=0; Q=20
10	ES/OUL; ROP=0; S=1
11	CR/OUL; ROP=0; S=1
12	ES/OUL; ROP=0; S=1
13	MA/OUL; ROP=0; S=1
14	MA/OUL; ROP=0; S=1
15	MA/OUL; ROP=0; S=15
16	ES/OUL; ROP=0; S=1

Inversely, when a demand probability of 40% is used, it is better to use an order-up-to-level S of 1 unit when both the ordering cost and the unit shortage cost are low and the unit holding cost is high.

When comparing results for changing the parameters of the *demand size*, no significant impact of these changes on the results can be detected.

Changes in the cost structure of the inventory system have a significant impact on the results. When the *ordering cost* is equal to 100, an order-up-to-level inventory management policy is used with the order-up-to-level S equal to 1, except when the unit holding cost is low, the unit shortage cost is high and the demand probability of a certain period is 40%. The level of these three factors all favour holding more units in inventory. The combination of these three levels therefore changes the best policy to a policy with an order-up-to-level or fixed order quantity between 15 and 30, although the order cost is low. When the ordering cost is equal to 200, the order-up-to-level S or fixed order quantity Q is between 15 and 30, except when the unit holding cost is high, the unit shortage cost is low and the demand probability of a period equals 20%.

When the *unit holding cost* is equal to 2, an order-up-to-level S or fixed order quantity Q between 15 and 30 is used, unless both the ordering cost and the unit shortage cost are also low and the demand probability of a period equals 20%. When this combination of factor levels occurs, an inventory policy with an order-up-to-level S equal to 1 is better used because all these factor levels give preference to a lower inventory level. When the unit holding cost equals 4, an order-up-to-level S of 1 is the best choice, unless the ordering cost and unit shortage cost are also high and the demand probability of a period is 40%. This combination of factor levels favours a higher inventory level and thus an order-up-to-level or fixed order quantity between 15 and 30 is better used.

A *unit shortage cost* of 5 implies an order-up-to-level S of 1 unit, except when the unit holding cost is also low and the probability of demand for a certain period equals 40%. When the shortage cost is low, it is not necessary to keep a lot of units in inventory. Therefore, an order-up-to-level equal to 1 is the best policy. However, if the holding cost is also low and the intermittent character of demand is not so distinct, it is better to have more units in inventory even though the shortage cost is low. Doubling the unit shortage cost leads to an order-up-to-level S or fixed order quantity Q between 15 and 30, except when the unit holding cost is high and the demand frequency is equal to 20%.

Overall, it can be concluded that the uncontrollable factors have an impact on the best strategy for combining inventory decision-making and demand forecasting for intermittent demand. Furthermore, there is interaction between these factors.

To structure these findings, a decision support system is developed using a classification tree. The classification tree is constructed using the C4.5 algorithm, a well-known algorithm in data mining (Quinlan 1993). The classification tree can be found in Figure 1. Using this tree, it can be decided which of the two strategies is best: an order-up-to-level inventory management policy with S=1 or an order-up-to-level inventory management policy with S≥15 or a fixed order quantity model with Q≥15. Three factors are needed to determine the best strategy in combining inventory decision making and demand forecasting: the frequency of demand, the order cost and the inventory cost. If one of these three factors is not known, the knowledge of the stock-out cost is also sufficient to make a classification. Summarizing, it can be said that if three factors of the four just mentioned (frequency of demand, order cost, inventory cost and stock-out cost) are fixed, the best strategy is presented.

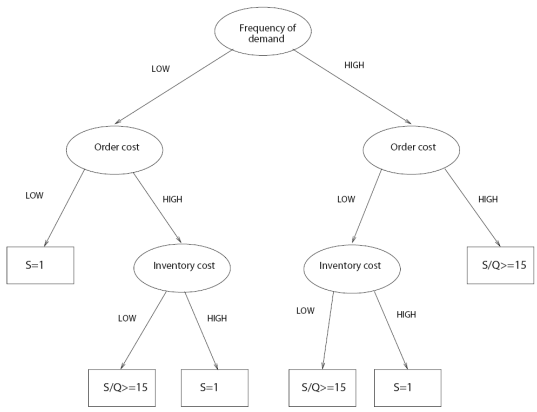


Figure 1: Classification tree

A good classification is necessary because there is a considerable increase in the costs of the inventory system when using the other strategy. When a fixed order quantity inventory management policy with Q=15 is used instead of an order-up-to-level inventory

management policy with S=1, total costs are on average 20% higher. In the opposite case, when an order-up-to-level inventory management policy with S=1 is used instead of an order-up-to-level inventory management policy with $S \geq 15$ or a fixed order quantity model with $Q \geq 15$, total costs increase with more than 40% on average.

5. CONCLUSIONS

In this paper a decision support system is presented to choose between several forecasting methods and inventory management policies for intermittent demand. A best strategy in combining inventory decision making and demand forecasting is proposed, using a simulation model. An experimental design is set up to determine the impact of uncontrollable factors: the cost structure and the demand. Depending on the experimental environment, two options for optimal strategies can be distinguished: an order-up-to level inventory management policy with an order-up-to level equal to 1 and a reorder point equal to 0 or an inventory management policy with a fixed order quantity Q>1 or an order-up-to level S>1 and a reorder point equal to 0. Four factors of the experimental environment have an influence on which of the two strategies is best chosen: the frequency of demand, the inventory holding cost, the order cost and the stock-out cost. When the level of three factors out of these four are fixed, it is possible to determine the optimal strategy. To structure these findings, a decision support system is developed using a classification tree. It is important to know which of both strategies is best because there is a significant increase in total costs of the inventory system if the wrong strategy is chosen.

REFERENCES

Chelouah R. and Siarry P.: Tabu search applied to global optimization, *European Journal of Operational Research* 123, 2000, pp. 256-270.

Croston J.D.: Forecasting and stock control for intermittent demands, *Operational Research Quarterly* 23, 1972, pp. 289-303.

Dengiz B. and Alabas C.: Simulation optimisation using tabu search, *In Proceedings of the 2000 Winter Simulation Conference*, Orlando, 10-13 December 2000, pp. 805-810.

Glover F.: Tabu search- Part I, *ORSA Journal on Computing* 1(3), 1989, pp. 190-206.

Quinlan J.: *C4.5 Programs for Machine Learning*, 1993, Morgan Kaufmann Publishers.

Siarry P. and Berthiau G.: Fitting of tabu search to optimize functions of continuous variables, *International Journal of Numerical Methods in Engineering* 40, 1997, pp. 2449-2457.

Syntetos A.A. and Boylan J.E.: On the bias of intermittent demand estimates, *International Journal of Production Economics* 71, 2001, pp. 457-466.

Syntetos A.A. and Boylan J.E.: The accuracy of intermittent demand estimates, *International Journal of Forecasting* 21, 2005, pp. 303-314.

Willemain T.R., Smart C.N. and Schwarz H.F.: A new approach to forecasting intermittent demand for service part inventories, *International Journal of Forecasting* 20, 2004, pp. 375-387.

BIOGRAPHY

Katrien Ramaekers graduated as Master of Business Economics at the Limburg University Centre in 2002. In 2007, she obtained her Ph.D. in Applied Economic Sciences at Hasselt University. Her Ph.D. research is on the integration of simulation and optimisation, especially as a support for complex logistics decision-making. Currently, she is a post-doctoral researcher at Hasselt University and is working on the modelling of freight transport. She is a member of the Transportation Research Institute of Hasselt University.

Gerrit K. Janssens received degrees of M.Sc. in Engineering with Economy from the University of Antwerp (RUCA), Belgium, M.Sc. in Computer Science from the University of Ghent (RUG), Belgium, and Ph.D. from the Free University of Brussels (VUB), Belgium. After some years of work at General Motors Continental, Antwerp, he joined the University of Antwerp until the year 2000. Currently he is Professor of Operations Management and Logistics at Hasselt University within the Faculty of Business Administration. He also holds the CPIM certificate of the American Production and Inventory Control Society (APICS). His main research interests include the development and application of operations research models in production and distribution logistics.

A LINEAR PROGRAMMING FORMULATION FOR AN INVENTORY MANAGEMENT DECISION WITH A SERVICE CONSTRAINT

Gerrit K. Janssens[(a)], Katrien M. Ramaekers[(b)]

[(a) (b)]Hasselt University - Campus Diepenbeek
Wetenschapspark 5 – bus 6
B-3590 Diepenbeek, Belgium

[(a)] gerrit.janssens@uhasselt.be [(b)] katrien.ramaekers@uhasselt.be

ABSTRACT
Inventory systems with uncertainty go hand in hand with the determination of a safety stock level. The decision on the safety stock level is based on a performance measure, for example the expected shortage per replenishment period or the probability of a stock-out per replenishment period. The performance measure assumes complete knowledge of the probability distribution during lead time, which might not be available. In case of incomplete information regarding the lead-time distribution of demand, no single figure for the safety stock can de determined in order to satisfy a performance measure. However, an optimisation model may be formulated in order to determine a safety stock level which guarantees the performance measure under the worst case of lead-time demand, of which the distribution is known in an incomplete way. It is shown that this optimisation problem can be formulated as a linear programming problem.

Keywords: inventory management, linear programming, incomplete information

1. INTRODUCTION

Some uncertainty in an inventory system (such as lead time, quantity and quality) depends on the suppliers. If the suppliers introduce too much uncertainty, corrective action should be taken. Some uncertainty, however, is attributable to customers, especially demand. If insufficient inventory is hold, a stock-out may occur leading to shortage costs. Shortage costs are usually high in relation to holding costs. Companies are willing to hold additional inventory, above their forecasted needs, to add a margin of safety.

Determination of an inventory replenishment policy, of the quantities to order, of the review period are typical decisions to be taken by logistics managers. Decisions are made through optimisation models taking a performance measure into consideration which might be cost-oriented or service-oriented. Performance measures of the service-oriented type may be expressed relatively as a probability of a stock-out during a certain replenishment period, or may be expressed absolutely in terms of number of units short, which is a direct indication for lost sales. Both performance measures are taken into consideration and special attention will be paid to feasible combinations of company's objectives regarding both performance measures.

For a definition of both measures we refer to chapter 7 in Silver, Pyke and Peterson (1998) and define the measures as:

The *expected shortage per replenishment cycle* (ESPRC) is defined as (with t the amount of safety stock):To avoid any difficulties during the publishing process, authors must not modify any of the styles.

The *expected shortage per replenishment cycle* (ESPRC) is defined as (with t the amount of safety stock):

$$ESPRC = \int_{d}^{+\infty} (x-t) f(x) dx \tag{1}$$

If ordered per quantity Q the fraction backordered is equal to ESPRC/Q and a performance measure, indicated as *P2*, is defined as

$$P_2 = 1 - ESPRC / Q \tag{2}$$

The other performance measure is the *probability of a stock-out in a replenishment lead time*, defined as:

$$1 - P_1 = \Pr\{x \geq t\} = \int_{d}^{+\infty} f(x) dx \tag{3}$$

From a production or trading company's point of view, a decision might be formulated to answer the following question: *given a maximum expected number of units short and/or a maximum stock-out probability the company wants to face, what should be the safety inventory at least (or at most)?* The question with the 'at most' option might be only of academic nature, as it reflects the most optimistic viewpoint. In human terms, this question would be interpreted as: 'would there exist any probability distribution so that I can still reach my preset performance criteria given a specific safety inventory?'. This type of question is not relevant for a manager facing a real-life situation.

In case the distribution of demand is known, determining the inventory level, given a maximum shortage or maximum stock-out probability, reduces to the calculation of the inverse cumulative probability function. The decision problem becomes more difficult if incomplete information exists on the distribution of demand during lead time, for example only the range of

demand, or the first moment, or the first and second moments are known. In such a case no single value can be determined but rather an interval.

In classical textbooks not too much attention is paid to the shape of the distribution of the demand during lead time. Mostly, based on the first and second moments, the safety stock level is determined using the normal distribution. When of relevance, one rather should look for a distribution, which is defined only for non-negative values and allows for some skewness. In the literature on inventory control, frequent reference is made to the Gamma distribution.

It is generally known that, given a shape of the demand distribution, the higher the coefficient of variation the more a company needs inventory to reach a given service level. In an investigation on the relevance of the demand shape Bartezzaghi, Verganti, and Zotteri (1999) find out that the shape is very relevant. In extreme cases the impact of different demand shapes on inventories is comparable to the effect of doubling the coefficient of variation.

This research deals with the case where the demand distribution during lead time is not completely known. This situation is realistic either with products which have been introduced recently to the market or with slow moving products. In both cases not sufficient data are available to decide on the functional form of the demand distribution function. Some but not complete information might exist like the range of the demand, its expected value, its variance and maybe some knowledge about uni-modality of the distribution.

In case incomplete information is available regarding the demand distribution the integrals of the performance measures P_1 and P_2 cannot be evaluated in an analytical manner. This means that also the inverse problem of determining the safety stock level to satisfy the performance measures cannot be obtained analytically. However, the integrals can be approximated by a linear programming formulation with a large set of constraints.

2. BOUNDS ON THE PERFORMANCE MEASURES IN THE CASE OF INCOMPLETE INFORMATION

In this section the ESPRC measure is focused. First, a link is identified with a similar integral formulation which appears in the field of actuarial sciences. Second, some results, which were obtained in actuarial sciences, are transferred to our type of application.

2.1 Towards an analogy in insurance mathematics

In insurance mathematics, an insurance company using the option of re-insurance is confronted with a stop-loss premium. A stop-loss premium limits the risk X of an insurance company to a certain amount d. If the claim size is higher than d the re-insurance company takes over the risk X-d. The stop-loss premium is based on the expected value of X-d, which in case of a known claim size distribution may be defined as:

$$\int_0^\infty (x-t)_+ \, dF(x) \tag{4}$$

where $F(x)$ represents the claim size distribution (Goovaerts, De Vylder, and Haezendonck 1984).

The same formula (4) may be useful in the performance evaluation of inventory management in case of uncertain demand during lead time. When a company holds t units of a specific product in inventory starting a period between order and delivery, any demand less than t is satisfied while any demand X greater than t results in a shortage of X-t units. A lesser number of units short results in a better service to the customer. In this way formula (4) is a measure for customer service in inventory management.

In the following sections lower and upper bounds are obtained for the performance measure under study, given various levels of information about the demand distribution. From a production or trading company's point of view, a decision might be formulated to answer the following question: *given an expected number of units short the company wants to face, what should be the safety inventory at least or at most?*

2.2 The case of known range, mean and variance

Let the size of the demand X for a specific product in a finite period have a distribution F with first two moments $\mu_1 = E(X)$ and $\mu_2 = E(X^2)$.

From a mathematical point of view, the problem is to find the following bounds:

$$\sup_{F \in f} \int_0^\infty (x-t)_+ \, dF(x) \tag{5a}$$

and

$$\inf_{F \in f} \int_0^\infty (x-t)_+ \, dF(x) \tag{5b}$$

where $Ø$ is the class of all distribution functions F which have moments μ_1 and μ_2, and which have support in R^+. Let further $s^2 = \mu_1 - \mu_2^2$. We assume t to be strictly positive.

For any polynomial $P(x)$ of degree 2 or less, the integral

$$\int_0^\infty P(x) dF(x)$$

only depends on μ_1 and μ_2, so it takes the same value for all distributions in $Ø$. There exists some distribution G in Ø for which the equality holds:

$$\int_0^\infty P(x) dG(x) = \int_0^\infty (x-t)_+ \, dG(x). \tag{6}$$

As distribution G a two-point or three-point distribution is used. The equality (6) is attained when $P(x)$ and $(x-t)_+$ are equal in both points of G. The best upper and lower bounds on this term with given moments μ_1 and μ_2 are derived. The method is inspired by papers of Janssen, Haezendock, and Goovaerts (1986) and by Heijnen and Goovaerts (1989). In the following we assume the known range of the distribution to be a finite interval $[a,b]$.

A probability distribution F is called n-atomic if all its probability mass is concentrated in n points at most. The points are called the atoms of the distributions. The problem (5a) has a 2-atomic solution and (5b) has a 3-atomic solution.

If α, β are two different atoms of the 2-atomic probability distribution F satisfying the first-order moment constraint $\int xdF = m_1$, then the corresponding probability masses p_a and p_b are

$$p_a = \frac{m_1 - b}{a - b}, \; p_b = \frac{m_1 - a}{b - a}. \tag{7}$$

If α, β, γ are three different atoms of the 3-atomic probability distribution F satisfying the moment constraints $\int xdF = m_1, \int x^2 dF = m_2$, then the corresponding probability masses p_a, p_b and p_g are

$$\begin{aligned} p_a &= \frac{s^2 + (m_1 - b)(m_1 - g)}{(a - b)(a - g)}, \\ p_b &= \frac{s^2 + (m_1 - a)(m_1 - g)}{(b - a)(b - g)}, \\ p_g &= \frac{s^2 + (m_1 - a)(m_1 - b)}{(g - a)(g - b)} \end{aligned} \tag{8}$$

The domain of the parameters is

$$\begin{aligned} &a \le m_1 \le b, \quad 0 \le s^2 \le (m_1 - a)(b - m_1) \; or \\ &m_1^2 \le m_2 \le m_1(a+b) - ab \end{aligned}.$$

Further the following abbreviations are used: $s_{mt}^2 = s^2 + (m_1 - t)^2$ and $c = \frac{1}{2}(a+b)$. Further let m_1 and m_2 be chosen that the previous inequalities hold, then let $r' = \frac{m_2 - m_1 r}{m_1 - r}$ for every $r \in [a,b]$ and $r \ne m_1$.

Before moving towards the application, it should be stated that the bounds and their use in applications can be translated from any distribution defined on $[a,b]$ into the bounds with a distribution defined on $[0,b_0]$, where $b_0 = b - a$. Further let $t_0 = t - a$, $m_{10} = m_1 - a$ and $m_{20} = m_2 - 2am_{10} - a^2$. In the following paragraphs we work, without loss of generalisation, with distributions defined on $[0,b_0]$.

The use of the bounds is illustrated by means of a numerical example. Let the demand be defined on the interval [25,75]. The demand follows a distribution with only the following characteristics known: $\mu_1 = 45$ and $\mu_2 = 975$. This means that in Tables 1 and 2, the following values have to be used for $m_{10} = 20$, $m_{20} = 600$, $b_0 = 50$, $b_0' = 13.333$, and $0' = 30$. The values for upper and lower bounds are shown in tables 3 and 4. From Tables 3 and 4, a decision-maker may be decide which level of inventory to hold, given a target value on the number of units short W as a performance measure. From these tables he can derive upper bounds on t_0, which correspond to a pessimistic viewpoint and lower bounds on t_0, which correspond to an optimistic viewpoint. The values corresponding to both viewpoints for the numerical example under study are given in tables 5 and 6.

Table 1: Lower Bounds on the Stop-loss Premium in an interval $[0,b_0]$

Lower bounds	Conditions
$m_{10} - t_0$	$0 \le t_0 \le b_0'$ or $0 \le t_0 \le (m_{20} - m_{10}b_0)/(m_{10} - b_0)$
$(m_{20} - m_{10}t_0)/b_0$	$b_0' \le t_0 \le 0'$ or $(m_{20} - m_{10}b_0)/(m_{10} - b_0) \le t_0 \le m_{20}/m_{10}$
0	$0' \le t_0 \le b_0$ or $m_{20}/m_{10} \le t_0 \le b_0$

Table 2: Upper Bounds on the Stop-loss Premium in an interval $[0,b_0]$

Upper bounds	Conditions
$m_{10}(m_{20} - m_{10}t_0)/m_{20}$	$t_0 \le 0'/2$ or $t_0 \le m_{20}/(2m_{10})$
$(m_{10} - t_0 + \sqrt{(m_{20} - m_{10}^2) + (t_0 - m_{10})})/2$	$0'/2 \le t_0 \le (b_0 + b_0')/2$ or $m_{20}/(2m_{10}) \le t_0 \le (m_{20} - b_0^2)/(2(m_{10} - b_0))$
$(m_{20} - m_{10}^2)(b_0 - t_0)/((m_{20} - m_{10}^2) + (b_0 - m_{10})^2)$	$(b_0 + b_0')/2 \le t_0$ or $(m_{20} - b_0^2)/(2(m_{10} - b_0)) \le t_0$

Table 3: Lower Bounds on the Number of Units Short for the illustrative example

Lower bounds	Conditions
$20 - t_0$	$0 \le t_0 \le 13.333$
$12 - 2/5\, t_0$	$13.333 \le t_0 \le 30$
0	$30 \le t_0 \le 50$

Table 4: Upper Bounds on the Number of Units Short for the illustrative example

Upper bounds	Conditions
$20 - 2/3\, t_0$	$0 \le t_0 \le 15$
$(20 - t_0 + \sqrt{200 + (t_0 - 20)})/2$	$15 \le t_0 \le 31.667$
$(100 - 2t_0)/11$	$31.667 \le t_0 \le 50$

Table 5: Lower bounds on the safety inventory level

Lower bounds	Requirements
$30 - 5/2\, W$	$W \le 6.667$
$20 - W$	$6.667 \le W$

Table 6: Upper bounds on the safety inventory level

Upper bounds	Requirements
$50 - 11/2\, W$	$W \le 3.333$
$(50 - W^2 + 20W)/W$	$3.333 \le W \le 10$
$(60 - 3W)/2$	$10 \le W$

3. A METHOD TO DETERMINE SAFETY STOCK IN THE CASE OF INCOMPLETE INFORMATION ON DEMAND

It has been shown in Janssens and Ramaekers (2008) how the optimisation problem (5a) with constraints interms of first and second moment of the demand distributions, has a dual program which is a linear program with an infinite number of constraints. In Goovaerts, Haezendonck, and De Vylder (1982) an idea is launched to replace the set of constraints by a large finite subset and then to solve the so obtained linear program.

The method assumes that integral constraints can be transformed into a sequence, with increasing number of evaluation points, of optimisation problems and where the integral is replaced by an infinite sum. Instead of evaluating the objective function on a continuous interval [low,high], the functions are evaluated in a discrete number of points x_i ($i = 1..N$). The assumption reflects the idea that if $N \rightarrow \infty$ the solution of the continuous problem is found.

This leads to an optimisation problem, where:

t = the level of the safety inventory

p_i = the probability mass in point x_i

z_1 = the expected value of X

z_2 = the absolute second moment of X

z_3 = the maximum allowed number of items short.

The optimisation problem might be formulated as:

[P1] Min t

Subject to

$$\sum_i p_i = 1$$

$$\sum_i x_i p_i = z_1$$

$$\sum_i x^2{}_i p_i = z_2$$

$$\sum_i (x_i - t)_+ p_i \le z_3$$

where $(x_i\text{-}t)_+$ stands for $\max(x_i\text{-}t,0)$. The decision variables in [P1] are t and p_i (i= 1..N), where N represents the number of discrete points which have been chosen in the experiment.

Problem [P1] offers the answer to the following question: what is the minimal amount of inventory so that a distribution with given characteristics exists in which the expected number short maximally equals the value z_3.

The non-linear constraint may be approximated by letting the value of t coincide with one of the x_i –values (so as $N \rightarrow \infty$, the approximation takes the correct value). In such a way the constraint in linearised.

In the case t coincides with a point x_j then

$$\sum_{i=1}^{n} p_i (x_i - x_j)_+ \le z_3 \ .$$

A binary variable needs to be introduced to indicate the condition ‘t = x_j’. In the case t does not coincide with a point x_j, a general truth should be indicated, for example, ‘the expected number short cannot be larger than the expected demand’, expressed by a binary variable y_j.

$$y_j = 1 \; if \quad t = x_j$$
$$else \quad 0 \quad .$$

As t can coincide with only one x_j-value, the additional constraint is introduced:

$$\sum_{j=1}^{n} y_j = 1.$$

The *y*-variable is introduced in the last constraint as:

$$\sum_{i=1}^{n} p_i (x_i - x_j)_+ \le z_3 y_j + z_1 (1 - y_j)$$

Finally a link should be made between *t* and the value of *x* with which *t* coincides

$$t \ge x_i y_i, \forall i$$

If *y* = 0, a universal truth is mentioned.

This elaboration will be illustrated by means of the example used in the previous section. With b_0 = 50, the first and second moments in the interval [0,50], the following values μ_{10} = 20 and μ_{20} = 600 are used. The worked out example, in LINDO code, is shown in figure 1. In this approximation 10 intervals of equal length in the interval [0,50] are chosen. Inclusion of both boundaries of the interval, the linear program makes use of 11 x_i-variables.

Take for example the maximum number of units short W = 6. From table 5, it can be obtained that the lower bound for t equals t = 15. The linear program in figure 1 leads to a minimum of t = 15, with probability mass in three evaluation points x_1 (X=0), x_4 (X=15) and x_{11} (X=50). The respective probability masses are: p_1 = 0.06667, p_4 = 0.76195 and p_{11}= 0.17143.

```
min t
subject to
p1 + p2 + p3 + p4 + p5 + p6 + p7 + p8 + p9 + p10 + p11 = 1
0 p1 + 5 p2 + 10 p3 + 15 p4 + 20 p5 + 25 p6 + 30 p7 + 35 p8 + 40 p9 + 45 p10 + 50 p11 = 20
0 p1 + 25 p2 + 100 p3 + 225 p4 + 400 p5 + 625 p6 + 900 p7 + 1225 p8 + 1600 p9 + 2025 p10 + 2500 p11 = 600
5 p2 + 10 p3 + 15 p4 + 20 p5 + 25 p6 + 30 p7 + 35 p8 + 40 p9 + 45 p10 + 50 p11 + 14 y1 < 20
5 p3 + 10 p4+ 15 p5 + 20 p6 + 25 p7 + 30 p8 + 35 p9 + 40 p10 + 45 p11 + 14 y2 < 20
5 p4 + 10 p5 + 15 p6 + 20 p7 + 25 p8 + 30 p9 + 35 p10 + 40 p11 + 14 y3 < 20
5 p5 + 10 p6 + 15 p7 + 20 p8 + 25 p9 + 30 p10 + 35 p11 + 14 y4 < 20
5 p6 + 10 p7 + 15 p8 + 20 p9 + 25 p10 + 30 p11 + 14 y5 < 20
5 p7 + 10 p8 + 15 p9 + 20 p10 + 25 p11 + 14 y6 < 20
5 p8 + 10 p9 + 15 p10 + 20 p11 + 14 y7 < 20
5 p9 + 10 p10 + 15 p11 + 14 y8 < 20
5 p10 + 10 p11 + 14 y9 < 20
5 p11 + 14 y10 < 20
14 y11 < 20
1 t > 0
1 t - 5 y2 > 0
1 t - 10 y3 > 0
1 t - 15 y4 > 0
1 t - 20 y5 > 0
1 t - 25 y6 > 0
1 t - 30 y7 > 0
1 t - 35 y8 > 0
1 t - 40 y9 > 0
1 t - 45 y10 > 0
1 t - 50 y11 > 0
y1 + y2 + y3 + y4 + y5 + y6 + y7 + y8 + y9 + y10 + y11 = 1
end
int y1 .. int y10
```

Figure 1: LINDO Code for the Illustrative Example

4. AN APPLICATION IN THE SINGLE-PERIOD (NEWSVENDOR) INVENTORY PROBLEM

The single period-inventory problem or newsvendor problem aims to decide the stock quantity of an item when there is a single purchasing opportunity before the start of the selling period and the demand for the item is unknown. A trade-off exists between the risk of overstocking (forcing disposal below the unit purchasing cost) and the risk of understocking (losing the opportunity of making a profit) (Gallego and Moon 1993). Many extensions to the newsvendor problem have been proposed in the last decades, including dealing with different objectives and utility functions, different supplier pricing policies, different news-vendor pricing policies and discounting structures, different states of information about demand, constrained multi-products, multiple-products with substitution, random yields, and multi-location models (Khouja 1999).

Assume a single product is to be ordered at the beginning of a period and can only be used to satisfy demand in that period. The relevant costs on basis of the ending inventory are:

```
min 0.35 Q + 0.9 w1 + 0.9 w2 + 0.9 w3 + 0.9 w4 + 0.9 w5 + 0.9 w6 + 0.9 w7 + 0.9 w8 + 0.9 w9 + 0.9 w10
subject to
p1 + p2 + p3 + p4 + p5 + p6 + p7 + p8 + p9 + p10 + p11 = 1
0 p1 + 5 p2 + 10 p3 + 15 p4 + 20 p5 + 25 p6 + 30 p7 + 35 p8 + 40 p9 + 45 p10 + 50 p11 = 20
0 p1 + 25 p2 + 100 p3 + 225 p4 + 400 p5 + 625 p6 + 900 p7 + 1225 p8 + 1600 p9 + 2025 p10 + 2500 p11 = 600
5 p2 + 10 p3 + 15 p4 + 20 p5 + 25 p6 + 30 p7 + 35 p8 + 40 p9 + 45 p10 + 50 p11 - w1 + 1000 y1 < 1000
5 p3 + 10 p4 + 15 p5 + 20 p6 + 25 p7 + 30 p8 + 35 p9 + 40 p10 + 45 p11 - w2 + 1000 y2 < 1000
5 p4 + 10 p5 + 15 p6 + 20 p7 + 25 p8 + 30 p9 + 35 p10 + 40 p11 - w3 + 1000 y3 < 1000
5 p5 + 10 p6 + 15 p7 + 20 p8 + 25 p9 + 30 p10 + 35 p11 - w4 + 1000 y4 < 1000
5 p6 + 10 p7 + 15 p8 + 20 p9 + 25 p10 + 30 p11 - w5 + 1000 y5 < 1000
5 p7 + 10 p8 + 15 p9 + 20 p10 + 25 p11 - w6 + 1000 y6 < 1000
5 p8 + 10 p9 + 15 p10 + 20 p11 - w7 + 1000 y7 < 1000
5 p9 + 10 p10 + 15 p11 - w8 + 1000 y8 < 1000
5 p10 + 10 p11 - w9 + 1000 y9 < 1000
5 p11 - w10 + 1000 y10 < 1000
1 Q - 0 y1 > 0
1 Q - 5 y2 > 0
1 Q - 10 y3 > 0
1 Q - 15 y4 > 0
1 Q - 20 y5 > 0
1 Q - 25 y6 > 0
1 Q - 30 y7 > 0
1 Q - 35 y8 > 0
1 Q - 40 y9 > 0
1 Q - 45 y10 > 0
1 Q - 50 y11 > 0
y1 + y2 + y3 + y4 + y5 + y6 + y7 + y8 + y9 + y10 + y11 = 1
end
int y1 .. int y11
```

Figure 2: LINDO Code of the Newsvendor Example

c_0 = cost per unit of positive inventory remaining at the end of the period (overage cost)

c_1 = cost per unit of unsatisfied demand (underage cost).

Further let:

Q: order quantity

D: random demand with a distribution F with density f defined on a finite interval $[a,b]$ with $a \geq 0$ and $b > a$.

Define $G(Q,D)$ as the total overage and underage cost incurred at the end of period when Q units are ordered at the start of the period and D is the demand. Then it follows that

$$G(Q,D) = c_o \max(0, Q-D) + c_u \max(0, D-Q) \quad (9)$$

The expected cost G(Q) = E[G(Q,D)] can be calculated as:

$$G(Q) = c_0 \int_0^Q (Q-x) f(x)dx + c_u \int_Q^\infty (x-Q) f(x)dx \quad (10)$$

(Nahmias 1993).

The newsvendor formulation also can be used to make decisions in a profit framework. This formulation needs information about the unit cost c, a mark-up m indicating the relative return per currency unit sold and a discount d indicating the loss per currency unit unsold (Gallego and Moon 1993) :

c: unit cost ($c > 0$)

p: unit selling price ($p = (1+m)c$, $m > 0$)

s: unit salvage value ($s = (1-d)c$, $d > 0$).

The expected profit in function of the order quantity, $P(Q)$, can be written as:

$$P(Q) = p\, E \min(Q,D) + sE(Q-D)_+ - cQ \quad (11)$$

since $min(Q,D)$ units are sold, $(Q-D)^+$ are salvaged, and Q units are purchased. Gallego and Moon (1993) show that maximizing $P(Q)$ is equivalent to minimizing

$$dQ + (m+d)E(D-Q)_+ \quad (12)$$

Similar to the case in section 3, the non-linear part of the objective function may be approximated by letting the value of Q coincide with one of the x_i –values (so as $N \rightarrow \infty$, the approximation takes the correct value). In such a way the constraint in linearised. The objective function takes the form:

$$dQ + (m+d)\sum_{j=1}^{N} w_j \quad (13)$$

in which the newly introduced variables w_j take the values

$$w_j = \sum_{i=1}^{n} p_i (x_i - x_j)_+$$

in the case Q coincides with a point x_j and 0 otherwise.

This logic can be introduced in some of the constraints making use of a binary variable introduced

to indicate the condition '$Q = x_j$', which is weighted with a big coefficient M as follows:

$$\sum_{i=1}^{n} p_i (x_i - x_j)_+ - w_j - M(1 - y_j) \leq 0$$

where the binary variable y_j.

$$y_j = 1 \; if \quad t = x_j$$
$$else \quad 0 \quad .$$

In case $y_j = 0$, the constraint induces no reason for w_j to take a positive value, so $w_j = 0$ and in case $y_j = 1$, the equality is obtained.

This elaboration will be illustrated by means of the same example as used in section 3. The worked out example, in LINDO code, is shown in figure 2. The following coefficients are used: $m = 0.35$, $d = 0.55$ and $M = 1000$. The code solves the program to optimality when Q = 30, facing a demand distribution of $p1 = 1/3$ and $p7 = 2/3$. It leads to an objective function value of 10.5 with no shortages. If however $m = 0.70$, $d = 0.20$ and M = 1000, then Q = 15, facing a demand distribution in three mass points with $p1 = 0.66667$, $p4 = 0.761905$ and $p11 = 0.171429$. This case leads to an objective function value of 15.9, with a shortage $w6 = 6$.

5. CONCLUSIONS

It is shown how decision regarding inventory management in the case of incomplete information on the demand distribution can be supported by making use of a linear programming formulation of the problem. At first it is illustrated using 'the expected number short during lead-time' as a performance measure, and the same idea is also applied in the newsvendor problem. A similar formulation can be developed making use of 'the probability of a stock-out during lead-time' and, of course, also the combination of both performance measures is also a challenging input to this decision-making problem.

REFERENCES

Bartezzaghi, E., Verganti, R., Zotteri G., 1999, Measuring the impact of asymmetric demand distributions on inventories. *International Journal of Production Economics*, 60-61:395-404.

Gallego, G., Moon, I., 1993, The distribution free newsboy problem: review and extensions. *Journal of the Operational Research Society*, 44(8):825-834.

Goovaerts, M.J., Haezendonck, J., De Vylder, F., 1982, Numerical best bounds on stop-loss premiums. *Insurance: Mathematics and Economics*, 1:287-302.

Goovaerts, M.J., De Vylder, F. and Haezendonck, J., 1984, *Insurance Premiums*. Amsterdam: North-Holland.

Heijnen, B., Goovaerts, M.J., 1989, Best upper bounds on risks altered by deductibles under incomplete information. *Scandinavian Actuarial Journal*, 23-46.

Janssen, K., Haezendonck, J., Goovaerts, M.J., 1986, Upper bounds on stop-loss premiums in case of known moments up to the fourth order. *Insurance Mathematics and Economics*, 5:315-334.

Janssens, G.K. Ramaekers, K.M., 2008, On the use of bounds on the stop-loss premium for an inventory management decision problem. *Journal of Interdisciplinary Mathematics*, 11(1):115-126.

Khouja, M., 1999, The single period (news-vendor) inventory problem: a literature review and suggestions for future research. *Omega: International Journal of Management and Science*, 27:537-553.

Nahmias, S., 1993, *Production and Operations Analysis*. Homewood:Irwin.

Silver, E.A., Pyke, D.F. and Peterson, R. 1998, *Inventory Management and Production Planning and Scheduling*. 3rd ed.. New York:Wiley & Sons.

AUTHORS BIOGRAPHY

GERRIT K. JANSSENS (°1956) received degrees of M.Sc. in Engineering with Economy from the University of Antwerp (RUCA), Belgium, M.Sc. in Computer Science from the University of Ghent (RUG), Belgium, and Ph.D. from the Free University of Brussels (VUB), Belgium. After some years of work at General Motors Continental, Antwerp, he joined the University of Antwerp. Currently he is Professor of Operations Management and Logistics at Hasselt University (UHasselt) within the Faculty of Business Administration. He also holds the CPIM certificate of the American Production and Inventory Control Society (APICS). During the last twenty years he has been several times visiting faculty in universities in South-East Asia and Africa. His main research interests include the development and application of operations research models in production and distribution logistics.

KATRIEN M. RAMAEKERS (°1979) graduated as Master of Business Economics, option Technology at the Limburg University Centre in 2002. In October 2007, she obtained a Ph.D. in Business Administration at Hasselt University. Her main research interest is on the integration of simulation and optimisation, especially as a support for complex logistics decision-making and for decision support with limited information in supply and demand. She is a member of the research group Data Analysis and Modelling and the Transportation Research Institute of Hasselt University (IMOB).

WAREHOUSE MANAGEMENT: INVENTORY CONTROL POLICIES COMPARISON

Agostino G. Bruzzone[(a)], Duilio Curcio[(b)], Giovanni Mirabelli[(b)], Enrico Papoff[(b)]

[(a)]DIPTEM Department of Industrial, Production & Thermoenergetics Engineering and Mathematical Models
University of Genoa
Via Opera Pia, 15, 16145 Genova GE, Italy
agostino@itim.unige.it – URL www.simulationteam.com

[(b)]Modeling & Simulation Center - Laboratory of Enterprise Solutions (MSC – LES)
M&S Net Center at Department of Mechanical Engineering
University of Calabria
Via Pietro Bucci, Rende, 87036, ITALY
dcurcio@unical.it, g.mirabelli@unical.it, e.papoff@unical.it

ABSTRACT
This paper aims at investigating the effect of some parameters (i.e. the demand intensity, the demand variability and the lead time) on three different inventory control policies. To this end a parametric simulator is implemented in order to perform what-if analyses and scenarios investigation. The performance parameter chosen for the inventory management policies comparison is the unit inventory management cost. Finally, the analytical relationship between the unit inventory management cost and the input parameters is determined.

Keywords: Warehouse Management, Inventory Policies, Simulation, DOE, ANOVA

1. INTRODUCTION

According to literature warehouse management and internal logistics planning and control received, during the last years a great deal of attention. Van den Berg (1999) presents a literature survey on methods and techniques for planning and control of warehouse systems. Planning and control deal from one side with long-term goals, supply chain organization and warehouse design, from the other with inventory management and control policies with the aim of storing the correct quantity of products as well as determining the optimal time for placing purchase orders (considering production and transportation lead times).

Ashayeri and Gelders (1985) propose a review of different warehouse design models. Other research studies deal with the inventory management problem within warehouses: Hariga and Jackson (1996) present a review of inventory models for warehouse management while Van Oudheusden, Tzen, and Ko (1988), Frazelle, Hackman, Passy and Platzman (1994), Brynzér and Johansson (1995) investigate the advantages (in terms of productivity enhancement) due to a correct warehouse planning and control.

In addition, recent research studies regard data/information management in warehouse systems: Eben-Chaime and Pliskin (1997) investigate the effect of operations management tactics on performance measures of automatic warehousing systems with multiple machines.

The main goal of this paper is to compare three different inventory control policies in a warehouse located within an industrial plant devoted to produce different types hazelnuts based products. The inventory control policies are compared under different demand and lead time constraints using as performance measure the inventory management cost. As support tool the authors implemented as simulator that recreates stochastic scenarios based on different demand intensity, demand variability and lead time values.

The overall structure of the paper is as follows. Section 2 describes the hazelnut production process and the main warehouse technical characteristics. Section 3 describes the simulation model implementation. The inventory policies adopted are discussed in Section 4. Section 5 reports simulation results analysis and scenarios comparison. Finally, concluding remarks are given in Section 6.

2. THE WAREHOUSE SYSTEM

This research work aims at investigating and comparing three classical inventory control policies within a warehouse used to store hazelnuts in order to select the more efficient policy in terms of unit inventory management costs (*UICs*).

The warehouse has a rectangular shape with a surface of about 300 m^2 (the industrial plant surface is about 2000 m^2). Figure 1 shows the industrial plant layout (red arrows shows the material flow through the different work station).

The plant layout is subdivided in 8 different areas/department each one including different workstations carrying out the following main operations:

- ✓ pre-cleaning;
- ✓ drying;
- ✓ calibration;
- ✓ shelling;
- ✓ selection;
- ✓ roasting;
- ✓ graining;
- ✓ pasting;
- ✓ packaging.

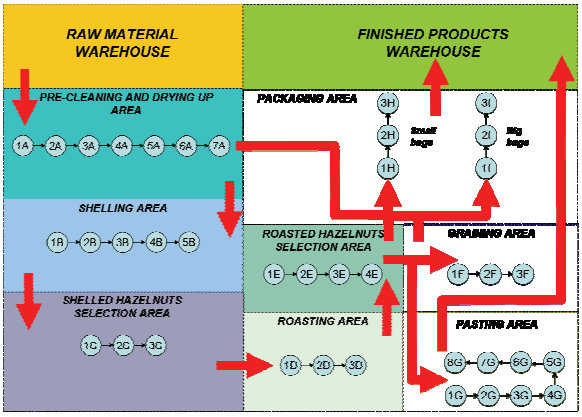


Figure 1: The Layout of the Manufacturing System

Figure 2 shows the flow chart of the production process including all the main operations.

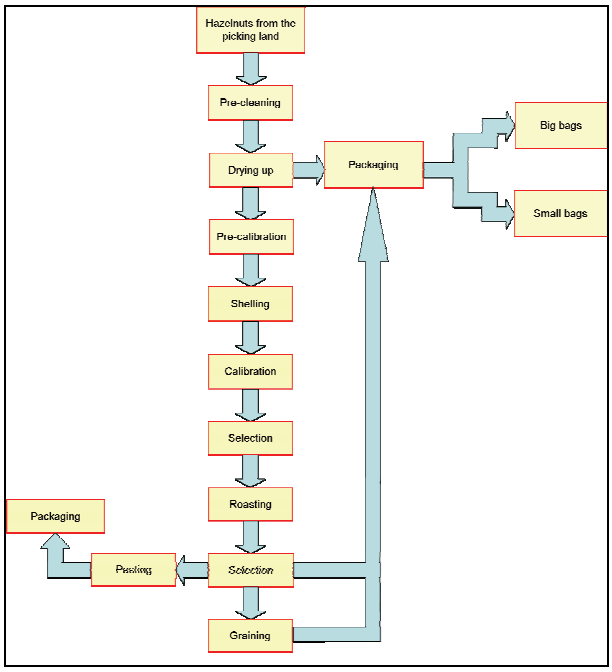


Figure 2: The production process flow chart

As before mentioned, the goal of this research work consists in monitoring the performance of three classical inventory management policies for understanding how some critical parameters (demand intensity, demand variability and lead time) affect the unit inventory management costs.

3. THE WAREHOUSE SIMULATION MODEL

Longo and Mirabelli (2008) and De Sensi et al. (2008) in their research work highlight the importance of simulation as an effective tool for inventory management problems and control policies comparison.

In fact, the use of simulation allows to explore and experiment possibilities for evaluating how the system under consideration reacts in correspondence of internal or external changes. As a consequence, one specific feature of the simulation model must be the flexibility for a complete scenarios design and analysis.

Bocca et al. (2008) implement a simulation model of a real warehouse highlighting the importance of building flexible simulation models while Cimino et al. (2008) analyze the performance of a real warehouse by monitoring its performance under different system configurations and by considering as performance measure the fill rate level.

The warehouse simulation model presented in this paper has been implemented by using Anylogic™ by *XJ Technologies* and it reproduces all the main warehouse operations. Warehouse main operations include trucks arrival and departure for items deliveries and internal materials handling operations. The simulation model implements different performance measures, including waiting times for suppliers' trucks and inventory costs.

4. THE INVENTORY CONTROL POLICIES

The objective of an inventory control policy is twofold:

- ✓ evaluation of the time for order emission;
- ✓ evaluation of the quantity to be ordered.

The focus of this paper is to test the effect of the demand intensity, the demand variability and the lead time on three different inventory control policies by using as support tool the simulation model before presented. The authors implement within the simulation model the following three inventory control policies:

- the reorder time-order quantity policy (*RTOQ*);
- the reorder point-order quantity policy (*RPOQ*);
- the (s, S) policy.

4.1. The analytical models of each inventory policy

Before introducing the analytical model of each inventory control policy, let us introduce the following notation:

- *s(t)*, the re-order level at time *t*;
- *S(t)*, the target level at time *t*;
- *SS(t)*, the safety stock level at time *t*;
- *DF(t)*, the demand forecast at time *t*;
- *OHI(t)*, the on-hand inventory at time *t*;

- *OQ(t)*, the quantity already on order at time *t*;
- *SQ(t)*, the quantity to be shipped at time *t*;
- *Q(t)*, the quantity to be ordered at time *t*;
- *L(t)*, the lead time;
- *DFL(t)*, the demand forecast over the lead time;
- *IP(t)*, the inventory position at time *t*.

The inventory position *IP(t)* is the on-hand inventory plus the quantity already on order minus the quantity to be shipped. In particular, it is defined as:

$$IP(t) = I(t) + QO(t) - QS(t) \quad (1)$$

4.1.1. The Reorder Time-Order Quantity (RTOQ) Policy

The RTOQ inventory control policy is based on a periodic check. If *T(t)* is the review period, the quantity to order is defined by *S(t)* minus *IP(t)*. The value of *T(t)* can be defined using the inverse formula usually used for evaluating the Economic Order Quantity (EOQ), refer to Silver et al. (1998).

$$Q(t) = S(t) - IP(t) = DFL(t) + SS - IP(t) \quad (2)$$

In this policy, *S(t)* represents the target level. This policy should be used when the inventory level is not automatically monitored, there are advantages related to scale economy, and orders are not regular.

4.1.2. The (s(t), S(t))Policy

This policy can be derived from the previous policy. According to literature, there are two main parameters:

- *s(t)*, the re-order level at time *t*;
- *S(t)*, the target level at time *t*.

IP(t) is checked periodically on the basis of the review period so two cases can occur:

- *IP(t)* is at or below the re-order point *s(t)*;
- *IP(t)* is above *s(t)*.

In the first case the quantity to be ordered (see equation 3) should be enough to raise the $IP_i(t)$ to $S_i(t)$ while in the second case no orders are placed.

$$Q(t) = S(t) - IP(t) \quad (3)$$

According to Silver et al. (1998), it is demonstrated that, under specific assumptions on demand pattern and cost factors, the *(s(t), S(t))* policy generates total costs lower than other inventory control policies.

4.1.3. The Reorder Point-Order Quantity (RPOQ) Policy

In this control policy, the inventory level is continuously checked according to production/demand requirements.

If *IP(t)* falls below the *s(t)*, a purchase order must be placed. The quantity to be ordered is defined using the Economic Order Quantity (*EOQ)* approach as reported in Silver et al. (1998).

$$s(t) = DFL(t) + SS(t) \quad (4)$$

$$Q(t) = EOQ(t) \quad (5)$$

Such policy should be adopted when inventory level is automatically monitored. There are no scale economies advantages and purchase orders can be regularly placed.

5. SCENARIOS DEFINITION AND DESIGN OF SIMULATION EXPERIMENTS

As before mentioned, the objective of this research work consists in evaluating how some input parameters affect the performance of the three inventory control policies before presented, in terms of unit inventory management costs (*UICs*). For each scenario the input parameters vary between specific values and conditions. In particular, the input parameters are:

- demand intensity (*DI*) which can assume three different conditions (low, medium, high);
- demand variability (*DV*) which can assume three different conditions (low, medium, high);
- lead time (*LT*) which can assume the following values be changed respectively in one day, three and five days.

The experiments planning is supported by the Design of Experiments (*DOE*) methodology; in particular, the Full Factorial Experimental Design is adopted.

Factors and levels for the design of experiments (*DOE*) are showed in Table 1.

Table 1: Factors and Levels of DOE

Factors	***Level 1***	***Level 2***	***Level 3***
DI	Low	Medium	High
DV	Low	Medium	High
LT	1	3	5

Each factor has three levels: Level 1 indicates the lowest value for the factor, Level 2 the medium value while Level 3 its greatest value.

To test all the possible factors levels combinations, the total number of the simulation runs is 3^3. Each simulation run has been replicated three times, so the total number of replications is 81 (27x3 = 81).

6. SIMULATION RESULTS ANALYSIS

This section presents the simulation experiments results. The behaviour of the inventory control policies has been

studied by using the Analysis of Variance (*ANOVA*) supported by some statistical charts. A similar approach is also proposed in Curcio and Longo (2009).

The ANOVA is used for understanding the analytical relationship between the input factors and the unitary inventory management costs by introducing an analytical relationship (the *meta-model* of the simulation model) between the performance measure and the factors being considered.

Let x_i ($x_1 = DI$, $x_2 = DV$, $x_3 = LT$) be the factors, equation 6 expresses the UIC as linear function of the x_i.

$$UIC = \beta_0 + \sum_{i=1}^{3} \beta_i x_i + \sum_{i=1}^{3} \sum_{j>i}^{3} \beta_{ij} x_i x_j + \\ + \sum_{i=1}^{3} \sum_{j>i}^{3} \sum_{h>j}^{3} \beta_{ijh} x_i x_j x_h + \varepsilon_{ijhkpn} \quad (6)$$

Equation 6 only considers the interaction terms up to order 3 plus an error term without considering the fourth and fifth order effects (usually such effects can be neglected). The main goal of the analysis of variance is twofold:

- to identify those factors which affect the UIC (sensitivity analysis);
- to evaluate the coefficients of equation 6 for defining the analytical relationship between the input and the output parameter.

Table 2 reports the results of the simulation experiments (the UIC) in correspondence all the factors levels combinations. The first three columns represents the experiments design matrix, while the last three columns reports the unit inventory management costs for each inventory control policy.

6.1. Simulation results for the RTOQ inventory control policy

In this section, simulation results for the first inventory policy (the reorder time-order quantity policy) are presented. The first step aims at detecting all those factors that influence the UIC. The confidence level adopted for estimating output data significance level is α=0.05 (according to the ANOVA theory the non-negligible effects are characterized by a *p-value* ≤ α where *p* is the probability to accept the negative hypothesis, i.e. the factor has no impact on the performance index). The results of the sensitivity analysis are reported in Table 3: the most significant effects are the first order effects because (their p-value is lower than the confidence level).

According to the ANOVA results, the second phase consists in introducing the analytical relationship between the input and the output parameters. Table 4 shows the coefficients for the input-output meta-model (equation 6).

Table 2: Simulation Results for the UIC

IN	*VAR*	*LT*	*RTOQ*	*(s,S)*	*RPOQ*
Low	Low	1	5,092	5,03	5,52
Medium	Low	1	5,06	5,008	5,181
High	Low	1	5,281	4,856	5,283
Low	Medium	1	6,06	5,62	6,08
Medium	Medium	1	5,633	5,673	5,727
High	Medium	1	5,965	5,655	5,923
Low	High	1	6,69	6,16	6,61
Medium	High	1	6,343	5,992	6,307
High	High	1	7,582	6,229	6,649
Low	Low	3	5,66	5,727	5,973
Medium	Low	3	5,4	5,287	5,66
High	Low	3	5,613	5,244	5,548
Low	Medium	3	6,59	6,551	6,701
Medium	Medium	3	6,74	5,997	6,405
High	Medium	3	6,692	6,267	6,398
Low	High	3	7,35	7,3	6,8
Medium	High	3	7,118	6,71	7,019
High	High	3	7,252	7,111	6,989
Low	Low	5	6,28	6,2	6,09
Medium	Low	5	5,79	5,653	5,954
High	Low	5	6,041	5,665	5,907
Low	Medium	5	7,14	6,92	7,02
Medium	Medium	5	6,493	6,61	7,007
High	Medium	5	7,182	6,959	6,594
Low	High	5	7,87	6,95	7,63
Medium	High	5	7,399	7,544	8,259
High	High	5	7,582	7,869	7,478

Table 3: Sensitivity Analysis Results – *RTOQ* policy

Source	*DF*	*Adj SS*	*Adj MS* (10^{-4})	*F*	*P*
DI	2	0,67	0,33	5,70	0,029
DV	2	12,49	6,24	106,1	0,000
LT	2	3,65	1,82	31,01	0,000
DI*DV	4	0,06	0,01	0,30	0,873
DI*LT	4	0,34	0,08	1,45	0,302
DV*LT	4	0,17	0,04	0,75	0,585
Error	8	0,47	0,05		
Total	26				

In particular, for each factor the first coefficient value represents the slope of the straight line between its low and medium levels while the value in the column (0) is the slope of the straight line for medium and high factor levels.

Table 4: ANOVA Coefficients – *RTOQ* policy

Term	*Coefficient*	
	(-1)	*(0)*
Constant	6,44067	
DI	0,08511	-0,22111
DV	-0,86100	0,05878
LT	-0,47333	0,04989

Equation 7 reports the input-output meta-model for the performance parameter (the unit inventory

management cost) when input parameters change between the medium and high levels:

$$Y_{ijkn} = 6{,}44067 - 0{,}22111 * DI + \\ + 0{,}05878 * DV + 0{,}04989 * LT \quad (7)$$

Figure 3 shows how the unitary inventory management cost changes in function of the three main effects: the unitary inventory management cost decreases when demand intensity goes from its low to medium values; from the other side it increases with the increase when the demand intensity changes from its medium to high levels.

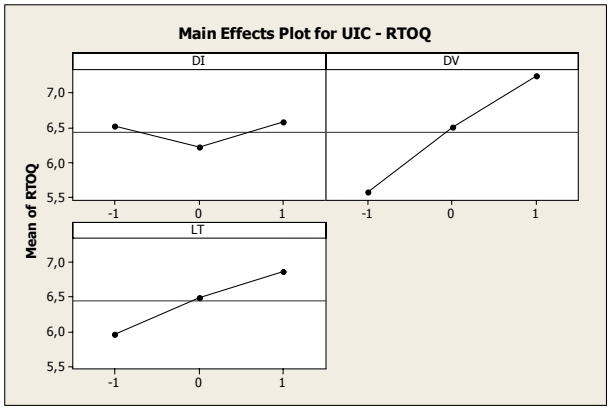


Figure 3: Unitary Inventory Management Costs versus Main Effects – *RTOQ* policy

6.2. Simulation results analysis for the (s,S) inventory control policy

The same analyses have been carried out for the (s,S) inventory control policy. From the sensitivity analysis the most significant effects are the first order effects (demand variability and lead time). The ANOVA coefficients for the (s,S) policy are reported in Table 5.

Table 5: ANOVA Coefficients – *(s,S)* policy

Term	***Coefficient***	
	(-1)	***(0)***
Constant	6,17730	
DV	-0,76952	0,07293
LT	-0,59696	0,06648

The input-output meta-model is reported in equation 8 (note that in this case low and medium levels parameters are reported).

$$Y_{ijkn} = 6{,}17730 - 0{,}76952 * DV + \\ - 0{,}59696 * LT \quad (8)$$

Figure 4 shows the Main Effects Plot for the (s,S) policy. In this case the unit inventory management costs increases when demand variability and lead time change from their low to their high levels.

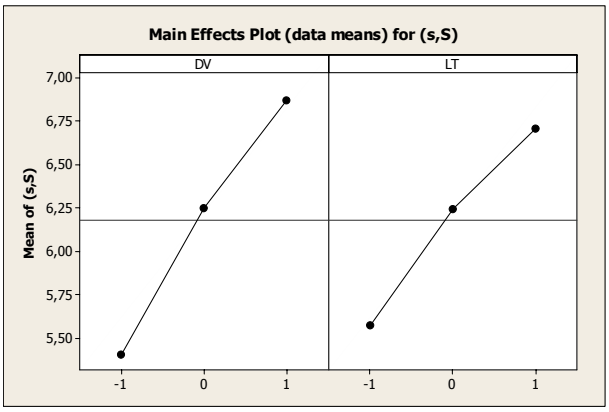


Figure 4: Unitary Inventory Management Costs versus Main Effects – *(s,S)* policy

6.3. Simulation results analysis for the RPOQ inventory policy

The third inventory control policy considered is the reorder point-order quantity policy.

Table 6 reports the sensitivity analysis results. Also in this case the most significant parameters are respectively the demand variability (*DV*) and the lead time (*LT*).

Table 6: Sensitivity Analysis Results – *RPOQ* policy

Source	***DF***	***Adj SS***	***Adj MS*** (10^{-4})	***F***	***P***
DI	2	0,1526	0,076	1,37	0,277
DV	2	8,868	4,434	79,63	0,000
LT	2	4,166	2,083	37,41	0,000
Error	20	1,113	0,055		
Total	26				

The ANOVA coefficients for the RPOQ policy are reported in Table 7 while equation 9 is the input-output meta-model (consider that parameters change between the low and medium levels).

Table 7: ANOVA Coefficients – *RPOQ* policy

Term	***Coefficient***	
	(-1)	***(0)***
Constant	6,39674	
DV	-0,71719	0,03159
LT	-0,47674	-0,00863

$$Y_{ijkn} = 6{,}39674 - 0{,}71719 * DV + \\ - 0{,}47674 * LT \quad (9)$$

Figure 5 shows the Main Effects Plot between demand variability and lead time effects.

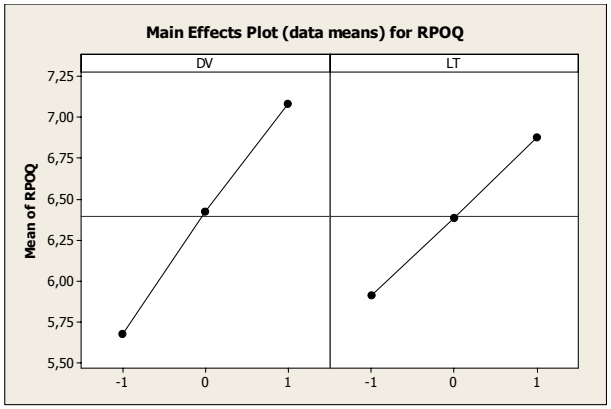


Figure 5: Unitary Inventory Management Costs versus Main Effects – *RPOQ* policy

For each scenario the validity of the analysis of variance results has been proved by using residuals analysis. To provide evidence on the results of the residuals analysis, the figure 6 shows the Normal Probability Plot of the Residuals for the RPOQ policy.

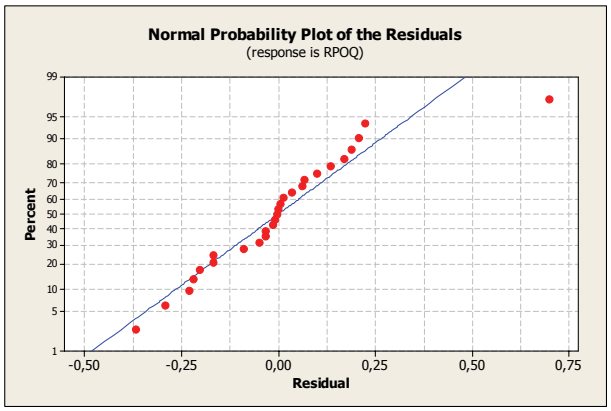


Figure 5: Normal Probability Plot of the Residuals – *RPOQ* policy

7. CONCLUSIONS

This paper aims at investigating the effect of some critical parameters (i.e. the demand intensity, the demand variability and the lead time) on three different inventory control policies within a real warehouse that support the production of hazelnuts based products. The performance measure taken into consideration is the unit inventory management cost. The simulation results have been studied by means of the Analysis of Variance both in terms of sensitivity analysis and in terms input-output analytical relationships. Such analytical relationships, one for each policy, express the unit inventory management costs as a function of the demand intensity, demand variability and lead time.

REFERENCES

Ashayeri, J., Gelders, L.F., 1985. Warehouse design optimization. *European Journal of Operation Research*, 21, 285–294.

Bocca, E., Curcio, D., Longo, F., Tremori, A., 2008. Warehouse and internal logistics management based on modeling & simulation. *Proceedings of the International Mediterranean Modelling Multiconference 2008,* pp. 41 – 48. September 17 – 19, Campora S.Giovanni (CS), Italy.

Bruzzone, A.G., Bocca, E., Longo, F., Massei, M., 2007. Training and Recruitment in Logistics Node Design by using Web Based Simulation. *International Journal of Internet Manufacturing and Services*, I(1), 2007, pp 32-50.

Bruzzone A.G., Massei M., Brandolini M., 2006. Simulation Based Analysis on Different Logistics Solutions for Fresh Food Supply Chain. *Proceedings of SCSC2006*, Calgary, Canada, July 30-August 3.

Bruzzone, A.G., Massei, M., Poggi, S., 2007. Simulation for Frozen Food Supply Chain. *Proceedings of Spring Simulation Multiconference*, Norfolk VA, March.

Brynzèr, H., Johansson, M.I., 1995. Design and performance of kitting and order picking systems. *International Journal of Production Economics*, 41, 115–125.

Cimino, A., Curcio, D., Mirabelli, G., Papoff, E., 2008. Warehouse inventory management based on fill rate analysis. *Proceedings of the International Mediterranean Modelling Multiconference 2008,* pp. 23 – 30. September 17 – 19, Campora S.Giovanni (CS), Italy.

Curcio, D., Longo, F., 2009. Inventory and internal logistics management as critical factors affecting supply chain performance. *International Journal of Simulation and Process Modelling*, 5 (2), 127–137.

De Sensi, G., Longo, F., Mirabelli, G., 2008. Inventory policies analysis under demand patterns and lead times constraints in a real supply chain. International Journal of Production Research, 46(24): 6997-7016.

Eben-Chaime, M., Pliskin, N., Sosna, D., 1997. Operations management of multiple machine automatic warehousing systems. *Production Economics*, 51, 83–98.

Eben-Chaime, M., Pliskin, N., Sosna, D., 2004. An integrated architecture for simulation. *Computer and Industrial Engineering*, 46, 159–170.

Frazelle, E.H., Hackman, S.T., Passy, U., Platzman, L.K., 1994. The forward-reverse problem. In: T. A. Ciriani and Leachman, R.C. (eds), *Optimization in Industry 2*. John Wiley, 43–61.

Hariga, M.A., Jackson, P.L., 1996. The warehouse scheduling problem: formulation and algorithms. *IIE Transactions*, 28 (2), 115–127.

Longo, F., Mirabelli, G., 2008. An Advanced supply chain management tool based on modeling and simulation. *Computer and Industrial Engineering*, 54/3: 570-588.

Silver, E.A., Pyke, D.F. and Peterson, R., 1998. Inventory Management and Production Planning

and Scheduling, 3rd ed., John Wiley & Sons, USA.
Van den Berg, J.P., 1999. A literature survey on planning and control of warehousing systems. *IIE Transactions*, 31 (8), 751-762.
Van Oudheusden, D.L., Tzen, Y., Ko, H., 1988. Improving storage and order picking in a person-on-board AS/R system. *Engineering Costs and Production Economics*, 13, 273-283.

AUTHORS BIOGRAPHY

AGOSTINO G. BRUZZONE received his Doctorship in Mechanical Engineering at the University of Genoa. Since 1991, he has taught 'Theories and Techniques of Automatic Control' and became member of the UniGE Industrial Simulation Work Group and then Full Professor. He applied M&S in harbour terminals, maritime trading and sailboat racing sectors. He is actively involved in the scientific community and served as Director of the McLeod Institute of Simulation Science, Associate Vice-President and Member of the Board of the Society for Modelling and Simulation International, President of the Liophant Simulation and Italian Point of Contact for the International Simulation Advisory Group and Sim-Serv.

DUILIO CURCIO was born in Vibo Valentia (Italy), on December the 15th, 1981. He took the degree in Mechanical Engineering from University of Calabria (2006). He is currently PhD student at the Mechanical Department of University of Calabria. His research activities include Modeling & Simulation and Inventory Management theory for production systems and Supply Chain design and management. He collaborates with MSC-LES at the Industrial Engineering Section of the University of Calabria to research projects for supporting Research and Development in SMEs.

GIOVANNI MIRABELLI was born in Rende (Italy), on January the 24th, 1963. He took the degree in Industrial Technology Engineering from University of Calabria. He is currently researcher at the Mechanical Department (Industrial Engineering Section) of University of Calabria. His research interests include Modeling & Simulation for workstations effective design, work measurement and human reliability.

ENRICO PAPOFF was born in Naples (Italy) on February the 03rd, 1948. He took the degree in Mechanical Engineering from University of Napoli Federico II, in 1973. He is currently Associate Professor at the Mechanical Department (Industrial Engineering Section) of the University of Calabria. His research interests regard project management, business plans and Modeling & Simulation applied to production and logistics systems.

MRP-DRP MODEL AS A BASE FOR NEGOTIATIONS IN TIMES OF RECESSION

Marija Bogataj[(a)(c)], Robert W. Grubbström FVR RI[(b)(c)] and Ludvik Bogataj[(a)]

[(a)] University of Ljubljana, Slovenia
[(b)] Linköping Institute of Technology, Sweden

[(c)] Mediterranean Institute for Advanced Studies, Šempeter pri Gorici, Slovenia

marija.bogataj@guest.arnes.si; robert@grubbstrom.com; Ludvik.bogataj@ef.uni-lj.si

ABSTRACT
In times of recession unemployment is increasing because of reduced demand, which influences the optimal production level. The impact of the production level on the cash flow, annuity stream and net present value generated by activities in a supply chain may be analysed in detail by employing MRP Theory.

In this paper we concentrate our attention on the question of (1) differences between planned production and realisation, which appear especially in the stochastic behavior of MRP-DRP systems, and (2) the oligopoly position of activity cells depending on the location and regional policies. Our extended MRP model enables us to derive consequences of these influences and interactions.

Keywords: MRP Theory, logistics, location, Net Present Value.

1. INTRODUCTION

The impact of the production level on the cash flow, annuity stream and net present value generated by activities in a supply chain may be analysed in detail by employing MRP Theory. MRP Theory has been developed in collaboration between Linköping Institute of Technology, Department of Production Economics, and other universities (in particular the University of Ljubljana) during the last two decades. The theory combines the use of Input-Output Analysis and Laplace transforms, enabling the development of a theoretical background for the dynamics of multi-level, multi-stage production-inventory systems together with their economic evaluation, in particular applying the Net Present Value principle (NPV) as the criterion function. In the late nineties, this theory has been extended from assembly to distribution (MRP-DRP) systems, and later also to include reverse logistics structures.

In this paper we concentrate our attention on the question of (1) differences between planned production and realisation, which appear especially in the stochastic behavior of MRP-DRP systems, and (2) the oligopoly position of activity cells depending on the location. A model is designed for predicting restructuring results and for the negotiation between regional authorities (where individual activity cells are located employing local human resources), and managers of the global supply chain. When in time of recessions the activity cells could be located at different regions, the regions differently participate to production level by their fiscal policies and level of subsidies mostly depends on the number of saved working places in their region, mostly proportional to the production level. But the policy of one region could influence the results of the total supply chain, also if the chain has activity cells allocated in several regions. Our extended MRP model enables us to derive consequences of these influences and interactions.

Among the elements that has a bearing on the suitability and viability of a community for capital investments in activity cells of the global supply chains are the following: (a) labour quality, availability and cost, (b) transportation cost and infrastructure, (c) labour union threats, (d) tax burden, (e) site and facility development and design, (f) development or acquisition cost and financing structure, (g) spatial planning restrictions and environmental legislation in region, (h) incentives, (i) access to infrastructure or other services, and other elements which influence profit and quality of life. Neoclassical theorists offer some insights into the spatial nature of industrial location. The more recent contributions of alternative location theorists explain the reasoning for such phenomena as decentralized production systems as a part of global supply chains. In our paper we wish to use some relevant pieces of neoclassical and modern theories to address the questions of industrial location, and decentralization applying MRP Theory when it is extended to supply chain models.

Central place theory, set forth by early location theorists like Weber, Christaller (Greenhut 1995), and Lösch (1954) is geometrically very simplified and based on the assumptions: (a) that population and resources are uniformly distributed over a homogeneous plane, (b) there exist free entry into the market, (c) the returns to scale are constant for all activities, (d) that perfect competition exists. In these models the production factors: labour and capital as well as transportation costs represent the keys to determine the optimal location. Firms locate in such a way that they maximize their profits. The models developed by these early location theorists fit reasonably well with observed reality. Lowry-like gravity models as an upgrade of these theories have been very well applied all over the world. A combination of this theory and the theory of land rent

developed by von Thünen (Beckmann 1997) and later embedded in the theory of urban growth by Alonso (1964) provides a step further to the results of modern location theories, emerging towards MRP-DRP models.

2. NETWORK MODELS

Production –distribution – reverse logistics network models provide us with an effective tool to model manufacturing and logistics activities of a supply chain.

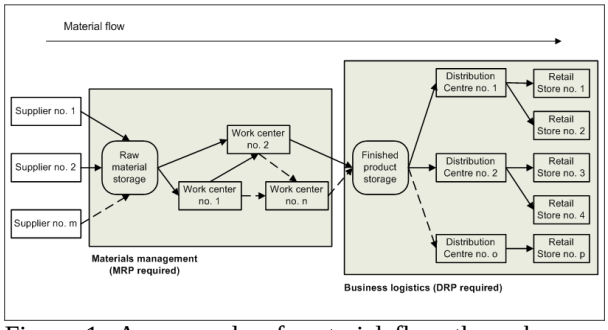


Figure 1: An example of material flow through many activity cells of a supply chain, (having different locations) divided into production and distribution segments. Between each pair of cells, there is transportation lead time.

In such a model, nodes represent vendors of raw materials or components, manufacturing and warehousing facilities in the production segment of a supply chain, and ports and distribution centres for semi-products and end items, warehouses and customers. We shall use the term "activity cell" for any of these. Arcs represent the infrastructure of flows between activity cells. The long-term performance goals for this production –distribution – reverse logistics system suggest strategic decision making regarding partners, playing different roles in the supply chain. The production –distribution – reverse logistics system design problem (PDRLSDP) involves the determination of the best configuration of the chain regarding location and capacity of the activity cells in the system. In such an attempt some activity cells have oligopoly positions and could be included in a network, which enable the flow of goods, or not. Here we will pay attention to the location of such activity cells and its impact on the net present value (NPV) and on a more general criterion function, when planning the flow could differ from its realisation.

The majority of analytical approaches for PDSDP utilizes discrete mixed integer programming models to represent facility design decision problems. Continuous models are successfully used in spatial economics and logistics, but there are only few papers that use continuous models for facility design (Daganzo 1998, Verter and Dincer 1995). Models of this type assume that customers are spread over a given market area and the optimal service region for each facility to be established is given. To develop a model for optimal strategic decisions on the location of activity cells, we shall start from MRP theory, developed by Grubbstrőm and others.

3. LOCATION OF ACTIVITY CELL IN MRP THEORY

Optimal decisions (i) where to produce, (ii) where to locate distribution centres (which of them could be included in a network) and (iii) where to organise reverse activities in integrated supply chain can be successfully discussed and evaluated in a transformed environment, where lead times and other time delays can be considered in linear form. An integrated approach is needed especially when we consider reverse logistics as an extended producer responsibility (Grubbström, Bogataj, and Bogataj 2007).

The site and capacity selection, as for instance the problems where it is best to locate a facility and what capacity is needed to achieve the most rapid response, can be discussed more easily in transformed environment (Aseltine 1958), using MRP (Orlicky 1975) and I-O analysis (Leontief 1951) in Laplace transformed space, as previously presented by Grubbström (1996, 1998, 2007) and as it has been discussed in many other papers of his Linköping research group.

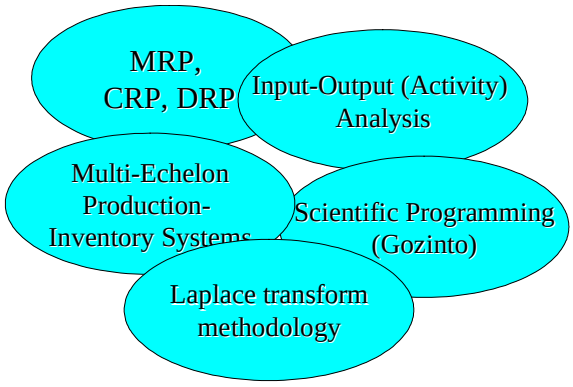


Figure 2: The ingredients of MRP Theory

MRP Theory has previously mainly dealt with *assembly structures* by which items produced downstream (on a higher level in the product structure) contain one or more sub-items on lower levels, but at each stage, the assembly activity produces only one type of output. This enables the *input matrix*, after enumerating all items suitably, to be organised as a triangular matrix, with non-zero elements only appearing below its main diagonal. The introduction of a diagonal *lead time matrix* capturing the advanced timing when required inputs are needed, enables compact expressions to be obtained, explaining the development of key variables such as available inventory and backlogs in the frequency domain. Central in this theory is the *generalised input* matrix showing when and how much the internal (dependent) demand amounts to for any production plan.

An extension of the production network has been made, including the distribution segment (Bogataj and Bogataj 2003) and later by Grubbström, Bogataj and Bogataj (2007) in close loop of product life cycle, including also reverse logistics. In these models transportation costs have been included in setup costs and transportation time lag was just extended production lead time, what is correct only if supply chain is linear or in radial form on the area described by Alonso's concentric models inherited from the Von Thünen model of agricultural land use, where all child nodes are equally remote from a certain activity cell. But this is very rear case. MRP theory was correct until transportation costs and transportation lead-time have been negligible and production lead time was the main reason for delays. Here we wish to improve MRP model to be able to use it for any supply chain evaluation, especially when we wish to study the impact of location and capacity of activity cells like ports are, on the certain objective function.

The labor cost and other costs of activities appear in every activity cell and depend on region where an activity cell is located. Together with transportation costs and costs of delay, which all depend on distances between two activity cells, (it means that it depends also on location of those cells), they influence NPV as the part of total of criterion function.

4. THE IMPROVED MRP THEORY FOR THE CASE WHEN REALISATION DO NOT FOLLOW THE PLANED ACTIVITIES IN A SUPPLY CHAIN

The line of research, now designated *MRP theory*, has attempted at developing a theoretical background for multi-level production-inventory systems, Material Requirements Planning (MRP) in a wide sense. Grubbström developed MRP theory on the basic methodologies of "Input-Output Analysis", (Leontief 1928) and *Laplace transform*. *Laplace transform* is a mathematical methodology dating back to the latter part of the 18th century and used for solving differential equations, for studying stability properties of dynamic systems, especially useful for evaluating the Net Present Value (NPV).

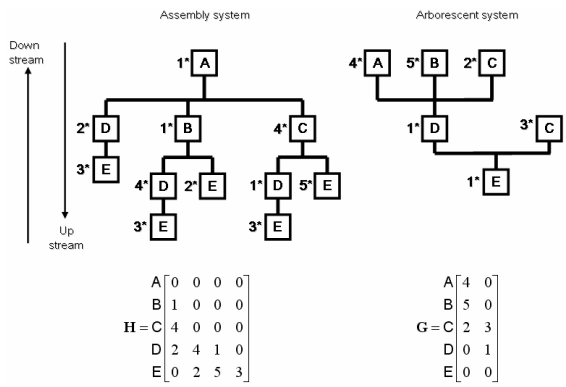


Figure 3. Examples of a pure assembly system and a pure arborescent system, in the form of product structures and their input and output matrices **H** and **G**, respectively (Grubbström, Bogataj, and Bogataj 2007).

Basic in MRP theory are the rectangular input and output matrices **H** and **G**, respectively, having the same dimension. Different rows correspond to different items (products) appearing in the system and different columns to different activities (processes). We let m denote the number of processes (columns) and n the number of item types (rows). If the jth process is run on activity level P_j, the volume of required inputs of item i is $h_{ij}P_j$ and the volume of produced (transformed) outputs of item k is $g_{kj}P_j$. The total of all inputs may then be collected into the column vector **HP**, and the total of all outputs into the column vector **GP**, from which the net production is determined as **(G - H)P**. In general **P** (and thereby net production) will be a time-varying vector-valued function of realized intensity of flows through the activity cells in a supply chain. In case when the plan of this intensity $\mathbf{P}_0$ is not equal to $\mathbf{P}$, we have to write the total of inputs by $\mathbf{HP}_0$, from which the net production is determined by $\mathbf{GP} - \mathbf{HP}_0$.

In MRP systems, lead times are essential ingredients. The lead time of a process is the time in advance of completion that the requirements are requested. If $P_j(t)$ is the volume (or rate) of item j planned to be completed at time t, then $h_{ij}P_j(t)$ of item i needs to be available for production (assembly) the lead time τ_j in advance of t, i.e. at time $(t-\tau_j)$. The volume $h_{ij}P_j$ of item *i*, previously having been part of *available inventory*, at time $(t-\tau_j)$ is reserved for the specific production $P_j(t)$ and then moved into *work-in-process* (*allocated component stock*, *allocations*). At time *t*, when this production is completed, the identity of the items type *i* disappear, and instead the newly produced items $g_{kj}P_j(t)$ appear. This approach has been developed for production systems, when transportation time did not influence lead time substantially. In case of transportation and production lead time τ_j should be split on two parts: production part of lead time τ_j^{pr} and transportation part τ_{ij}^{tr}. Therefore $h_{ij}P_j(t)$ of item *i* needs to be available for production (assembly) the lead time $\tau_{ij} = \tau_j^{pr} + \tau_{ij}^{tr}$ in advance of *t*, i.e. at time $(t - \tau_j^{pr} - \tau_{ij}^{tr})$.

In order to incorporate the lead times for assembly and arborescent processes in MRP systems without transportation time lags, Grubbström (1967, 1980, 1996, 1998, 2007) suggested transforming the relevant

time functions into Laplace transforms in the frequency domain.

5. WORKING IN FREQUENCY DOMAIN

When a time function repeats itself periodically, like it is often the case in the global supply chains, the length of a period being, says *T*, the transform of an infinite sequence of such time functions is:

$$\pounds\left\{\sum_{k=0}^{\infty} f(t-kT)\right\} = \frac{\tilde{f}(s)}{1-e^{-sT}} \qquad (1)$$

For sequence of discrete events within continuous processes, there is a need to introduce *Dirac's delta function* (*impulse function*) $\delta(t-t')$, having Laplace transform $\pounds\{\delta(t-t')\} = e^{-st'}$.

6. TRANSPORTATION LEAD TIME AND COST

Consider an *assembly* system, for which the components of process *j* need to be in place $\tau_{ij} = \tau_j^{pr} + \tau_{ij}^{tr}$ time units before completion according to the plan $\mathbf{P}_0$, applying the time translation theorem, the input requirements as transforms will be

1. in case of equal distances to the child nodes from i:

$$\mathbf{H}\begin{bmatrix} e^{s\tau_1^{pr}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{s\tau_m^{pr}} \end{bmatrix}\begin{bmatrix} e^{s\tau_1^{tr}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{s\tau_m^{tr}} \end{bmatrix}\tilde{\mathbf{P}}_0(s) =$$

$$= \mathbf{H}\tilde{\boldsymbol{\tau}}^{pr}(s)\tilde{\boldsymbol{\tau}}^{tr}(s)\tilde{\mathbf{P}}_0(s) = \tilde{\mathbf{H}}^{pt}(s)\tilde{\mathbf{P}}_0(s) \qquad (2)$$

where $\tilde{\boldsymbol{\tau}}^{pr}(s)$ and $\tilde{\boldsymbol{\tau}}^{tr}(s)$ are the so called production and transportation *lead time matrix*, and $\tilde{\mathbf{H}}^{pt}(s)$ the *generalised input matrix* capturing the volumes of requirements as well as their advanced timing. This vector describes in a compact way all component volumes that need to be in place for the production plan $\tilde{\mathbf{P}}_0(s)$ to be possible.

2. in case of different transportation time delays from the node i to its child nodes, we have to add to the components h_{ij} of matrix $\mathbf{H}$ corresponding delays in the form $h_{ij}e^{-s\tau_{ij}} = h_{ij}^{(\tau)}$ so that product of matrices

$$\begin{bmatrix} 0 & 0\,0 & \cdots & 0 \\ \ldots & h_{ij} & \ddots & \vdots \\ h_{m1} & h_{m2} & \cdots & 0 \end{bmatrix}\begin{bmatrix} e^{s\tau_1^{pr}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{s\tau_m^{pr}} \end{bmatrix}\begin{bmatrix} e^{s\tau_1^{tr}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{s\tau_m^{tr}} \end{bmatrix} \qquad (3)$$

is replaced by

$$\begin{bmatrix} 0 & 0 & 0\,0 & \cdots & 0 \\ \ldots & \ldots\ldots & h_{ij}e^{s\tau_{ij}^{tr}} & \ddots & \vdots \\ h_{m1}e^{s\tau_{m1}^{tr}} & & h_{m2}e^{s\tau_{m2}^{tr}} & \cdots & 0 \end{bmatrix}\begin{bmatrix} e^{s\tau_1^{pr}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{s\tau_m^{pr}} \end{bmatrix} = \tilde{\mathbf{H}}^{tr}\tilde{\boldsymbol{\tau}}^{pr}(s) \qquad (4)$$

and

$$\begin{bmatrix} 0 & 0 & 0\,0 & \cdots & 0 \\ \ldots & \ldots\ldots & h_{ij}e^{s\tau_{ij}^{tr}} & \ddots & \vdots \\ h_{m1}e^{s\tau_{m1}^{tr}} & & h_{m2}e^{s\tau_{m2}^{tr}} & \cdots & 0 \end{bmatrix}\begin{bmatrix} e^{s\tau_1^{pr}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{s\tau_m^{pr}} \end{bmatrix}\tilde{\mathbf{P}}_0 =$$

$$= \tilde{\mathbf{H}}^{tr}\tilde{\boldsymbol{\tau}}^{pr}(s)\tilde{\mathbf{P}}_0(s) = \tilde{\mathbf{H}}^{pt}(s)\tilde{\mathbf{P}}_0(s) \qquad (5)$$

where in $\tilde{\mathbf{H}}^{pt}(s)$ all kind of delays are included.

The similar split of lead time has to be made in arborescent system. Here the output of item *k* from running process *j* on the level $P_j(t)$ in terms of volume is $g_{kj}P_j(t)$.

Therefore, if $P_j(t)$ refers to the start of the process (initiation time) and the time of distribution (extraction) is Δ_{kj}, then the extracted items will appear at $t+\Delta_{kj}$,

$$\begin{bmatrix} e^{-s\Delta_1^{pr}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-s\Delta_n^{pr}} \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \vdots\ldots & & g_{kj}e^{-s\Delta_{kj}^{tr}} & & \vdots \\ 0 & & & \cdots & 0 \end{bmatrix}\tilde{\mathbf{P}}(s) =$$

$$= \tilde{\boldsymbol{\Delta}}^{pr}(s)\tilde{\mathbf{G}}^{tr}(s)\tilde{\mathbf{P}}(s) = \tilde{\mathbf{G}}^{pt}(s)\tilde{\mathbf{P}}(s) \qquad (6)$$

Here the diagonal matrix

$\tilde{\boldsymbol{\Delta}}^{pr}(s)$, $\tilde{\boldsymbol{\Delta}}^{tr}(s)$

are the lead time matrices of outputs and

$$\tilde{\mathbf{G}}^{pt}(s) = \tilde{\boldsymbol{\Delta}}^{pr}(s)\tilde{\boldsymbol{\Delta}}^{tr}(s)\mathbf{G}$$

is defined as the generalized output matrix.

The net production of such a system will conveniently be written:

$$\tilde{\mathbf{G}}^{pt}(s)\tilde{\mathbf{P}}(s)-\tilde{\mathbf{H}}^{pt}(s)\tilde{\mathbf{P}}_0(s)=$$

$$=\begin{bmatrix} e^{-s\Delta_1^{pr}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-s\Delta_n^{pr}} \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \vdots\ldots\ldots & & & g_{kj}e^{-s\Delta_{kj}^{tr}} & \vdots \\ 0 & & & \cdots & 0 \end{bmatrix}\tilde{\mathbf{P}}(s)-$$

$$-\begin{bmatrix} 0 & 0 & 0\,0 & \cdots & 0 \\ \ldots\ldots & \ldots\ldots\ldots & h_{ij}e^{s\tau_{ij}^{tr}} & \ddots & \vdots \\ h_{m1}e^{s\tau_{m1}^{tr}} & h_{m2}e^{s\tau_{m2}^{tr}} & & \cdots & 0 \end{bmatrix}\begin{bmatrix} e^{s\tau_1^{pr}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{s\tau_m^{pr}} \end{bmatrix}\tilde{\mathbf{P}}_0(s) \tag{7}$$

Given a plan $\tilde{\mathbf{P}}_0(s)$, available inventory $\tilde{\mathbf{R}}(s)$ will develop according to:

$$\tilde{\mathbf{R}}(s)=\frac{\mathbf{R}(0)+\tilde{\mathbf{G}}^{pt}(s)\tilde{\mathbf{P}}(s)-\tilde{\mathbf{H}}^{pt}(s)\tilde{\mathbf{P}}_0(s)-\tilde{\mathbf{F}}(s)}{s} \tag{8}$$

where $\mathbf{R}(0)$ collect initial available inventory levels. The term $\tilde{\mathbf{G}}^{pt}(s)\tilde{\mathbf{P}}(s)$ is the inflow of purchasing, production, extraction, distribution etc. into available inventory, the term $\tilde{\mathbf{H}}^{pt}(s)\tilde{\mathbf{P}}_0(s)$ is the required outflow representing needs generated by all processes (internal demand, dependent demand), where in both cases location influence technology matrices and the term $\tilde{\mathbf{F}}(s)$ represents deliveries (exports) from the system to the users on their existing locations.

This is an instance of the fundamental equations of MRP theory in case of extension to MRP-DRP case, where transportation delays influence behavior of supply chain and plan differ realization. In order for the plan $\tilde{\mathbf{P}}_0(s)$ to be feasible, we must always have fulfilled $\pounds^{-1}\left\{\tilde{\mathbf{R}}(s)\right\}\geq\mathbf{0}$. This is the *available inventory constraint*. If also capacity requirements are considered, a corresponding constraint for available capacities may be formulated like in some previous papers of Grubbström et al (Segerstedt 1996).

In the case that we wish to model *cyclical processes*, repeating themselves in constant time intervals T_j, j = 1, 2, … , m, we may write the plan $\tilde{\mathbf{P}}_0(s)$ in the following way, using two new diagonal matrices $\tilde{\mathbf{t}}(s)$ and $\tilde{\mathbf{T}}(s)$,

$$\tilde{\mathbf{P}}_0(s)=\tilde{\mathbf{t}}(s)\tilde{\mathbf{T}}(s)\hat{\mathbf{P}}_0$$

$$\tilde{\mathbf{t}}(s)=\begin{bmatrix} e^{-st_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-st_m} \end{bmatrix} \tag{9}$$

$$\tilde{\mathbf{T}}(s)=\begin{bmatrix} \left(1-e^{-sT_1}\right)^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \left(1-e^{-sT_m}\right)^{-1} \end{bmatrix}$$

where $\hat{\mathbf{P}}_0$ is a vector of constants, for instance describing the total amounts planned to be produced in (or delivered by) each process during one of the periods T_j, j = 1, 2, … , m, and where t_j, j = 1, 2, … , m, are the points in time when the first of each respective cycle starts. These latter times may be necessary in order for the system to have items on lower levels available as inputs on higher levels. And realization $\tilde{\mathbf{P}}(s)$ is going to be close to the plan as much as possible according to the criterion function.

Here a series expansion of $\tilde{\mathbf{T}}(s)$ leads to

$$\tilde{\mathbf{T}}(s)\hat{\mathbf{P}}=\begin{bmatrix} \dfrac{\hat{P}_1}{1-e^{-sT_1}} \\ \vdots \\ \dfrac{\hat{P}_m}{1-e^{-sT_1}} \end{bmatrix}=\frac{1}{s}\begin{bmatrix} \dfrac{\hat{P}_1}{T_1} \\ \vdots \\ \dfrac{\hat{P}_m}{T_m} \end{bmatrix}+\frac{1}{2}\begin{bmatrix} \dfrac{\hat{P}_1}{T_1^2} \\ \vdots \\ \dfrac{\hat{P}_m}{T_m^2} \end{bmatrix}+\mathbf{O}(s) \tag{10}$$

where O(s) is a vector vanishing at least with the speed of s.

7. CRITERION FUNCTIONS OF GLOBAL SUPPLY CHAIN AND SUBSIDIES IN TIME OF RECESSION

We now turn our attention to economic relationships. Activity cell j is assumed to produce item with value per item equal p_j. We collect these values per item into a price vector p being a row vector:

$$\mathbf{p}=\left[p_1, p_2, \ldots, p_n\right] \tag{11}$$

which could have different values at different locations. Because of state interventions in time of recession this values can be disturbed by

$$\Delta\mathbf{p} = \left[\Delta p_1, \Delta p_2, ..., \Delta p_n\right] \tag{12}$$

the values achieved on the market can be reduced by fees and taxes and increased by different kind of subsidies.

Although prices are normally positive, representing positive values to the holder of the asset, there may be instances when negative prices may be used. For instance, this is the case for waste items, which need to be disposed of at an expense, and having such items represents a negative value to the holder, which could differ from location to location. Prices p_i in general differ when changing location. Between location of activity cell i and following activity cell j it can differ for $b_{ij} \cdot \tau_{ij}^{tr}$ where b_{ij} presents transportation costs per item i per time unit, which we collect into a transportation price matrix per unit of product at j, $\mathbf{\Pi}_{G,}$ $\mathbf{\Pi}_{H,}$ so that the sum of transportation costs between activity cells TrC is equal to:

$$TrC = \mathbf{E}^T \left(\tilde{\mathbf{\Pi}}_G(s)\tilde{\mathbf{P}}(s) + \tilde{\mathbf{\Pi}}_H(s)\tilde{\mathbf{P}}_0(s)\right)$$

$$\tilde{\mathbf{\Pi}}_G(s) = \begin{bmatrix} e^{-s\Delta_1^{pr}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-s\Delta_n^{pr}} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & g_{kj}d_{kj}\tau_{kj} & \vdots \\ & & 0 & \cdots & 0 \end{bmatrix}$$

$$\tilde{\mathbf{\Pi}}_H(s) = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ .... & & .....h_{ij}d_{ij}\tau_{ij} & & \ddots & \vdots \\ h_{m1}d_{m1}\tau_{m1}... & & & h_{mj}d_{mj}\tau_{mj} & \cdots & 0 \end{bmatrix} \begin{bmatrix} e^{s\tau_1^{pr}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{s\tau_m^{pr}} \end{bmatrix} \tag{13}$$

and $\mathbf{E}^T = (\sum_{j=1}^{m} \mathbf{e}_j)^T$ is an *m*-dimensional row vector of unit values.

When the processes take place in discrete batches at times t'_{jk}, $k = 1, 2, \ldots$, for process *j*, we may also locate fixed costs (setup costs) at these times. Such setup times of process *j* are conveniently collected into a sequence of Dirac impulses $\delta\left(t - t'_{jk}\right)$ and the transform of such a sequence is $\tilde{v}_j(s) = \sum_k £\left\{\delta(t - t'_{jk})\right\} = \sum_k e^{-st'_{jk}}$. If there is a fixed out-payment attached to each such batch, say K_j, the NPV of these payments together will amount to $K_j\tilde{v}_j(\rho) = K_j \sum_k e^{-\rho t'_{jk}}$. In particular, if batches are completed in an infinite sequence and all batches are temporally located between constant time intervals of length T_j, the NPV of the setup (ordering) costs will be $\mathrm{NPV} = K_j e^{-\rho t_j} / \left(1 - e^{-\rho T_j}\right)$, assuming the first batch being timed at $t'_{j1} = t_j$. In our standard treatment, ordering costs are collected into the row vector $\mathbf{K} = \left[K_1, K_2, ..., K_m\right]$.

If all processes take place in discrete batches, letting $\tilde{\mathbf{v}}(s) = \begin{bmatrix} \tilde{v}_1(s) \\ \vdots \\ \tilde{v}_m(s) \end{bmatrix}$ denote the *m*-dimensional vector of all setup events, the NPV of all fixed ordering costs will be

$$\mathrm{NPV}_{\text{ordering}} = \\ = -\mathbf{K}\tilde{\mathbf{v}}(\rho) = -\mathbf{K}\begin{bmatrix} \tilde{v}_1(\rho) \\ \vdots \\ \tilde{v}_m(\rho) \end{bmatrix} = -\sum_{j=1}^{m} K_j \tilde{v}_j(\rho), \tag{14}$$

When all item flows in the system together with the parameters contained in $\mathbf{p}$, $\mathbf{\Pi}_{G,}$ $\mathbf{\Pi}_{H,}$ and $\mathbf{K}$ also accurately describe the relevant cash flow, the overall NPV may be written:

$$\mathrm{NPV} = \mathbf{p}\left(\tilde{\mathbf{G}}^{pt}(\rho)\tilde{\mathbf{P}}(\rho) - \tilde{\mathbf{H}}^{pt}(\rho)\tilde{\mathbf{P}}_0(\rho)\right) - \\ - \mathbf{E}^T\left(\tilde{\mathbf{\Pi}}_G(\rho)\tilde{\mathbf{P}}(\rho) + \tilde{\mathbf{\Pi}}_H(\rho)\tilde{\mathbf{P}}_0(\rho)\right) - \mathbf{K}\tilde{\mathbf{v}}(\rho)\mathbf{E}$$

$$\mathrm{NPV} = (\mathbf{p}\tilde{\mathbf{G}}^{pt}(\rho) - \mathbf{E}^T\tilde{\mathbf{\Pi}}_G(\rho))\tilde{\mathbf{P}}(\rho) - \\ -(\mathbf{p}\tilde{\mathbf{H}}^{pt}(\rho) + \mathbf{E}^T\tilde{\mathbf{\Pi}}_H(\rho))\tilde{\mathbf{P}}_0(\rho) - \mathbf{K}\tilde{\mathbf{v}}(\rho)\mathbf{E} \tag{15}$$

If we wish to control a supply chain system so, that the realization of the flow in the system is close to the planed one as much as possible, where each activity cell has individual importance when approaching to planned production or distribution intensity, we have to write the criterion:

$$\text{Min } ((\tilde{\mathbf{P}}(\rho) - \tilde{\mathbf{P}}_0(\rho))^T \boldsymbol{\theta}(\tilde{\mathbf{P}}(\rho) - \tilde{\mathbf{P}}_0(\rho))) \tag{16}$$

and a diagonal matrix $\boldsymbol{\theta}$ gives the importance to each production or distribution activity cell to approach to planned intensity.

In time of recession different state plans are trying to keep the activity on the level as it was before the recession by subsidies, reducing liquidity problems in time of recession. Let us assume that there was realization equal to plan before recession and now the local authorities where the activity cell j of supply chain is located wish to push the production to planned

one, to keep the human resources in region, where activity cell is located, close to the previous employment. At the same time the supply chain managers try to relocate activities to keep the NPV at reducing demand high as much as possible. Their NPV is described by (15). When demand is falling also production has to be reduced and therefore employment would be much lower if we are following only criterion (15). In the negotiation procedure between local authorities at different regions where activity cells are located, giving the subsidies to keep the production high as much as possible for avoiding unemployment in the region, and supply chain managers, who are still following equation (15) the main goal is to determine the ponderous $\boldsymbol{\Psi}$ and $\boldsymbol{\theta}$ when maximizing NPV. Therefore the ponderous should be find for the following :

$$\begin{aligned}&\max(((\mathbf{p}+\Delta\mathbf{p})\tilde{\mathbf{G}}^{pt}(\rho)-\mathbf{E}^T\tilde{\boldsymbol{\Pi}}_G(\rho))\boldsymbol{\Psi}\tilde{\mathbf{P}}(\rho)-\\&-((\mathbf{p}+\Delta\mathbf{p})\tilde{\mathbf{H}}^{pt}(\rho)+\mathbf{E}^T\tilde{\boldsymbol{\Pi}}_H(\rho))\boldsymbol{\Psi}\tilde{\mathbf{P}}_0(\rho)\\&-\mathbf{K}\tilde{\mathbf{v}}(\rho)\mathbf{E}-(\tilde{\mathbf{P}}(\rho)-\tilde{\mathbf{P}}_0(\rho))^T\boldsymbol{\theta}(\tilde{\mathbf{P}}(\rho)-\tilde{\mathbf{P}}_0(\rho)))\end{aligned} \tag{17}$$

What becomes the game between regional policies and global supply chains, especially needed to be consider in time of recession to determine acceptable production level, which is reducing unemployment. The approach given in Bogataj and Bogataj (2001) about supply chain coordination in spatial games can be used here in more general sense.

8. CONCLUSIONS

In this paper we have studied some aspects of differences between pairs of planned activities and realization in a global supply chain. In time of recessions this kind of difference appear in every region and mostly in all global supply chains.

To describe the approach to negotiations among regional authorities and managers in global supply chain extended MRP model, previously developed by Grubbström and later extended by distribution and reverse logistics component in a compact form, presented by Grubbström, Bogataj, and Bogataj (2007), has been suggested.

We have used the results of Bogataj, Grubbström, and Bogataj (2008) demonstrating the basic differences between MRP Theory describing the flows “Under the same roof” and model of global supply chain. At global supply chain lead time appear not only because of production and logistic activities. They also influence strongly NPV of supply chain activities because of transportation time delays. Therefore we have split lead time to production and transportation part, which appear on different ways in the model.

Negotiations among regional authorities and global supply chain managers about subventions to keep the human resources in a supply chain and therefore the production on the intensity level as planned, could base on the criterion function (17), where $\Delta\mathbf{p}, \boldsymbol{\Psi}, \tilde{\mathbf{P}}(\rho)$ and $\boldsymbol{\theta}$ are subject of negotiation. MRP Theory approach with MRP-DRP extension makes a supply chain more visible and more controllable. While in market economy $\boldsymbol{\Psi}$ is supposed to play the most important role, in eastern economies of last century $\boldsymbol{\theta}$ had been over-weighting $\boldsymbol{\Psi}$. The influence of this extreme policies as well as any combination of it can be well describer and studied using MRP-DRP approach presented above.

REFERENCES

Alonso, W., 1964. *Location & Land Use: Toward a General Theory of Land Rent*, Thesis (PhD Harvard.

Aseltine, J.A., 1958. *Transform Method in Linear System Analysis*, McGraw-Hill, New York, N.Y.

Beckmann, M.J., 1999. *Lectures on Location Theory*, Springer, Berlin.

Bogataj, D. and Bogataj, M., 2007. Measuring the supply chain risk and vulnerability in frequency space. *Int. j. prod. econ.*. [Print ed.], vol. 108, Issues 1-2, pp. 291-301.

Bogataj, M. and Bogataj, L., 2003. On the compact presentation of the lead-time perturbations in distribution networks, *Int.j.Prod. Econ*, Volume 88, Issue 2, pp 145 – 155.

Bogataj, M., 1996. Inventories in spatial models. *Int. j. prod. econ.*. vol. 45, pp. 337-342.

Bogataj, M. and Bogataj, L., 2001. Supply chain coordination in spatial games. *Int. j. prod. Econ.*, vol. 71, no. 1-3, pp. 277-286.

Bogataj, M., 1999. Inventory allocation and customer travelling problem in spatial duopoly. *Int.j. Prod. Econ.*59, pp.271-279.

Bogataj, M., Grubbström, R.W., Bogataj, L (2008), *Location of activity cells in distribution and reverse logistics*, ISIR Symposium 2008, Budapest, to appear: *Int.j. Prod. Econ.*

Daganzo, C.F., 1991. *Logistics Systems Analysis*, Springer, Berlin.

Greenhut, M. L., 1995. *Location Economics: Theoretical Underpinnings and Applications*. Edward Elgar Publ.Co .

Grubbström, R.W. 1967, On the Application of the Laplace Transform to Certain Economic Problems, *Management Science*, Vol.13, No.7, pp.558-567.

Grubbström, R.W., 1998. A Net Present Value Approach to Safety Stocks in Planned Production, *Int.j. Prod. Econ.* Vol. 56-57, pp. 213-229.

Grubbström, R.W., 1996. Material Requirements Planning and Manufacturing Resources Planning, in Warner, M., (Ed.), *International Encyclopedia of Business and Management*, Routladge, London,

Grubbström, R.W., 1980. A principle for determining the correct capital costs of work-in-progress and inventory, *International Journal of Production Research*, 18, 259-271.

Grubbström, R.W., Tang, O., 2000. An Overview of Input-Output Analysis Applied to Production-Inventory Systems, *Economic Systems Review*, 12, pp.3-25.

Grubbström, R.W., Bogataj, L., (Eds), 1998. *Input-Output Analysis and Laplace Transforms in Material Requirements Planning*. Storlien 1997. FPP Portorož.

Grubbström, R.W., 2007. Transform Methodology Applied to Some Inventory Problems, *Zeitschrift für Betriebswirtschaft*, 77(3), pp. 297-324.

Grubbström, R.W., Bogataj, M., Bogataj, L, *A compact representation of distribution and reverse logistics in the value chain*, WP, Ljubljana: Faculty of Economics, KMOR, 2007.

Leontief, W.W., 1951. *The Structure of the American Economy*, Oxford University Press, London.

Lösch, A., 1954. *The Economics of Location*, Yale University Press, New Haven, CT,

Orlicky, J.A., 1975. *Material Requirements Planning*, McGraw-Hill, New York, N.Y.

Puu, T., 1997**.** *Mathematical Location and Land Use Theory*, Springer, Heidelberg.

Segerstedt, A., 1996. A Capacity- Constrained Multi-level Inventory and Production Control Problem, *International Journal of Production Economics*, Vol. 45, pp.449-461.

Verter, V., Dincer, M.C., 1995. Facility location and capacity acquisition: An integrated approach, *Naval Research Logistics* 42, pp. 1141-1160.

AUTHORS BIOGRAPHIES

Marija Bogataj received Math. B.Sc. Degree (1974), MSc Degree (Operations Research, 1978) and Dr Engineering Science Degree from University of Ljubljana in 1983. She started as a lecturer at University of Maribor. She is Professor of Operations Research and Statistics at University of Ljubljana. During 33 years she was the member of many Scientific Committees in international, Slovenian and Croatian Journals and Proceedings. In the last 20 years she was visiting professor at many EU Universities. She is author, co-author, or editor/co-editor, of 19 books, author or co-author of 164 articles, essays and working papers, of which 54 have been published in scientific journals. She was leading researcher or participated substantially at 49 research projects. She is founding member of MEDIFAS. (see: www.medifas.net).

Robert W. Grubbstrom received Lic. Techn. Degree from Royal Institute of Technology, Stockholm in 1969 and Dr Econ. Degree from Gothenburg School of Economics and Business Administration in 1973, he is Knight 1 Cl. of the Order of the White Rose (Finland), Honorary Doctor of Lappeenranta University of Technology (Finland), Corvinus University (Budapest, Hungary) and Leopold-Franzens-Universät (Innsbruck, Austria), Editor-in-Chief of the International Journal of Production Economics (Elsevier), since 1979. Author, co-author, or editor/co-editor, of some 15 books, author or co-author of some 300 articles, essays and working papers, of which more than one hundred have been published in highly ranking scientific journals. He is founding member of MEDIFAS.

Ludvik Bogataj received MSc. Degrees: Operations Research, at Uni. Ljubljana, 1975, and Mathematics, at Uni. Zagreb, 1985 and Doctor of Science Degrees from University of Ljubljana in 1979: Information – Management Science and in 1989: Mathematics. He started his teaching at University of Maribor. He is Professor of Mathematics in Economics at University of Ljubljana and Head of KMOR.. He was member of many Scientific Committees in international and Slovenian Journals and Proceedings. He is author, co-author, or editor/co-editor, of 13 books, author or co-author of 134 articles, essays and working papers, of which 32 have been published in scientific journals. He was leading researcher or he has participated substantially at 24 research projects.

A SIMULATION MODEL FOR LIGHT RAIL TRANSPORTATION SYSTEM

Filiz Dumbek[(a)], Dilay Celebi[(b)]

[(a)(b)]Department of Management Engineering, Istanbul Technical University, Macka 34367, Istanbul, Turkey

[(a)]dumbek@itu.edu.tr, [(b)]celebid@itu.edu.tr

ABSTRACT
Public transportation has a strategic importance especially in large cities with respect to increasing population. Subway and light rail systems are efficient passenger transportation systems in large cities due to their large capacities and they are usually faster than other transportation tools. Increasing demand for these transportation systems has brought up the use of complex models to increase efficiency of the system by reducing time delays and passenger waiting times. In this study, a simulation model has been developed to determine best operation strategy in Istanbul Light Rail Transportation System (LRTS). Currently, Istanbul LRTS network has 18 stations and a total length of 19,6 km. This network provides transportation service for 240.000 passengers every day and operates according to predefined timetables which differ due to the operating season and day of the week. The system was developed as a discrete event simulation model and Arena Simulation Software was used.

Keywords: simulation, light rail transportation, passenger service quality

1. INTRODUCTION

Efficient management of rail systems that presents an effective solution for the big cities which are getting crowded day by day is very important in terms of operational efficiency and passenger service quality.

Public rail transportation planning process has been divided into strategic, tactical and operational levels (Ghoseiri, Szidarovszky, and Asgharpour 2004). Demand analysis and line planning are considered in the strategic planning level. Rolling stock planning, crew scheduling, crew rostering are considered in the tactical planning level. Train scheduling that includes departure times of the trains for specified routings take place between tactical and strategic planning levels. Therefore, scheduling process of the trains has significant importance to provide operational efficiency and passenger service quality.

Hooghiemstra and Teunisse (1998) indicate that travel times and halting times, plus possibly a maximum travel time between the end stations, constrain the timetable of an individual train.

While timetables are being constructed, service quality of the system should be taken into consideration. Vansteenwegen and Oudheusden (2007) indicate that passenger waiting time is a critical factor to evaluate passenger service quality.

Yalcinkaya and Bayhan (2008) developed a model optimize average travel time for a metro line by using simulation and response surface methodology. Headways which represent time interval between two consecutive trains were calculated to find optimum rate of carriage fullness and average travel time (Yalcinkaya and Bayhan 2008).

In this study, a simulation model was developed to determine optimum process strategy in Istanbul Light Rail Transporation System (LRTS). This network was founded in 1989 and been expanded over years. Currently, Istanbul LRTS network has 18 stations and total length of 19,6 km. This system has different time tables on weekdays, Saturdays, and Sundays.

2. SYSTEM STRUCTURE

Istanbul LRTS was founded in 1989 and currently has 18 stations. Istanbul Ulasim Corporation has 92 wagons and 80 of these wagons are being used. 17 vehicles operate in network and network has 240.000 passenger transport capacity per day. Figure 1 shows the rail network of Istanbul city. Red colored network indicates Istanbul light rail transportation system.

Figure 1: Istanbul Rail Network

In developed model seven stations (Aksaray, Emniyet, Ulubatli, Bayrampasa, Sagmalcilar,

Kartaltepe, Otogar) of this network were taken into consideration. System operates between 5:00 am and 1:00 am.

Simulation model was developed to reflect system structure according to these predefined constraints.

1. Number of vehicles
2. Minimum signalization time between train arrivals for one direction
3. Wagon capacity
4. Minimum train dwell times at stations
5. Minimum transfer time between two stations
6. Vehicle velocity

Currently 17 trains operate in LRTS. One of these trains have used for ring services between two stations and this movement of trains were not included in model. Therefore, 16 trains will operate between specified stations.

Signalization system that is used in the system allows train arrivals with minimum 2 minutes time intervals.

One wagon has 257 passengers carriage capacity. A vehicle operates with four wagons and total passenger carriage capacity is 1028.

Trains stop in stations for passenger alighting and boarding activities at the stations. Dwell times of the trains change according to the station, predefined time intervals during a day and train trip directions.

Minimum transfer times depend on train velocity and the distance between stations. Train velocities can change between stations but average velocities were used. Velocities of the trains did not change during the model. Therefore transfer times between stations have been defined constant according to the current data taken from Istanbul Ulasim Corporation.

3. METHOD

Rail network is a dynamic system that changes during time and variables in this system change at discrete times. Therefore, model was developed as discrete simulation model.

By taking system structure into consideration related data was collected as indicated in section 3.1.

Arena simulation software package was used to developed model and this model based on basic assumptions which are indicated in section 3.2. Section 3.3. explains the used Arena components and section 3.4. explains the model events which are occuring in the system.

3.1. Data Collection

Passenger arrival patterns were defined according to the collected data at stations within 15 minutes intervals between January 2008 and November 2008. Approximately half million records were loaded to Oracle database. Records that represent weekday arrivals to the stations were taken into consideration and average number of the passenger arrivals within 60 minutes was obtained.

Istanbul Ulasim Corporation conducted a survey to obtain percentages of alighting passenger at the stations. Table 1 shows the obtained results from this survey. For example, it was shown that %4,9 of passengers who starts their trip from Aksaray station alight at Emniyet station. This information was used to update the number of the passengers in the train and train's current capacity was calculated after this update.

Table 1: Percentages of Alighting Passengers at the Stations

Trip Ending Station	Trip Starting Station																	
	Aksaray	Emniyet	Ulubatli	Bayrampasa	Sagmalcilar	Kartaltepe	Otogar	Esenler	Terazidere	Davutpasa	Merter	Zeytinburnu	Bakirkoy	Bahcelievler	Atakoy	Yenibosna	DTM	Havalimani
Aksaray	0	6.1	4.5	16.1	35.1	27	20.6	28.8	22.7	19	21.4	4.5	9.7	0	14.5	10.8	12.5	15.6
Emniyet	4.9	0	4.5	3.2	8.1	5.4	4.4	7.7	9.1	0	7.1	1.5	3.2	4.8	4.3	3.6	0	0
Ulubatli	1	0	0	0	2.7	2.7	2.9	1.9	0	4.8	0	1.5	3.2	4.8	2.9	4.8	0	0
Bayrampasa	11.7	12.1	0	0	0	0	7.4	3.8	0	9.5	28.6	3	12.9	9.5	13	13.3	0	9.4
Sagmalcilar	6.8	15.2	9.1	0	0	5.4	5.9	3.8	0	0	14.3	3	6.5	4.8	8.7	8.4	0	3.1
Kartaltepe	9.7	3	4.5	9.7	0	0	4.4	5.8	0	4.8	0	4.5	9.7	4.8	4.3	4.8	0	0
Otogar	15.5	3	4.5	6.5	5.4	5.4	0	5.8	0	9.5	7.1	9.1	6.5	0	8.7	10.8	12.5	9.4
Esenler	13.6	21.2	22.7	9.7	13.5	5.4	5.9	0	0	0	0	7.6	6.5	9.5	7.2	12	12.5	6.3
Terazidere	1	0	0	12.9	5.4	2.7	0	0	0	4.8	0	6.1	6.5	4.8	5.8	2.4	0	0
Davutpasa	1.9	0	0	3.2	0	8.1	0	0	4.5	0	7.1	0	6.5	0	0	1.2	0	3.1
Merter	0	3	9.1	3.2	0	2.7	8.8	0	0	0	0	0	0	0	2.9	3.6	0	0
Zeytinburnu	5.8	9.1	0	3.2	2.7	8.1	4.4	15.4	18.2	4.8	0	0	16.1	23.8	17.4	10.8	12.5	15.6
Bakırkoy	3.9	3	13.6	6.5	8.1	8.1	1.5	5.8	0	4.8	0	16.7	0	0	0	1.2	0	6.3
Bahcelievler	1	6.1	0	3.2	5.4	8.1	5.9	1.9	13.6	0	7.1	7.6	0	0	4.3	2.4	12.5	0
Atakoy	7.8	0	4.5	0	2.7	2.7	11.8	5.8	9.1	33.3	0	19.7	0	14.3	0	1.2	25	9.4
Yenibosna	9.7	18.2	18.2	12.9	8.1	5.4	11.8	13.5	22.7	4.8	0	10.6	12.9	19	2.9	0	12.5	21.9
DTM	1	0	0	3.2	0	0	1.5	0	0	0	7.1	0	0	0	1.4	0	0	0
Havalimani	4.9	0	4.5	6.5	2.7	2.7	2.9	0	0	0	0	4.5	0	0	1.4	8.4	0	0

Current data represents number of passengers that enters to the stations within specified time periods. Destination of any passenger entering the system was assigned based on the probabilities determined through this conducted survey.

3.2. Model Assumptions

Simulation model was developed based on the following basic assumptions:

1. Trains operate on a single unidirectional track.
2. The system is a terminating system and works for predetermined hours of a day.
3. Speeds of the trains are system variables, but they do not change during a trip.
4. Inter-arrival times of passengers are exponentially distributed with rates given in 60 minutes intervals.
5. Dwell times at stations are function of arrival time of the passengers waiting in the station and getting off the train.
6. Only one train is allowed at a time in a station.
7. Passengers wait for trains in stations and get on the train according to the train's current capacity and number of passengers on it.
8. If there is no train to board at departure time due to a delay, passengers go on waiting for a train until train arrives. There will be no adjustment in timetable due to this delay.
9. Trains can follow different routes depending on the schedule. Half routes are allowed.
10. The system represents normal conditions. Failures in railway or trains are not included.
11. Headways are determined before the simulation model is run. Headways can be different for different time intervals, but the headway for a time interval is fixed during the running of the simulation model.

12. Simulation model of the system was developed for seven stations. Other part of the system was taken into consideration by calculating trip times according to the train cycles and adding this time to the delay times in the model.
13. All passengers in the train alight at the terminal stations.

3.3. Model Components

Number of stations, distances between stations, number of trains and their velocities are controllable input variables of the system.

In developed model, these Arena modules were used for modeling the system events: Create, Transporter, Request, Free, Dispose, Station, Assign, Route, Schedule, File, Read/Write, Signal, Hold, Delay, Decide, Separate, Record.

Create module was used to generate passengers at the system according to the defined arrival type schedules according to the collected and analyzed passenger for each station and direction. Also this module was used to direct drivers to start their trips.

Transporters were used to define limited number of trains in the system. Trains start their trips by using of request module and at the end of trips they were released with free module.

Dispose module was used to remove waiting passengers from the stations after the train arrivals.

Signal and hold modules work coordinately in the model. A signal value was sent to the related station after train arrivals and waiting passengers at hold module's queue was released and disposed.

Stations represent the predefined seven stations at the system. System was developed as a circle network and one station was defined as two stations in Arena model that represents different directions of train movements at the stations. Table 2 represents the defined stations in the model.

Table 2. Defined Stations in Developed Model

Station	Station Short Name	Station Code	Station Type
Aksaray Start	AKS Start	St0	Station
Aksaray	AKS1*	St1	Station
Emniyet	EMT1*	St2	Station
Ulubatli	ULU1*	St3	Station
Bayrampasa	BAY1*	St4	Station
Sagmalcilar	SAG1*	St5	Station
Kartaltepe	KRP1*	St6	Station
Otogar	OGA1*	St7	Station
Otogar Start	OGA Start	StA0	Station
Otogar	OGA2**	StA7	Station
Kartaltepe	KRP2**	StA6	Station
Sagmalcilar	SAG2**	StA5	Station
Bayrampasa	BAY2**	StA4	Station
Ulubatli	ULU2**	StA3	Station
Emniyet	EMT2**	StA2	Station
Aksaray	AKS2**	StA1	Station

* These stations are used for the trips at Otogar direction.
* These stations are used for the trips at Aksaray direction.

Assign module was used to make assignments or changes in system variables and attributes according to the defined equations.

Route module was used to transfer trains between stations at specified times.

Train schedules were determined outside of the system and train trip start times and directions were read from a predefined file. Separate module generates a new entity to read next departure while the entities were starting their trips.

Train dwell times at stations were kept at delay module.

Decide module was used to control signalization time factor and only trips that have minimum 2 minutes was allowed to start trips as shown in the first expression otherwise they waited until the minimum time interval was provided as shown in the second equation.

Terminal Station Departure Time >= Previous Departure Time + 2 (1)

Delay Time = 2 – (Terminal Station Departure Time – Previous Departure Time) (2)

Record module keeps the time between arrival statistics of passengers at each station and time between arrival statistics of trains for terminal stations.

Furthermore these attributes and variables were defined in the developed model:

3.3.1. Attributes

Attributes are used to attach specific values to the entities that can differ from one entity to another. In developed model these attributes were defined:

Train capacity: '1028' was assigned in the beginning of trips.

Number of passenger in the train: '0' was assigned in the beginning of trips and this value was updated at stations according to the number of passengers that got on the train at this station.

Number of alighting passengers in specified stations: Number of boarding passengers was multiplied by percentages of alighting passengers for the next stations and attributes were assigned.

Trip start time: Specified timetables include train trip start times for the terminal stations. This attribute kept these start times and it was used in delay module to define required waiting time as difference between assigned trip start time and system time.

Train route: Developed model allows defining various routes for different requirements. At terminal stations this attribute was assigned and it was controlled at defined decision points.

3.3.2. Variables

Variables are used to reflect the state of system for specific purposes and they are not attached to the entities. In developed model these user defined variable were defined:

Number of waiting passengers at stations: This variable was increased with per passenger arrival and decreased after train arrivals to the stations. Number of

waiting passengers at stations was updated after boarding of passengers to the trains as the difference between number of waiting passengers and number of boarding passengers.

Number of boarding passengers at stations: Trains have limited capacity to transfer passengers between stations. Number of boarding passengers at stations depends on available capacity of trains and the number of waiting passengers at stations. Minimum value of number of waiting passengers and available capacity of trains gave the number of boarding passengers.

Planned departure time from terminal stations: Signalization system necessitates controlling minimum time interval between departures. This variable was defined for both Aksaray and Otogar stations which are terminal stations in the system. Planned train departure time was kept in this variable.

Actual departure time from stations: This variable was used to keep the actual departure time from stations after required controls were done. Trains started their trips without any delays if difference between planned departure time and actual departure time is greater than or equal 2 minutes. Otherwise, trains departures were delayed to provide required time interval.

3.4. Model Events

Simulation model development was divided into 8 main events and these events were taken into consideration.

1. Trip startings are declared to the drivers. This event occurs according to the pre-defined timetables.
2. Drivers request the nearest train based on the smallest distance rule. Available trains are matched with drivers. If there are no any available trains, drivers will wait until trains become available.
3. Trains arrive to the stations within specified duration. These durations were obtained from Istanbul Ulasim Corporation based on previous data collections.
4. Passengers alight from the train. Number of passengers that alight at the station are determined by previously assigned attribute with these equations. The second equation is repeated until all previous stations' data was evaulated. ANINT function of Arena was used to define number of passengers as integer values.

*No. of alighting passengers at Station X = ANINT (No. of boarding passengers at beginning station * Alighting passenger ratio for station X)* (3)

No. of alighting passengers at Station X = No. of alighting passengers at Station X + ANINT (No. of boarding passengers at the second station in the train direction Alighting passenger ratio for Station)* (4)

5. Trains delay at stations for alighting and boarding activities. Delay times were obtained from Istanbul Ulasim Corporation based on previous data collections for the stations.
6. Passengers arrive to the stations and number of the waiting passengers increases.

No. of passengers at the station at the station = No. of passengers at the station at the station + 1 (5)

7. After the train arrivals, passengers get on the train and number of waiting passengers in the station decreases according to the train capacity. Following equation was used to provide this. MN function of Arena was used to get the minimum value of the available passenger capacity in the train and number of passengers at the specified stations.

No. of alighting passengers = MN ((Capacity – No. of passenger in the train), No. of passengers at the station at the Station X) (6)

8. In Aksaray station trains' route may vary due to the passenger demand and some trips are kept shorter. Routings of the trips are assigned with departure times in the beginning of the trips. Also, according to the last departure time from the terminal stations trains may be freed. Following expression was used to control this in decide module:

TNOW <= Last departure time (7)

TNOW represents the system time and increments during simulation run. Until TNOW is greater than last departure time, trains leave stations; otherwise they are freed.

Appendix A shows the structure of developed Arena simulation model.

4. EXPERIMENTATION

The model is verified by developing the model in a modular manner, using random debug points, substituting constants for random variables, manually checking the results, extreme case analysis, and animating the system.

Experiments were designed to provide a structured framework that explains simulation execution logic in terms of model components. Experimental design provides an estimation how changes in input factors affect the results, or responses of the experiment (Kelton 2000).

Scenario analyses have been conducted on the developed model. Developed model was validated with designed controlled experiments. Passenger arrivals, train schedules and number of trains that operates in railways was defined as parameters and experiments were performed on the developed Arena simulation model. Model was executed within 1200 minutes

replication length and model was replicated 40 times. Base time unit was defined as minutes.

In current system, trains are scheduled more frequenet in morning hours. Firstly, current system's average passenger waiting times have been found for spesified stations. Experiments have been conducted with 3, 5 and 10 minutes train arrival intervals. Figure 2 indicates the average passenger waiting times by changing train arrival frequencies.

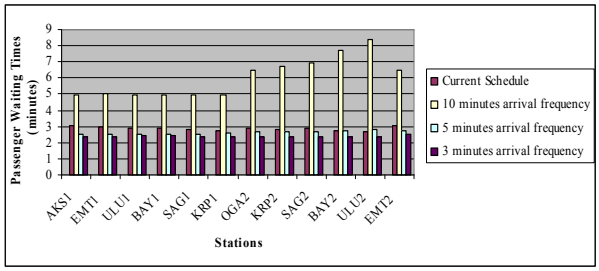


Figure 2: Average Passenger Waiting Times for Different Timetables

It was observed that decrease in headways cannot provide expected decrease in average passenger times without changing number of trains in some cases. 3 minutes time interval between trains' arrivals cannot be implemented with current capacity, 3,894 minutes average time between arrivals was measured in terminal stations.

In current system, there exists 17 trains active and 16 of these trains operate between defined 7 stations in the model. Current schedule has been used in experiments that number of trains have been changed as 5, 10, 20 and average passenger waiting times were observed.

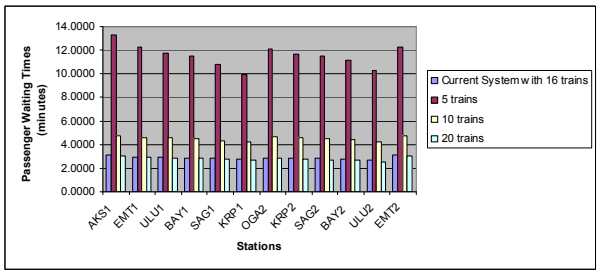


Figure 3: Average Passenger Waiting Times for Different Number of Trains

Increase in the number of trains while current timetables were being implemented did not affect average passenger waiting times. Decrease in the number of trains caused delays in predefined timetable and current table could not be implemented with 10 and 5 trains. Average passenger waiting times have increased due to increase in headways.

5. RESULTS AND CONCLUSION

Developed model was implemented for Aksaray-Havalimani light rail metro system for seven stations. In this study, the number of passengers that are transported in trains and dwell times of the trains in specified stations were analyzed and experiments were made. Passenger waiting times were obtained in the basis of existing system. Changes in the system's performances measures for different conditions were obtained by changing train frequencies and the number of trains in system.

In conclusion, increase in the train arrival frequencies during time periods when passenger arrival rate is high causes decrease in average passenger waiting times. It has been observed that the number of trains is an important factor to implement specified schedules and specified schedules cannot be applied in cases when the number of trains decreases. The simulation model built in this study is a generic model that can be easily changed to adapt the changes in the number of trains and train schedules. By running the model, several what-if questions of operations management can be replied to make revisions. It might also provide suggestions on strategic infrastructure development, and thus provide a starting point for the construction of any new infrastructure, namely capacity planning. It has been observed that the number of trains is an important factor to implement specified schedules and specified schedules cannot be applied in cases when the number of trains decreases.

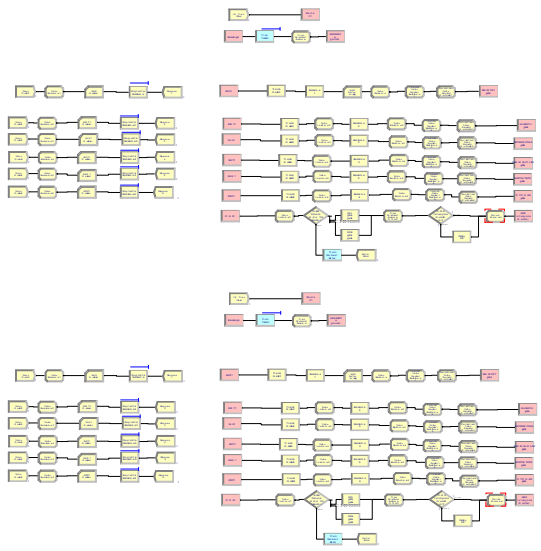
This study represents an efficient tool to measure passenger dwell times in stations by changing the train schedules and feasibility of pre-specified train schedules with by measuring its effect on rail transportation.

6. FUTURE RESEARCH

The scope of the second stage of the research is to expand the model by adding other stations of the network and provide integrity to evaluate the system performance. Also, transfer times between stations will be defined as user defined variable and other possible timetables' effects on the system will be measured.

APPENDIX

APPENDIX A. Developed Arena Model



REFERENCES

Ghoseiri, K., Szidarovszky, F., Asgharpour, M.J., 2004, A multi-objective train scheduling model and solution. *Transportation research*, 38 (10), 927-952.

Hoohiemstra, J.S., Teunisse, M.J.G., 1998. The use of simulation in the planning of the Dutch railway services. *Proceedings of the Winter Simulation Conference*, pp. 1139-1146. December 24-28, Washington (Washington, USA).

Kelton, W.D., 2000. Experimental design for simulation. *Proceedings of the 2000 Winter Simulation Conference*, pp. 32-37. December 10-13, Orlando (Florida, USA).

Vansteenwegen, P., Oudheusden, D.V., 2007. Decreasing the passenger waiting time for an intercity rail network. *Transportation research,* 41 (4), 478-492.

Yalcinkaya, O., Bayhan, G. M., 2008, Modeling and optimization of average travel time for a metro line by simulation and response surface methodology, *European journal of operational research*, 196 (1), 225-233.

AUTHORS BIOGRAPHY

FILIZ DUMBEK graduated from Management Engineering department of Istanbul Technical University in 2009.

DILAY CELEBI is academician of Management Engineering department at the Istanbul Technical University. She received her Ph.D. from Istanbul Technical University in 2008.

CONFIGURATION OF A SERVICE AND DISTRIBUTION NETWORK IN THE AEROSPACE INDUSTRY

Marcos Simas Magalhães [(a)], Miguel Cezar Santoro [(b)]

[(a)(b)] Dept. of Production Engineering, University of São Paulo

[(a)]msimasm@gmail.com, [(b)]santoro@usp.br

ABSTRACT
The evaluated company develops, manufactures and supports aircraft for the commercial, executive and military markets, and is developing new products for the executive market. In order to support this fleet in the American market, a new spare parts distribution and maintenance services network is being developed. This work has the objective of defining the quantity and location of service centers and the corresponding distribution center. To do that the spare parts inventory levels in each location were taken in consideration in order to minimize transportation and inventory carrying costs and attend the established service levels in an alignment with the competitive and supply chain strategies. The methods employed were simulation, to evaluate each alternative, and meta-heuristics, to search for the best solutions. The results led to locating the distribution center in Menphis and to a quantity of 7 service centers, much less than the current solution.

Keywords: logistic strategy, distribution network, service network, spare parts in the aerospace industry

1. INTRODUCTION

The problem evaluated in this study is to define the best location for the distribution center and the quantity and best location for the service centers in the U.S. These centers will support the new aircraft fleet, which is currently being built by an executive aircraft industry. The distribution center and the service centers will contain stocks of spare parts and their levels will be taken under consideration for the location decisions. The stocks will be used as source of replacements components for defective ones removed from aircraft, when so is needed. All decisions should be taken so that transportation costs, stock carrying costs and fixed costs of the service centers are minimized, given a minimum level of service required.

The scenario for the problem is an industry characterized as highly cyclical. This led the company to seek in the executive aviation market, among others, a way to soften the impact of unexpected fluctuations in demand in their business. That market was chosen having in mind the predictions that foresee its steadily growth in the coming years, due to factors such as the airports increasingly security procedures for commercial flights, which is leading a number of executives to seek the flexibility, exclusivity and comfort provided by the executive aviation, and the economic globalization, which demands frequent business trips.

This new executive aircraft that are being developed will require a completely different support from the one the company usually provides for civil and military aviation. The aircraft future owners are mainly private executives with no available infrastructure for maintenance, spare parts logistics, pilots training, etc. Furthermore, the market has signaled that a decisive factor for the purchase decision is the perception on the company's ability to provide an adequate support for the fleet in action. To meet this demand it is necessary, among other things, to establish an appropriate network of logistics and services. This article studies the distribution and services network in the U.S., for the new executive aircraft. The network is showed in the following schedule:

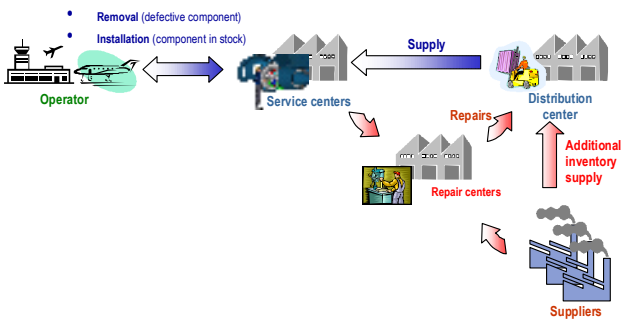


Figure 1: Flow of materials for the support to the aircraft fleet

The process begins when a failure is detected in the aircraft during its operation or during maintenance. When the failure happens during operation, a spare part is obtained from the DC (Distribution Center) or from the SC (Service Center) where maintenance is taking place, after the defective component has been identified and removed. This means that the SCs can also serve regional distribution centers, giving them a very quick response.

If the removed component is a consumable one, (i.e., cannot be repaired), the material sent by the SC is replaced by the DC which, in turn, is restocked by the supplier. If it is a repairable item then it is also replaced by the DC, but only in this case the component removed from the aircraft is sent to a repair center in order for them to be restored and then sent back to the distribution center. This happens when the repairable part cannot be repaired for technical or commercial reasons. It is then discarded and a new one is sent to the DC by the supplier.

2. PROBLEM DEFINITION

The issue covers the following decisions on the distribution network and services:

- The spare parts distribution center location;
- The amount and location of the service centers, which besides providing maintenance services, will also serve as advanced inventories of spare parts;
- The amount of the spare parts stocks (it will be considered a sample of items, to make it possible to take into account the different levels of inventory for each alternative configuration of the DC and the SCs).

This is a strategic-level issue, and therefore has long-term impacts on the results of the company's operation. The question is how to minimize the cost of supporting the fleet which is ready to fly in the United States, sustaining a given level of service. The cost is composed of three values:

1. The cost of the spare parts stock for defective components of the aircraft.
2. The cost of the spare parts transport.
3. The fixed cost of the service centers.

The restrictions of the problem are:

1. The fleet availability should be equal to or greater than the target set by the company's competitive strategy. The fleet availability will be the indicator to measuring the level of customer service, in the cases of parts that could prevent the aircraft from flying. The fleet availability is measured by the percentage of time that the aircraft are available for flight. An airplane is considered unavailable when undergoing maintenance or being transported to a repair center. In the figure below, the slot of time for obtaining spare parts is represented by the fragment called "Logistic Process", which is part period of time necessary for maintenance. This part can be influenced by the design of the support network.

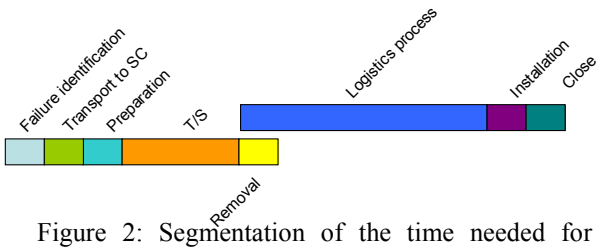


Figure 2: Segmentation of the time needed for aircraft maintenance.
Source: Evaluated company

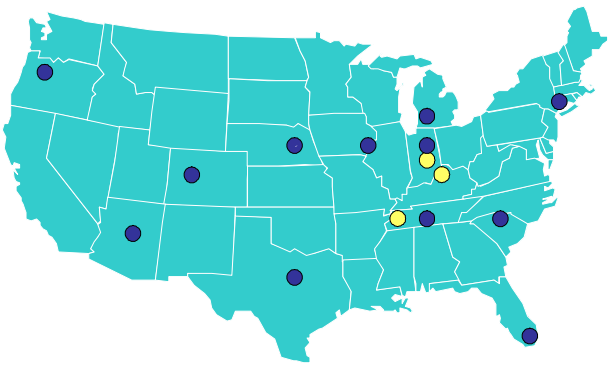
2. For those parts which do not prevent the aircraft from flying, the fill rate will be used as a service level measure, i.e., the percentage of requests fulfilled within a set time. Despite the fact that those parts do not hinder the flight, the situation may have a strong impact on customer satisfaction. It can impact, for example, in the comforts inside the aircraft, and therefore this kind of maintenance will also be considered as a restriction.

3. Adherence to the standards of service provided to business clients over a maximum distance of a two-hour trip from the customer's base to the nearest service center. This additional restriction on the fleet availability is due to the fact that a medium level of service is taken under consideration when availability is concerned, while our objective here is to establish a maximum distance to prevent any customers from having an inadequate service.

4. The service centers shall contain at least one piece for items classified as "Aircraft On Ground" (AOG), according to what has been disclosed to the market by the company. These are parts that, if defective, can prevent the plane from flying.

5. The support network must contain at least one piece for those ones which are not classified as GBS. Here again, the concern is not the average level of service, but to avoid situations where the service time will be beyond a maximum acceptable level, even if in special situations.

Those are the possible locations that will be evaluated:



Distribution centers
Service centers

Figure 3: Possible locations for distribution center and service centers.

3. LITERATURE REVIEW

As the main restriction to the problem concerns the service level offered to customers, it is fundamental that the proposed solutions are strongly based on the competitive and functional strategy theories. This will make it possible for us, in an unfolding process, to define in a structured way which values should be a parameter restriction.

To meet the desired alignment between competitive strategy and supply chain strategy, it is necessary to understand the client needs in each one of the competitive strategy target segments, to fully understand the supply chain, and then achieve the alignment between the two. Therefore, we need first to define as attributes, the customers' needs which, in accordance to the competitive strategy, we need to attend. They can be classified by price or by differentiation attributes.

As a second step, in order to understand the supply chain, Chopra (2001) proposes a methodology similar to that used to understand the demand: to use a measure of how responsive the supply chain is, in order to provide a consolidated view of how the chain attributes are configured. The major point for attention here is that the more responsive a supply chain is, the less efficient it tends to be. And by efficiency he means the cost incurred to produce and deliver products. Finally, in order to achieve the desired alignment between competitive strategy and supply chain strategy, both of them should aim at the so-called company's zone of strategic alignment. In this zone, the supply chain's responsiveness increases according both to the level of the demand's uncertainty and to the product's life-cycle phase (initial stages require more responsive chains).

Ballou (1998) sustains the hierarchy of decisions concerning the supply chain taking the time frame which they are linked to, as a reference. Decisions at strategic level involve long horizons of time, usually greater than one year, and are based upon aggregate data. Another important point mentioned by Ballou is the trade-offs within the supply chain decisions, both at the same and among different levels of hierarchy such as: the location of deposits, the levels of stock and the size of lots to be used to resupply.

The heuristic and optimization methods are the most commonly used in location and configuration of logistics networks surveys.

Hale et al. (2003) offers a historical perspective on the development of research related to location of facilities. He makes a selection of articles that undoubtedly contributed to its development: Weiszfeld (1937); Weber (1909); Chrystal (1885); Hotelling (1929); Hakimi (1964); Cooper (1963).

Powell (2005) points out advantages and disadvantages in the use of the methods of simulation and optimization via linear programming. On one hand, optimization methods tend to be preferred by scholars due to the intelligence involved in solving problems and the guarantee to find the best solution. Turns out this best solution assumes various factors as previously defined, but they are not always fully known in real world. On the other hand, the simulation methods can represent, with a quite great adhesion, and with the use of rules, the high complexity and the various uncertainties involved in an operation. Since the rules can be adjusted and tested with relative easiness, people directly involved in evaluating operations, tend to prefer the simulation due to this flexibility. The article suggests that both types of methods can be used jointly, in the case of problems involving resources allocation. One should, thus, try to take advantages from both methods.

Several authors, such as Ballou (1984), Goetschalckx et al. (2002), Nozick et al. (2001), Shen et al. (2005), Daskin (2002), Shen et al. (2007), Hinojosa et al. (2008) and Syam (2002) suggest that it is important to consider the inventory levels and its cost, in the location model so that the models in fact consider all the costs involved in the problem. In order to achieve this, the heuristic and simulation models are often pointed out as good alternatives. Keskin et al. (2007) use a heuristic algorithm based on scatter search, including local search and path-relinking routines to solve the problem of finding two stages, with capacity restriction. In comparison to the exact method, there is gap lower than 1%, in relation to the optimal solution, and significant gains in computational performance were obtained. According to the literature search that the author has made, the heuristic methods used in problems of locating the p-median are: heuristics interchange, genetic algorithms, tabu search, simulated annealing, scatter search, greedy randomized adaptive search and a heuristic based on dynamic programming.

According to Glover (2003), the meta-heuristics methods are the ones that organize an interaction between local improvement procedures and high-level strategies in order to create a process which is able to escape from local optimum and to perform a deep search in the solutions area. Lately, these methods have been widely used, in solving various types of complex problems, in particular those of combinatorial nature, for it has proven to be extremely effective in providing near optimal solutions, in a reasonable processing time. Although they do not guarantee as good solutions as in exact models, these last ones often result in processing times which prevent their use. Moreover, some successful applications of exact models have incorporated meta-heuristics strategies.

Muckstadt (2004) puts a great focus onto a type of stock policy called the base stock (s-1, s), in which every time there is any consumption of material, the same amount is resupplied. As the author shows in his book, this is the most suitable model for items of rarefied demand and of high aggregated value, where the order cost is irrelevant when compared to the value of the parts that are being considered.

According to Muckstadt (2004), a safety stock in a stock policy, should be determined both according to the demand's variation and to the resupply time's variation as it has the function of absorbing both of

them. As shown by Muckstadt, the definition of the optimum levels of stock in a two-link network with rarefied demand, which behaves as sparse distribution of Poisson probabilities, using exact models is not feasible from a computer. Alternatively, the author suggests using the classic METRIC (Multi-Echelon Technique for Recoverable Item Control) model. The author demonstrates mathematically that the optimal stock of the central deposit is almost always in the interval [(λiDi, 2 * (λiDi) 1 / 2 + λiDi], where λi is the rate of demand for item i at the central depot and Di is the time to resupply the central deposit. This is, therefore, the range in which the search method will look for the best solution.

Lee et al. (2005) propose the use of dynamic simulation to evaluate the inventory policies for repairable spare parts. An optimization model is used to determine the policies to be tested. Zanoni et al. (2005) use the dynamic simulation software ARENA in conjunction with the OptQuest resource within the same tool, in order to find optimal decision variables. According to Kelton et al (2002), besides using whole programming, this resource uses various types of heuristics such as tabu search, scatter search and neural networks, in order to define the values of variables. Dias et al. (1998) proposed the use of dynamic simulation coupled with a search method for determining stock policies for spare parts.

4. LITERATURE DISCUSSION

To start the discussion about the evaluated literature, we can notice it provides the basis for carrying out the work, but in a fragmented way. This means that all the conceptual and methodological elements are present in the literature, but in an ungrouped way, in relation to the set of issues addressed in this article.

The competitive strategy and its deployments to the supply chain functional strategy, has a clearly structured theory. It provides the desired performance parameters to support the company's strategic vision. There has been much research on the strategic role that the service area, with its peculiar supply chain, is currently assuming to companies, due to its potential to provide differentiation and high levels of profitability. Still, the connection among the issues, including the deployment of competitive strategy in the supply chain strategy of specific service, are not yet addressed in an appropriate level of depth. For this research, the entry into a new market, the executive aircraft, aiming at differentiation, is very clearly part of the company's competitive strategy. In such market the products have extremely high added-value, and the quality of after-sales services is crucial to the customer's satisfaction and also to differentiate the company from its competitors. Therefore, it is justified considering as restriction the service levels characteristic of the largest market in the world. Furthermore, we are dealing with a new product in a market which demands great variability. According to Chopra (2001) these two factors suggest the need for a responsive supply chain. Thus, the distribution and services network should be defined in line with the company's strategy.

We have found an extensive literature, most relatively new, on location models which consider logistics costs as costs of inventory loading, transport and applications' processing. By doing so, more results with adherence to the operational scenario of the location problem are expected, thus adding more credibility for the company, among the users of the models. Similarly, the same happens with spare parts planning models, which have received much attention in recent years, due to their inherent complexity, distinctiveness (mainly because of typically sparse and difficult to predict demand) and relevance to the business. However, there was found no article or reference to location models which would take into account levels of spare parts inventory. This, of course, represents a further complexity once, as commented before, the stock levels could not be estimated in the traditional way, even considering non-linearity. Different models for service-provider facilities and materials distribution centers were found in literature, but not for both of them altogether. This is what the present case needs, once the aircraft maintenance centers will also be used as storage facilities.

Thus, from a conceptual and methodological stance, the research was based on a sequence of the evaluated theory on competitive strategy and its deployments for the operational supply chain strategy, the strategic relevance of the service area, the location models which consider logistic costs, and planning models for spare parts.

5. METHOD

The method used to solve the problem was the dynamic simulation, in conjunction with a search heuristic with the objective of achieving, within the highly complex problem scenario, a high level of adherence to the real operation system model, and a processing time able to analyze different alternatives. Such proposal is based on findings from the literature review, as outlined below.

According to the so far the evaluated literature, which addresses the importance of considering the inventory levels while dealing with location problems, the use of methods traditionally applied to location problems, is not adequate to handle the complexity of the interdependence of decisions of location and inventory: Ball (1984); Shen et al. (2005); Shen et al. (2007); Powell (2005); Nozick et al. (2001); Syam (2002); Goetschalckx (2002) and Hinojosa et al. (2008). Among them, the mixed integer programming one has a special place: Owen (1998); Hale (2003) and Klose (2003).

Moreover, there are several uncertainties involved in the operation, especially regarding the demand behavior. This is intensified by the fact that the spare parts stocks have an erratic and highly unpredictable demand. Therefore, they demand some specific planning models, which are often based on heuristic methods which aim at determining the stocks optimum

levels, as suggested by Cohen (2006), Muckstadt (2004), Wong et al. (2005), Sherbrooke (2004), Lee et al. (2005), Batchoun et al. (2003), Zanoni et al. (2005) and Dias et al. (1998). In fact, some of the authors of articles with location models that consider the costs of stock loading, draw attention to the simplifications they had to do to make the model viable and to the specific cases in which such simplifications can be applied: Nozick et al. (2001) and Shen et al. (2007).

To apply the simulation method here described with the heuristic search the ARENA system was used in its 10:00:00 version, property of the U.S. company Rockwell Software. With this program it is possible to model an operation in a very realistic way, considering the fleet size, the levels of inventory, time, costs and their probability distributions for each stage of the process. The results of the simulation can be obtained through the creation of variables.

The operation of the model specifically developed for this work occurs as follows:

1. Each demand for a replacement part represents an entity which is generated within the model. This takes place following specified periods of time, which have been determined in accordance to the rate of defaults in the parts, the probability distribution, the fleet size and the number of parts used in each aircraft.
2. Since the entity has been created, the next step is to characterize the demand for which the part was created, the location, priority and type (scheduled or not scheduled demand).
3. In the case of a demand coming from an airport, the time for sending the piece from the distribution center is simulated (when the piece is not available in stock, the time it takes to arrive there is also added).
4. In the case of a scheduled maintenance, the aircraft flight time to the service center is simulated.
5. When the demand comes from a service center, its consumption is immediate when the piece is in stock. Otherwise, the time taken for the piece to arrive from the DC is added. If the piece is not available in the DC as well, the time it takes until it arrives is equally counted.
6. As soon as each piece is consumed and the demand is met, the replenishment process starts. If the stock is consumed in the SC, it is replenished by the DC, and if it is consumed in the DC, it is replenished by a supplier or a repair center. In the case of repairable parts, a process of sending the defective part to the repair center is also initiated.
7. Throughout the simulation, the transportation costs and the inventory levels are monitored during the periods when the aircraft is unavailable, in order to provide data for a performance assessment for the considered alternative.

Besides the simulation itself, the ARENA system has a module called Optquest that uses, as exposed in Glover (1995), heuristics search, such as "scatter search", in order to determine some pre-determined parameters, so that an objective is maximized or minimized. The variables that will go through the Optquest seeking process are the opening and closing of the various service centers that can be used in the distribution and services network, as well as the parameters for each and every piece base stock (level below which the stock must be restocked), for the distribution center as well as the service center. The objective is to minimize the total cost composed by the inventory carrying cost, the logistic transport cost (calculated based on the cost per shipment plus the cost per miles of each specific transport) and the fixed cost of each open service center.

To evaluate the three alternatives of location for the distribution center, it will be done different runs of the model with appropriate changes of distance and cost of transport. Since each location is specific to a logistics provider, it is associated with different costs related to each provider. In the parameterization of the model, each simulation will be made for a period of 1 year using the average of 3 replications. For each scenario of the distribution center 500 different simulations will be run within the Optquest for finding the best solution. This number was defined from the observation that there are not significant gains in the results after 300 simulations.

For all parts the distribution center should have inventories inside the interval [(λiDi, 2 * (λiDi) 1 / 2 + λiDi], where λi is the rate of demand for the item i and Di is the time to resupply in the central deposit. As mentioned in the literature review, Muckstadt (2004) showed that the optimum level of stock in the central deposit is almost always in that interval. Inventories bases and centers of service may vary between 0 and the maximum amount set for the distribution center if they are non-critical parts. For critical parts, the minimum quantity in stock is one unit in each service centers. The data used in the simulation obtained from the evaluated company and had to be changed for confidentiality purpose.

The parameters related to the level of service offered to the customers were defined based on the theoretical foundations of competitive strategy and deployment to services and supply chain strategy and the specific settings of the market and the company. The conclusion of the discussion on literature analysis is that the focus of supply chain strategy should be more on differentiation of the service than in reduced costs. The level of service should therefore produce results equal or higher than the best found on the market. On this basis, the parameters to be used as restraints in simulations are:

- Minimum availability of the fleet: 98%
- Minimum Fill rate for non-critical items: 95%
- The percentage of flights to a minimum center services performed in less than 2 hours: 99% (there are specific places as Hawaii and Alaska where the target of 2 hours is not economically feasible, that's why 100% was not used)

6. RESULTS AND DISCUSSION

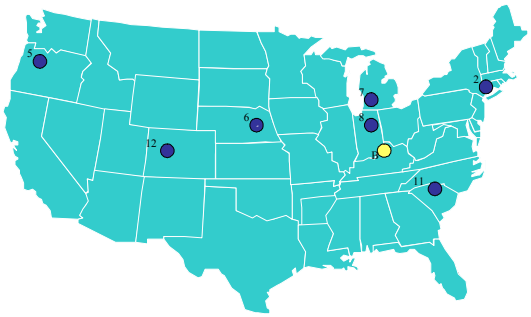
The searches were executed in Optquest for 3 different scenarios, each of them for a distribution center in a different location. Whereas the computer used, a Toshiba Satellite notebook with an Intel Pentium 1.73 GHz and 1 GB of RAM, is default now on the market, the performance can be considered quite satisfactory (times between 32 and 36 minutes). It should be noted, however, that significant simplifications were adopted, such as the inclusion of only 10 pieces in the simulation. To facilitate benchmarking between the 3 scenarios, follows a table with a summary of key results:

Table 1: Main results

Amounts in US$ Millions	Scenario 1	Scenario 2	Scenario 3
Stock carrying cost	2.62	2.56	2.64
Service centers fixed costs	0.6	0.7	0.6
Transport costs	2.46	1.78	2.89
Total cost	5.68	**5.04**	6.13

The final solution to the problem, found after 500 simulations performed in a period of 35 minutes, was the location of distribution center in Memphis and the use of 7 service centers located as below map at:

2 - Windsor Locks (CT)
5 - Hillsboro (OR)
6 - Lincoln (NE)
7 - Battle Creek (MI)
8 - Indianapolis (IN)
11 - Greenville (SC)
12 - Grand Junction (CO)



Distribution centers
Service centers

Figure 4: Selected service centers in the solution of the problem

The restrictions in the level of service (availability of the fleet, fill rate and time of flight to the service center) are all satisfied with this solution. The biggest operational cost is the inventory carrying cost (resulted from an average stock of $ 25.61 million), followed by the cost of transport and finally the fixed cost of service centers.

The fact that the biggest cost is the inventory carrying cost is consistent with the type of problem treated, commonly called asset intensive, where the amount invested in assets, in stocks of spare parts, usually costs more than other operational factors, such as the transport. This result reinforces the need for considering the cost of carrying inventory in such problems.

7. CONCLUSIONS

The solution obtained as a result of the proposed model allows the achievement of the defined levels of service to ensure the company's differentiation in the market, and at the same time minimizing operational costs, taking into account the concept of total logistics cost.

Considering the motivation of this work, we were able to meet the original purpose of proposing a viable, realistic and effective method. The method is feasible because it was demonstrated its use in a specific problem with reasonable processing times. It was possible to make its application to be realistic because, despite the simplifications used, the simulation model reproduced with great fidelity the operational environment to support the aircraft fleet. Its effectiveness comes from both the alignment of the solution with the competitive strategy of the company as well as the correct consideration of the costs and time involved in that, given the level of service defined as a concept that actually reflects what impacts the satisfaction of customers of executive aircraft.

Possible extensions of this research are:

- Research in meta-heuristics that in the case of location problems considering spare parts inventory levels have better performance. This will enable greater performance and subsequent acceptance of the method proposed in this paper for the type of problem here treated.
- Construction of simulation models that result in less than the simplifications adopted in this work. It will always be possible to include new considerations in the simulation model so that it is a better representation of the reality.
- Studies comparing the performance of the method used with traditional methods of optimization. It would require the evaluation of results in real applications of the methods so that the conclusions are not distorted by the various simplifications made in each case.
- Advancement in the study of the concept of fleet availability for measuring the level of service to the market of executive aviation. Currently, this is used mainly for military aircraft, while in other markets fill rate is more common. As a measure of satisfaction from customers of aircraft executives, it seems to make more sense to use the percentage of time the aircraft is available than the percentage of parts that were delivered within a specified period. Also, using the measure of availability allows including the time of flying the aircraft to the centers of service, when the plane is not available to the owner.

8. REFERENCES

Ballou, Ronald H., 1984. DISPLAN: A Multiproduct Plant/Warehouse Location Model with Nonlinear Inventory Costs. *Journal of Operations Management* Vol. 5. No. I, 75-90.

Ballou, Ronald H., 1998. Business Logistics Management. *New Jersey: Prentice Hall, Upper Saddle River.*

Batchoun, P., J. A. Ferland, & Cléroux, R., 2003. Allotment of aircraft spare parts using Genetic Algorithms. *Pesquisa Operacional* 23(1:141–159.

Chopra, Sunil & Meindl, Peter, 2001. Supply Chain Management, Strategy, Planning and Operation. *Prentice-Hall, Upper Saddle River, New Jersey.*

Cohen, Morris A., Agrawal, Narendra & Agrawal, Vipul, 2006. Winning in the Aftermarket. *Harvard Business Review OnPoint* 4311. 129-138.

Dias, George Paulus Pereira & Corrêa, Henrique Luiz, 1998. Uso de simulação para dimensionamento e gestão de estoques de peças sobressalentes. *Anais do SIMPOI 1998, FGVP.*

Glover, F., 1995. Scatter Search and Star-Paths: Beyond the Genetic Metaphor. *OR Spektrum*, vol. 17, 125-137.

Glover, F. & Kochenberger, G., editors, Handbook of Metaheuristics, 2003. *Boston: Kluwer's International Series.*

Goetschalckx, Marc, Vidal, Carlos J. & Dogan, Koray, 2002. Modeling and design of global logistics systems: A review of integrated strategic and tactical models and design algorithms. *European Journal of Operational Research* 143, 1–18.

Hale, Trevor S. & Moberg, Christopher R., 2003. Location Science Research: A Review. *Annals of Operations Research* 123, 21–35.

Kelton, W. D., Sadowski, R. P. & Sadowski, D. A., 2002. Simulation with Arena, *McGraw-Hill, Boston.*

Keskin, Burcu B. & Üster, Halit, 2007. A scatter search-based heuristic to locate capacitated transshipment points. *Computers & Operations Research* 34, 3112–3125.

Keskin, Burcu B. & Üster, Halit, 2007. Meta-heuristic approaches with memory and evolution for a multi-product production/distribution system design problem. *European Journal of Operational Research* 182, 663–682.

Klose, Andréas & Drexl, Andréas, 2003. Facility location models for distribution system design. *European Journal of Operational Research*.

Lee, Loo Hay, Teng, Suyan, Chew, Ek Peng, Lye, Kong Wei, Lendermann, Peter, Karimi, I. A., Chen & Yankai, Koh, Choon Hwee, 2005. Application of multi-objective simulation-optimization techniques to inventory management problems. *Proceedings of the 2005 Winter Simulation Conference.*

Muckstadt, John A., 2004. Analysis of Service Parts Supply Chain Systems. *School of Operations Research and Industrial Engineering,* Cornell University. Ithaca, NY 14853.

Nozick, Linda K. & Turnquist, Mark A., 2001. Inventory, transportation, service quality and the location of distribution centers . *European Journal of Operational Research* 129, 362-371.

Owen, Susan Hesse & Daskin, Mark S., 1998. Strategic facility location: A review. *European Journal of Operational Research*, 111, 423-447.

Powel, Warren B., 2005. The Optimizing-Simulator: Merging Simulation and Optimization Using Approximate Dynamic Programming. *Proceedings of the 2005 Winter Simulation Conference.*

Shen, Zuo-JunMax & QI, Lian, 2007. Incorporating inventory and routing costs in strategic location models. *European Journal of Operational Research* 179, 372–389.

Shen, Zuo-Jun Max & DASKIN, Mark S., 2005. Trade-offs Between Customer Service and Cost in Integrated Supply Chain Design. *Manufacturing & Service Operations Management* 7(3, 188-207.

Sherbrooke, Craig C., 2004. Optimal Inventory Modeling of Systems. *Boston: Kluwer*.

Syam, Siddhartha S., 2002. A model and methodologies for the location problem with logistical components. *Computers &Operations Research* 29, 1173-1193.

Wong, H., Houtum, G.J., Cattrysse, D. & Oudheusden, D. Van, 2005. Simple, efficient heuristics for multi-item multi-location spare parts systems with lateral transshipments and waiting time constraints. *Journal of the Operational Research Society* 56, 1419–1430.

Zanoni, Simone, Ferretti, Ivan & Zavanella, Lucio, 2005. Multi echelon spare parts inventory optimization: a simulative study. *Proceedings 19th European Conference on Modelling and Simulation.*

AUTHORS BIOGRAPHY

Marcos Simas Magalhães was graduated in production engineering in the University of São Paulo in 2000 and has been working in the spare parts planning and logistics department of a brazilian aerospace and defense company since then. In 2002 he received the CPIM (Certified in Production and Inventory Management) certification from APICS and in 2009 he received the master in logistics systems engineering from the University of São Paulo.

Miguel Cezar Santoro is Associate Professor of Operations and Logistics Management in the Polytechnic School, University of São Paulo. He earned his PhD in Industrial Engineering from the University of São Paulo. He teaches undergraduate and graduate courses in Operations and Logistics Management. His research areas of interest include production planning and scheduling, inventory and distribution.

MONTE CARLO VALUATION OF POWER GENERATING UNITS

Ivan Ourdev[(a)], Simaan M. AbouRizk[(b)], Mohammed Al-Bataineh[(c)]

[(a)]Research Associate, Hole School of Construction Engineering, University of Alberta, Edmonton, Alberta, Canada
[(b)]Professor, Hole School of Construction Engineering, University of Alberta, Edmonton, Alberta, Canada
[(c)]Analyst, S.M.A. Consulting, Edmonton, Alberta, Canada

[(a)]iourdev@construction.ualberta.ca, [(b)]abourizk@ualberta.ca, [(c)]mohammed@smaconsulting.ca

ABSTRACT
Electricity deregulation in North America has enabled owners of small electric generators to become independent power producers (IPP) by selling electricity to the grid. Uncertainty affects both the profit maximizing utilization schedule and the related valuation of the underlying asset, the power generating unit (PGU). In addition to price uncertainty, the small sized power generating units are subject to different constraints and fuel supplies compared to those of large power plants, with higher production costs. Monte Carlo simulation is applied to evaluate the profitability of small PGUs as potential electricity producers and stochastic methods are used to model the prices of electricity and fuel. We extended the mean-reverting stochastic differential equation for the electricity prices with term modeling the price spikes. The spike distribution was done using the 'peak over threshold' approach of the Extreme Value Theory. The profitability measures, net present value and internal rate of return, were calculated by modeling the term structure of the interest rate with extended Nelson-Siegel form.

Keywords: Monte Carlo simulation, price models, Nelson-Siegel method, extreme value theory

1. INTRODUCTION

The electricity industry is concerned with the processes of electricity generation, electricity transmission from location (e.g. power plant) to location (e.g. populated area), and electricity distribution to the customers. Until the beginning of the 1990s, the industry was characterized by a relatively small number of government regulated and vertically integrated monopolies that covered the whole process. The pressure of globalization and the desire for lowering the price of electricity brought about restructuring and deregulation in the electricity markets in many industrial countries. One of the first areas to adopt these changes was the province of Alberta, Canada. The transition to fully deregulated electricity generation in Alberta took several years and established a competitive market, where the spot price of electricity is determined by the forces of supply and demand.

Deregulation of the electricity industry opened the door for owners of small electric generators to become independent power producers and sell electricity to the grid. A typical PGU would be a part of an emergency power system and would be idle most of the time. Most PGUs have a generating capacity of 0.5 – 2MW and are driven by diesel engines, resulting in a high cost for the produced electricity. With uncertainty in the price of fuel and electricity, there is no simple way to evaluate the potential profitability of using such generators to sell electricity to the grid.

One approach is to apply Monte Carlo simulation within the framework of capital budgeting. The first step is to build models for the price processes of the electricity and fuel. For a given realization in these prices the generator would be 'switched on' if the price of electricity exceeds the price of the fuel necessary to produce it. The present value of the generated cash flows are discounted by the appropriate required rate of return to obtain the values of two profitability measures: net present value and the internal rate of return. Generating a large number of sample price paths allows accurate estimates of the profitability measures to be obtained.

The major challenge in this approach is to build a good model for the price of electricity, accommodating its unique features. We modeled these features by specifying a seasonal mean-reverting stochastic model on the basis of standard geometric Brownian motion. To account for the spikes, we first used the 'peak over threshold' method of the Extreme Value Theory to find a high threshold above which the data is well described by the generalized Pareto distribution. We used the threshold excesses to fit an independent 'spike' term that is a product of spike intensity and spike frequency.

An additional source of uncertainty was the discount factor used in the calculation of the net present value and the internal rate of return. We opted for using these two discounted cash flow measures, as opposed to more refined approaches such as real options, because of their intuitive simplicity and ease of communication with decision makers. The simplest approach would have been to specify a required rate of return to use as a discount factor throughout the project lifetime. Because the typical length of the latter is 10-15 years, different

heuristic discount factors would yield quite different results. We decided to use instead a benchmark approach with the risk-free rate as a discount factor. We modeled carefully the full term structure of interest rates by fitting the extended Nelson-Siegel form on the available Government of Canada bonds.

2. POWER PLANT VALUATION

The current widely accepted capital budgeting method for assessment of capital investments is the net present value (NPV). It captures the intuitive guiding principle that a project should not be undertaken if the investments in the project outweigh the revenues from the project, or more formally – if the expected rate of return on the investment is less than the opportunity cost of capital. In practice, the calculation of NPV uses the discounted cash flow (DCF) method to obtain the present value (PV) of expected cash flows C_t from the project and compare it with the estimated (initial) capital outlay C_0:

$$NPV = PV\left(\sum_t E\left[C_t\right]\right) - C_0. \tag{1}$$

A project is accepted if its NPV is positive and rejected otherwise. In general, for project that has T+1 number of discrete cash flows that can be both positive (inflows) and negative (outflows) over the project lifetime T, the formula can be written as:

$$NPV = \sum_{t=0}^{T} \frac{E\left[C_t\right]}{(1+r_t)^t}, \tag{2}$$

where the cash flows C_t are explicitly discounted at an appropriate hurdle rate rt (required rate of return) at time, t. For the case of power generating units, the cash flows C_t are the net difference between the price at which the generated electrical energy was sold and the price, paid for the fuel that was used to generate that unit of electricity, i.e.

$$C_t = q(P_t^e - hP_t^f) \tag{3}$$

where P_t^e is the spot price of electricity (\$/MWh), P_t^f is the price the fuel was bought at time $\tau < t$ (\$/MMBtu), h is the unit heat rate (MMBtu/MWh) and q is the unit output level, which, for small PGUs, is usually close to 1. The quantity in the brackets is called the spark spread. Since the generator is only switched on if it is deemed profitable to do so, the spark spread in this case is always positive.

A second popular measure is the internal rate of return (IRR), which is defined as the discount rate corresponding to zero NPV, i.e. it is the rate that solves the equation:

$$\sum_{t=1}^{T} \frac{E\left[C_t\right]}{(1+IRR)^t} - C_0 = 0. \tag{4}$$

For a small number of periodic cash flows, Equation (4) can easily be solved numerically. In the case of frequent cash flows over a long period of time, T, a different approach is needed. We adopt a Future Value (FV) point of view (Ourdev and AbouRizk 2006) and calculate the total amount of cash generated over the lifetime of the project by summing up all the cash flows compounded by the corresponding forward rate,

$$FV = \sum_{t=1}^{T} E\left[C_t\right] F^c(0, t_0, T - t_0). \tag{5}$$

The continuously compounded forward rate, $F^c(0, t_0, T - t_0)$, is the rate as seen from time $t=0$ that starts at time t_0 with residual maturity T-t_0. The total return rate over the period is obtained by solving

$$(1 + IRR)^T = FV / C_0. \tag{6}$$

The two measures, NPV and IRR, require a continuous set of interest rates. A continuous function representing the spot rates, r_t, for different times to maturity is referred to as term structure of interest rates. There is a limited number of fixed income securities, and only a few of them are zero-coupon bonds, i.e. can provide spot rates. The rest are coupon-bearing bonds, and their yield to maturity depends on the coupon rates.

The zero-coupon bonds are typically represented by the Treasury bills, available with maturities up to one year. The remaining portion of the spot rate curve has to be estimated from the available longer term Government bonds, decomposing each coupon bond into a portfolio of zero-coupon bonds, with each zero-coupon bond corresponding to a coupon payment. The implied forward rates can be obtained from the spot rate curve by representing each forward rate as a synthetic portfolio of two bonds with equivalent cash flows (Martellini et al. 2003).

The most flexible estimation for the spot and forward rates is given by the Extended Nelson–Siegel form (Svensson 1994). Introducing, for readability, the auxiliary functions

$$A(\theta_k) = exp(-\theta_k), \quad B(\theta_k) = \frac{1 - exp(-\theta_k)}{\theta_k}, \tag{7}$$

the fitting function that represents the instantaneous forward rate for the time θ is defined as:

$$F^c(0, \theta) = \beta_0 + \beta_1 A(\theta_1) + \beta_2 \theta_1 A(\theta_2) + \beta_3 \theta_2 A(\theta_3) \tag{8}$$

Here, $\theta_k = \theta / \tau_k$ is the scaled time, and the parameters are two groups: scaling parameters τ_1, τ_2, and shape parameters, $\beta_k, k = 0:3$. The implementation of the method involves fixing a priori the scaling parameters and obtaining the shape parameters through optimization.

The continuously compounded spot rate $R^c(t,T)$ at trade time t and maturity time T can be interpreted as the average of the instantaneous forward rates with settlements between time t and T, hence

$$R^c(t,T) = \frac{1}{T-t}\int_t^T F^c(t,\theta)d\theta. \tag{9}$$

The spot rate is derived from Eq.(8) using Eq.(9) as follows:

$$\begin{aligned} R^c(0,\theta) = \beta_0 + \beta_1 B(\theta_1) + \beta_2[B(\theta_1) - A(\theta_2)] \\ + \beta_3[B(\theta_2) - A(\theta_2)].. \end{aligned} \tag{10}$$

One problem with using the discounted cash flow method for project valuation is that the method is essentially static. First, it assumes that the state of the world remains unchanged during the life of the project and ignores the uncertainty of parameters, such as time and the size of the cash flows. More importantly, the DCF method assumes passive commitment to the project at the inception and ignores the ability of the management to make decisions that adjust project parameters in response to changing market conditions. These issues can be addressed by the *real options method* (Ronn 2003). In this work, we use only the traditional DCF measures, NPV and IRR, because of their simplicity and ease of communication with decision makers.

3. MOTIVATION FOR THE STUDY

The motivation for this study lies in the idea of using small scale power generating units as commercial electrical generators. Typically such PGUs will be part of the Emergency (standby, backup) Power Systems (EPS), alternate reserve sources of electric power that provide energy to critical loads in case of power outages.

Depending on the load, power sources for EPS can be batteries, uninterruptible power systems (UPSs) or PGUs, also called *gensets*. PGUs are electric power generators driven by engines or turbines. Most widely used are diesel engine generator sets with sizes in the range from 100 *kW* to 2,000 *kW* (Kusko 1989). Because PGUs vary in both size and efficiency, we take as a representative example a mid-range genset manufactured by SDMO with a Volvo Penta engine yielding 500 *kW* stand-by power. The fuel consumption at this level is 122.1*L*/*h*. Although the efficiency slightly increases at lower levels of utilization, the difference does not justify a separate treatment in this case.

The requirements for the design and implementation of EPS's are specified in the national standards and codes of good practices (NFPA 2005). One of the requirements of the National Fire Protection Association (NFPA) limits the capacity of unenclosed day tanks to 660 *gal* (2,498*L*) (NFPA 1998). In accordance with this rule, we assume that the genset for our example is supplied with a tank with a capacity of 2,442 *L*, which is equivalent to 20 hours of uninterrupted work. At the end of the period, the tank is refilled at the average retail price of diesel at the time.

4. MODELING ENERGY PRICES

The two commodities involved in this project valuation are diesel and electricity. As with any commodity, their prices are determined by the interplay of the forces of supply and demand, but both retail diesel and electricity exhibit quite specific characteristics that affect the price formation.

4.1. Characteristics of Electricity Prices

4.1.1. Electricity Supply

In the deregulated electricity market design, supply meets demand on a short-term spot market that is organized either in the form of a pool or as a power exchange. In both cases, the price formation is realized via a bidding process for each hour or half-hour of the following day 0.

Electricity has some features that make it quite distinct from other commodities. Electricity is a non-storable commodity and the matching of supply and demand has to be done in real time. Both the supply and demand are subject to unexpected shocks from weather changes, congestion, unit failures, etc. Demand fluctuations are exacerbated by the fact that the deregulated producers have no incentives to keep excess capacity idle, except for old and inefficient generating facilities with high production costs.

One additional factor that affects the supply is the electricity transportation. Electricity is transported in a single transmission network, usually referred to as the *grid,* that in many cases is geographically localized. The lack of alternative transportation channels increases the risk of failure and introduces an upper limit of the transmission capacity.

4.1.2. Electricity Demand

The demand side is characterized by recurring patterns over multiple time scales. The longest time scale is the one that reflects the seasonal changes in the specific geographic region (including demands related to heating and cooling). The second time scale deals with weekly variations due to industrial activity and thirdly, there are intraday patterns of variation.

The intraday variability is reflected in the terminology used by the exchanges. The 24 hour day is divided into three time periods: one *on-peak* and two *off-peak* periods. The on-peak period encompasses the time of the day with higher electricity demand and its boundaries vary with the geographic region. For Alberta, on-peak is defined as the period from HE 9 to HE 21 inclusive, Mountain Prevailing Time during a business day. Here "HE" (hour ending) refers to the 60-minute period ending that hour. For example HE 21 includes the time interval between 20:00 and 21:00. The

average price over the whole 24-hour daily interval is referred to as the *baseload*. Figure 1 illustrates the weekly variability of the three different time periods – on-peak load, off-peak load, and baseload – for 2005. The different time periods are preserved even for the prices over the weekends and holidays, denoted by 'Hol'.

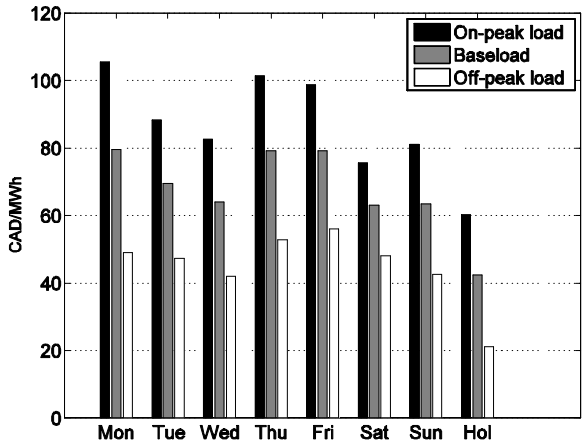


Figure 1: Average weekday and holiday prices for 2005.

Because the baseload is, in fact, the average of the off– and on–peak period prices we can use it as a representative of the daily price variations. Figure 2 shows the typical pattern of baseload price changes for Alberta.

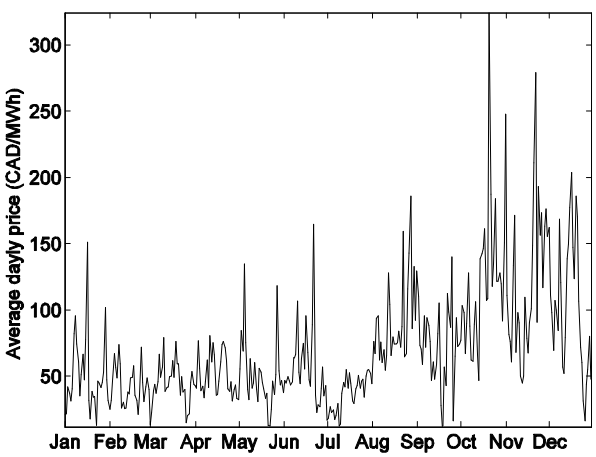


Figure 2: Baseload electricity prices for Alberta, 2005

In engineering terms, the realized current demand is referred to as the *load* on the system and represents the total amount of electricity consumed by the customers at any given moment. There are two main categories of load: base load and peak load. This differentiation attains importance because of its relation to the corresponding types of generating stations. *Base load* is the minimum amount of electricity that is pulled out of the grid by the consumers. This type of demand is covered by nuclear and hydro-power units that typically have lower marginal production costs, but also longer dispatch and response times. *Peak load* is the maximum amount of electricity demanded, usually understood as the marginal difference over and above the normal base load. It is typically covered by oil and gas power plants.

The demand is usually categorized as short-run and long-run. The *short-run demand* refers to time scales during which there is no change in the existing loads, such as residential appliances, industrial infrastructure, technological processes, etc. Currently, it is accepted that the short-run demand for electricity is income and price inelastic (Lafferty et al. 2001).

Ordering the generating units by their size and marginal production cost yields the supply stack. In the short-run, peak level electricity demand can only be matched by the remaining generating units from the far right side of the stack. Since these units are typically inefficient and have high marginal costs, the supply also becomes inelastic.

4.1.3. Features of Electricity Prices

The system exhibits significant uncertainty due to unexpected weather changes, generation failures, transmission congestion, etc. The combination of inelastic demand with inelastic supply in an environment with high levels of uncertainty determines the following unique features of electricity prices:

- **Cyclical seasonality patterns** on three different time scales – yearly, weekly, and daily – driven by the regular variation of the demand,
- **Random price spikes** – sudden jumps in prices relative to the average level for the season, due to unanticipated shocks in either demand or supply, e.g. unusually hot weather or congested transmission,
- **Mean reversion** – the tendency of the prices to revert to a dependent on the season average price level.

Figure 2 illustrates both the mean reversion and the price spikes, as well as the yearly seasonal dependence. Figure 3 clearly shows both the weekly patterns and the difference in the peaks intensity over two different representative months: May and November.

4.2. Alberta Energy Market

The design of a typical electrical system involves coordination between the independent power producers and importers on the supply side and the retailers and exporters on the demand side, including the interests of the owners of the transmission facilities and the owners of the distribution system. All transactions are channelled through the Power Pool operated by Alberta Electric System Operator (AESO), a statutory corporation. It has about 200 participants and over $7 billion in annual energy transactions. The Pool's Operating Policies and Procedures (AESCO 2006) require pool participants to submit their hourly bids and offers for the next seven days every day before noon. Distributors and exporters submit bids specifying the

amount of energy they are willing to buy at specified prices. Power suppliers submit offers for blocks of power at prices they are willing to accept. The quotes are binding only for the next trading day. All valid bids and offers for a settlement interval are sorted in ascending order of how expensive they are and generation and load units are dispatched according to this merit order. There is a price cap currently set at *CAD*$1,000*MWh*. The *System Marginal Price* (SMP), i.e. the price of the last block of energy dispatched to meet the load, is updated every minute in real time. The ex-post time-weighted average of the SMPs gives the market pool price for that hour.

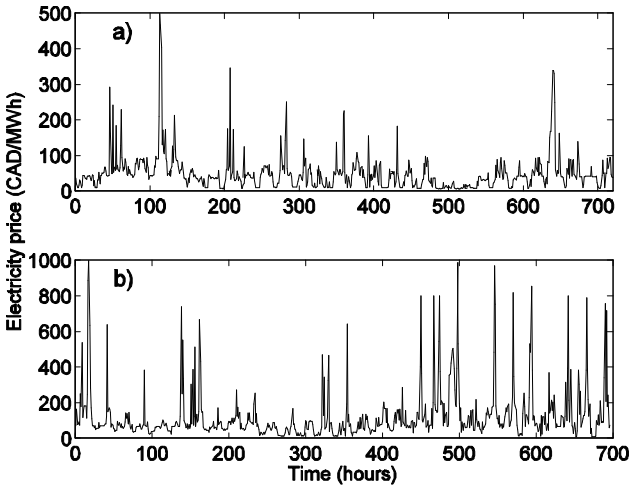


Figure 3: Hourly electricity prices for Alberta for two periods of 2005, a) May 01 – Jun 01, and b) Nov 01 – Dec 01.

4.3. Price Models

The taxonomy of electricity price models includes various approaches broadly classified in six groups depending on the mathematical methods used for process description, namely: time series models, game theory models, structural models, non-parametric models, statistical models, and production cost models (Gonzalez et al. 2005). For the purposes of the Monte Carlo simulations, the most appropriate models are the ones based on financial time series methods.

Modern finance theory postulates that price follows specific random process X_t and uses the methods of stochastic calculus for its description (see e.g. Bjork 2004). Here the subscript, *t*, shows time dependence. The main building block in the theory of stochastic processes is the (standard) Brownian motion (BM) and its mathematical formulation, *Wiener process*, W_t. BM is the continuous limit of the random walk and is the basic model of the effect of continuous noise over time *t*. The process has independent increments, and its change, dW_t, over a short period of time Δt is normally a distributed random variable with zero mean and a variance proportional to the length of the time interval, i.e.

$$W_{t+\Delta t} - W_t \sim N(0, \sqrt{\Delta t}). \tag{11}$$

If we denote the random draws from the standardized normal distribution $N(0,1)$ as ε_t the discrete version of the increments of the Wiener process can be written as $\Delta W_t = \varepsilon_t \Delta t$, which is the expression used in the Monte Carlo simulations.

The simplest non-trivial generalization of the Wiener process is the arithmetic Brownian motion, which has a deterministic term with *drift*, ν, and a random term that is multiplied by a scale factor, σ, that has the meaning of standard deviation of the process

$$dX_t = \nu dt + \sigma dW_t. \tag{12}$$

This equation already allows for the modeling of the variable drift and variance and is appealing due to its simplicity, but because the random variable X_t can take negative values it cannot be interpreted as price. The simplest solution is to make the instantaneous mean and standard deviation proportional to the random variable, S_t, formulating, in fact, the most popular model for description of price evolution, the *geometric Brownian motion* (GBM)

$$dS_t = \mu S_t dt + \sigma S_t dW_t. \tag{13}$$

Using Ito's lemma, it is quite easy to show 0 and Pindyck 1994) that this is equivalent to setting $X_t = \ln S_t$ and defining the drift as $\nu = \mu - \sigma^2/2$. Thus the distribution of the price return over a time interval Δt is normal with mean $\mu\Delta t$ and variance $\sigma^2 \Delta t$

$$X_{t+\Delta t} - X_t \sim N(\nu\Delta t, \sigma^2 \Delta t) \tag{14}$$

i.e. the random variable S_t from Eq.13 can be interpreted as (log-normally distributed) price.

GBM is commonly accepted as a model for stock price, but in its basic form it cannot adequately describe random variables with mean reverting behaviour, such as electricity prices. The first mean-reverting models were developed to describe the *term structure of interest rates*, i.e. the interest rate yield curve for different times to maturity (Vasicek 1977, Hull and White 1990). The simplest approach is to modify the drift term of the basic model Eq.13 as follows:

$$dS_t = \lambda(\alpha - S_t)dt + \sigma dW_t. \tag{15}$$

The first term of this equation changes its sign depending on the difference between the magnitude of the price, S_t and the long-term mean-reversion level α thus ensuring reversion towards this level with speed proportional to λ. The second term is the usual

normally distributed stochastic term. This one-factor model is known as the the Vasicek model, or for the special case of zero level mean reversion, i.e. $\alpha = 0$ as the Ornstein-Uhlenbeck process. Still, it does not have a mechanism to describe price seasonality.

Price seasonality is usually modeled by adding a deterministic term to the mean-reverting part of the model, S_t^{MR} as

$$S_t = S_t^{MR} + f_t. \tag{16}$$

The most common choice of the deterministic function f_t is some combination of sinusoids (Roncoroni and Geman 2003), sometimes with the addition of deterministic linear trend and/or terms to account for the seasonalities at different time scales (Lucia and Schwartz 2002). One fairly general form of the deterministic function is (Escribano et al. 2002)

$$f_t = \alpha + \beta t + \gamma D_t + \sum_n \zeta_n sin\left((t + \theta_n) \frac{2\pi n}{365} \right), \tag{17}$$

for a set of constant parameters $\alpha, \beta, \gamma, \zeta_n, \theta_n$. Usually the sum is restricted to only two terms, which is enough to describe two annual maxima. The term γD_t is sometimes used to capture the day-of-week variability. With this adjustment for the existence of deterministic periodicity the stochastic differential equation describing the process takes the form

$$dS_t = \lambda \left(\alpha_t - S_t \right) dt + \sigma dW_t, \tag{18}$$

with the deterministic function α_t defined as

$$\alpha_t = f_t + f_t' / \lambda. \tag{19}$$

A comparison with Equation 15 shows that the model specified by Equation 18 can be regarded as an extended Vasicek model (Hull and White 1990) with a time dependent mean-reverting level given by the deterministic function α_t.

None of the models considered so far exhibit spikes, which are so prominent in the Alberta electricity market. The most common approach is the formulation of a jump-diffusion model (Eydeland and Geman 1999) by adding a jump component $J_t dq_t$ to Equation 18, with jump size J_t and intensity given by q_t. The problem with these models is that the duration of the electricity spikes is rarely longer than one hour, while jumps have much longer durations. Different approaches of forcing the jump back (Roncoroni and Geman 2003) increase model complexity and introduce statistical distortions. Therefore, an alternative set of approaches has been tested. In the hidden Markov models (HMM) an approach towards the price process is split in two regimes: the stable regime and the spike regime, and the spot price switches between the regimes (Gonzalez et al. 2005). The drawback of HMM is the difficulty of incorporating the seasonality. Yet another approach involves direct modeling of the supply and demand processes (Eydeland and Geman 1999, Burger at al. 2004).

For the purposes of this study, where the importance of the price spikes on hourly resolution scale is crucial, the model must incorporate the features of both the base price (Equation 18), and the frequency and intensity of the spikes.

5. EXTREME VALUE THEORY

Extreme Value Theory (EVT) addresses the behavior of stochastic processes that exhibit heavy tails. The classical EVT is based on the asymptotic Generalized Extreme Value (GEV) function that models the distribution of maxima over a specified period. An alternative approach, called 'peaks over threshold' (POT), models the observations that exceed a high threshold, the threshold exceedances. For a sequence $X_1, X_2, \cdots$ of independent and identically distributed random variables with an unknown underlying distribution function $F(x) = P\{X_i \le x\}$ we define the distribution of the excesses over a high threshold u as the conditional probability:

$$F_u(x) = P\{X - u \le x \mid X > u\} \tag{20}$$

It can be observed that the excess distribution can be written in terms of the underlying distribution F as:

$$F_u(x) = \frac{F(x+u) - F(u)}{1 - F(u)}, \tag{21}$$

which implies that if the underlying distribution is known, then its excess distribution is easily computed. It has been shown (Balkema and Haan 1974) that for a wide class of underlying distributions and sufficiently high threshold the excess distribution F_u can be approximated by the generalized Pareto distribution (GPD) given by:

$$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi x / \beta)^{-1/\xi} & \xi \ne 0, \\ 1 - \exp(-x/\beta) & \xi = 0, \end{cases} \tag{22}$$

with a scale parameter $\beta > 0$. The value of the shape parameter ξ determines the particular type of distribution function subsumed in GPD. When $\xi = 0$, the GPD is the exponential distribution; when $\xi < 0$, it is known as Pareto type II distribution, and when $\xi > 0$, it is the ordinary Pareto distribution, which has a long history in actuarial mathematics as a model for large losses.

5.1. Parameter estimation

In order to take advantage of the extreme values

distribution tail approximation

$$F_u(x) \approx G_{\xi,\beta}(x) \tag{23}$$

we need to estimate the parameters ξ and β for a properly chosen threshold u. The most common technique for parameter estimation is the Maximum Likelihood Estimation (MLE) method which is part of most statistical software packages. Figure 4 shows the MLE results for shape parameter ξ and scale parameter β of the GPD, with the 95% confidence intervals for different threshold values.

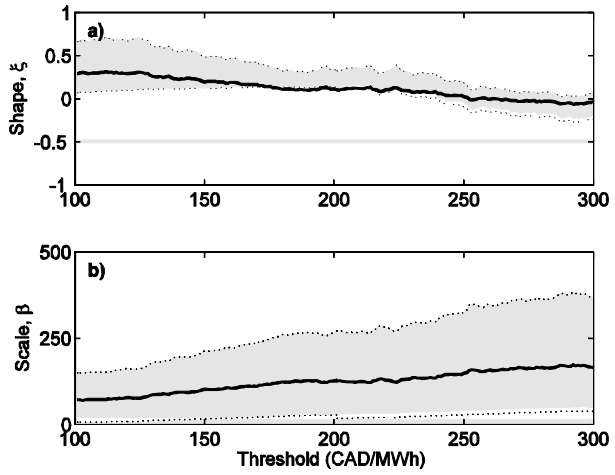


Figure 4: Maximum likelihood estimates of the parameters of GPD, a) shape parameter ξ and b) scale parameter β with 95% confidence intervals for different threshold values.

The choice of an appropriate threshold value is more difficult. On the one side, the threshold has to be high enough for the approximation of Equation 23 to hold, but on the other, we need a large enough number of exceedances to ensure good statistics. In addition to the QQ plots another useful tool is the conditional *sample mean excess* (SME) function defined as

$$e(u) = E(X - u \mid X > u), \quad u \geq 0, \tag{24}$$

An important property of GPD for values $\xi > -1$ is that SME is a linear function of the threshold, u

$$e(u) = \frac{\beta}{1+\xi} - \frac{\xi}{1+\xi} u. \tag{25}$$

SME can be empirically estimated as the total distance of exceedance over the threshold divided by the number of points exceeding that threshold, i.e.

$$e_n(u) = \sum_{k=1}^{N} (X_k - u)^+ \Big/ \sum_{k=1}^{N} I_{\{X_k > u\}}, \tag{26}$$

where n is the sample size, the '+' sign denotes the positive part of the expression in the brackets, i.e. $A^+ = max(A, 0)$, and the indicator function $I = 1$ if $X_k > 0$ and 0, otherwise. The empirical estimate of the SME, Equation 26 can be used to infer an appropriate threshold value by identifying a linear part of the function (Embrechts et al. 2004) according to the linearity property, Equation 25. Figure 5 shows the variation of the SME function on the electricity data for 2005.

At threshold, u=130 CAD/MWh, the maximum likelihood estimates for the shape and scale parameters of the GPD are $(\hat{\xi}, \hat{\beta}) = (0.257,\ 85.47)$ with confidence intervals: $CI(\hat{\xi}_{95\%}) = (0.167,\ 0.346)$ and $CI(\hat{\beta}_{95\%}) = (76.65,\ 95.31)$.

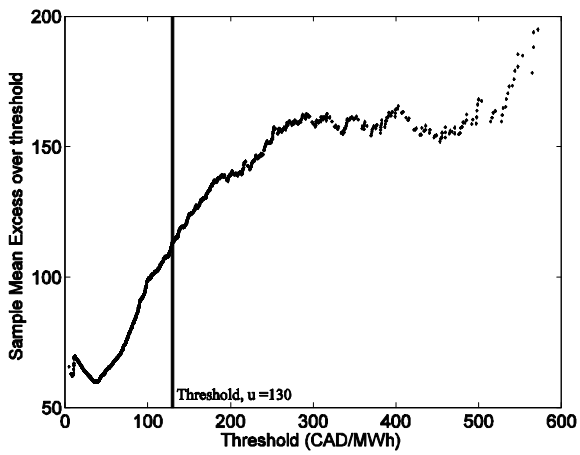


Figure 5: Sample Mean Excess function over threshold for hourly Alberta electricity prices for 2005

The graphical comparison between the GPD calculated with these parameters and the empirical histogram created on the basis of the same tail data is presented in Figure 6. Visual inspection confirms the goodness of fit for the chosen parameter.

In addition to the rather subjective visual inspection we applied the more formal goodness of fit test. We formulated the null hypothesis that the empirical tail distribution is a sample coming from the theoretical GPD against the alternative that it does not come from that distribution. We applied two-sample Kolmogorov–Smirnov tests to the data. Bearing in mind the limited power of the test (Choulakian and Stephens 2001, Zempleni 2004), it also confirmed the null hypothesis at 95% confidence level with p–value of $p \approx 0$.

Figure 7 shows the position of the threshold that separates the extreme values of the hourly electricity prices according to EVT. There are 937 extreme points over the threshold out of total of 8,471 observations, i.e. 11%.

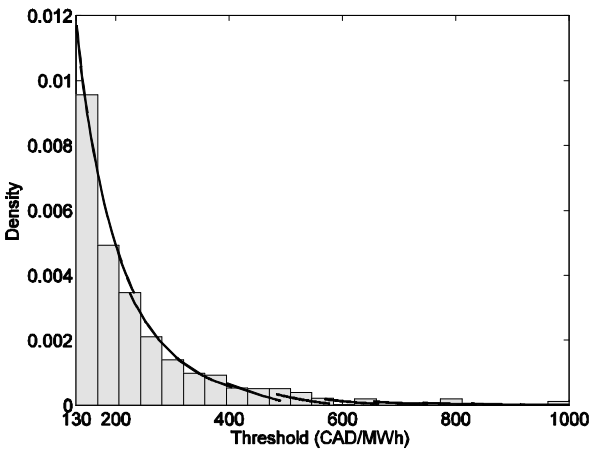


Figure 6: Diagnostic plot of GPD probability density and the empirical histogram for the values of Alberta hourly electricity prices for 2005 over the threshold u=130.

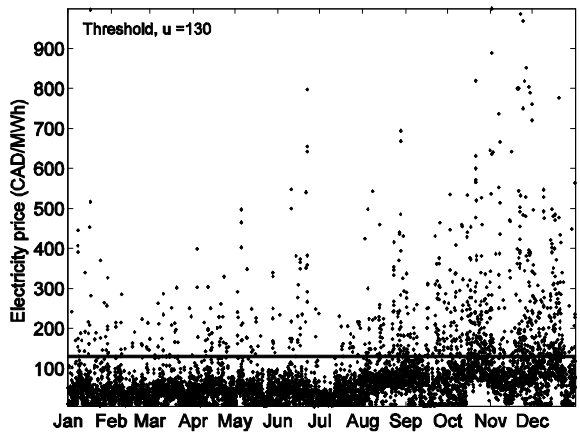


Figure 7: Threshold for the extreme value theory plotted against the hourly prices for 2005

6. MONTE CARLO SIMULATION

6.1. Algorithms

In this study, we apply Monte Carlo simulation within the framework of capital budgeting. The problem involves three sources of uncertainty, the price of fuel, the price of electricity, and the level of interest rate. The prices are modeled as stochastic processes. The hourly price of fuel is obtained by sequentially applying the discretized version of Equation 18. The hourly price of electricity is obtained from the same equation with the added 'spike' term.

The simulation is run N_r number of times. Each time, an hourly comparison of the simulated prices of fuel and electricity is performed. If the price of electricity is higher than the price of fuel used to produce it, the spark spread, the power generating unit is 'dispatched' for that hour. If the spark spread is negative, no energy is sold to the grid. The cash flows generated during the day are used to perform a DCF analysis of the investment.

The interest rate used in the DCF analysis, the third source of uncertainty, is modeled by fitting the extended Nelson-Siegel form to the data of tradable Canada Government bonds, as described below. The spot rate, Equation 10, is used to calculate the present value of the cash flow for every day, $PV(CF_t)$. The sum of all cash flows is used to determine the net present value for the particular realization NPV_r. In parallel, the forward rate, Equation 8, is used to calculate the future value of the cash flow for every day, $FV(CF_t)$ the sum of which is used to determine the total future value for the particular realization FV_r^{tot}. The latter is plugged into Equation 4 in order to calculate the internal rate of return from the asset for the particular realization IRR_r. The values of both NPV_r and IRR_r are retained at the end of each realization, thus at the end of the simulation we can calculate their average values (point estimators for the population means)

$$\hat{V} = \frac{1}{N_r}\sum_{r=1}^{N_r} V_r, \tag{27}$$

where V stands for both NPV and IRR. The standard error of the estimator $\hat{V}$ is calculated as (Hines et al. 2003)

$$\varepsilon(\hat{V}) = \sigma(\hat{V}) / \sqrt{N_r}, \tag{28}$$

where the sample standard deviation is given by

$$\sigma(\hat{V}) = \sqrt{\frac{1}{N_r - 1}\sum_{r=1}^{N_r}\left(V_r - \hat{V}\right)^2} \tag{29}$$

6.2. Data

The data for the electricity and fuel prices in Edmonton covers the period of 2001–2006, i.e. the period starting after the full deregulation of the electricity market in Alberta. The electricity data has an hourly frequency and some representative samples of the data are shown in Figure 2 and Figure 3.

The diesel price data refers to the weekly average prices at Exxon Mobil gas stations. Because the price differentials between the gas stations of the different oil companies is typically in the range of 1–2 cents, we take these prices as representative for the city. The price levels for the period are shown in Figure 8.

These weekly diesel prices were further re-sampled by linear interpolation in order to obtain hourly values. The low weekly variation allows obtaining the hourly price data by linearly interpolating between the weekly data points.

The data used to fit the extended Nelson-Siegel model consists of four Government of Canada Treasury bills with maturities of 1, 3, 6 and 12 months, and five selected Government of Canada benchmark bonds with maturities of 2, 5, 7, 10, and 30 years, summarized in Table 1.

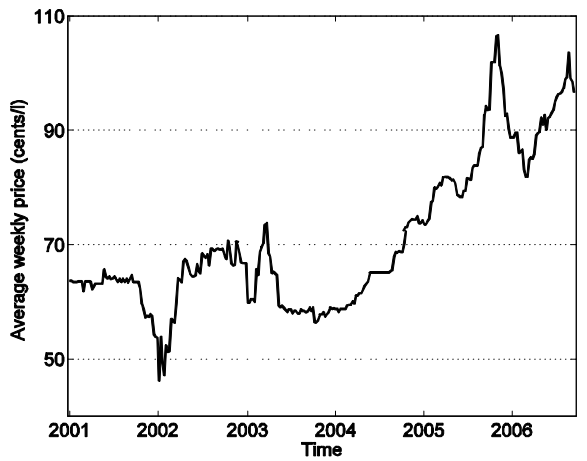


Figure 8: Weekly average retail prices of diesel in Edmonton for 2001-2006.

Table 1: Selected Government of Canada T-bills and benchmark bonds used to construct the term structure of the interest rates. The yields are based on mid-market closing values.

Maturity (years)	Yield (% per anum)	Coupon (% per anum)
0.08	4.1	0
0.25	4.15	0
0.50	4.12	0
1	4.00	0
2	3.785	4.25
3	3.76	4.25
5	3.745	3.75
7	3.79	5.25
10	3.86	4.00
30	3.96	5.75

6.3. Calibration

Diesel prices are modeled as a mean-reverting process described by Equation 18. In order to estimate the parameters of this model from data we write the equation in discretized form as

$$X_{k+1} - X_k \lambda(\alpha - X_k)\Delta t + \sigma\varepsilon_k . \tag{30}$$

where the stochastic term, $\varepsilon_k \sim X(0,1)$ are draws from standard normal distribution. This equation is recast as a linear function of the price levels:

$$X_{k+1} - X_k = a_0 + a_1 X_k + \sigma\varepsilon_k , \tag{31}$$

and the parameters $a_0 = \lambda\alpha\Delta t$, and $a_0 = -\lambda\Delta t$, a_1=-λΔt, are estimated by the least squares method. Using the log-prices data, the estimation yields $a_0 = 0.0419 \pm 0.0711$, and $a_0 = -0.0088 \pm 0.0163$. From here we determine the following parameters for the fuel model $\lambda_F = 0.4593$, $\alpha_F = 4.7393$, and $\sigma_F = 0.125$.

The calibration of electricity prices is similar to the calibration of the diesel prices, but involves additional steps because of the presence of spikes. Using the threshold u=4.8675, corresponding to logarithm of 130/MWh, as determined in the previous section, we separate the 'normal' regime from the spike regime. The latter is modeled as a product of two distributions, one describing the frequency of the spikes, and the other describing their intensity. The frequency component was modeled with the Poisson distribution, and the intensity component was modeled with the Gamma distribution. The maximum likelihood estimates for the shape parameter, α_G and the scale parameter λ_G of the Gamma distribution are $(\hat{\alpha}_G, \hat{\lambda}_G) = (89.2453, 0.0576)$ with confidence intervals: $\hat{\alpha}_{G|95\%} = (85.1712, 93.5142)$ and $\hat{\lambda}_{G|95\%} = (0.0549, 0.0603)$ Similarly, for the Poisson distribution we obtain $\hat{\lambda}_P = 5.1392$ with values at 95% confidence interval $\hat{\lambda}_{P|95\%} = (5.0642, 5.2142)$.

The fitting of the term structure of interest rate, Equation 10 , to the data in

Table 1 was formulated as a minimization problem (Soderlind and Svensson 1997)

$$\min_{(\beta,\tau)\in R^4\times R^2} \sum_{k=1}^{N_b} \left(R(t_k) - \hat{R}(t_k) \right)^2 , \tag{32}$$

where the squared price deviations are calculated for all bonds, $N_b = 10$, and the corresponding times to maturity t_k, subject to

$$\beta_0 > 0, \quad \beta_0 + \beta_1 > 0, \quad \tau_{1,2} > 0. \tag{33}$$

The solution of this nonlinear minimization problem is quite sensitive to the starting point. We used a Matlab implementation of the Nelder-Mead simplex method, and our choice of the starting points was guided by the interpretation of the parameters (β, τ). In practice, β_0 is interpreted as a long-term interest rate, β_1 is the spread between long- and short-term interest rates, β_2 and β_3 determine the height and the directions of the curvatures, and the scale parameters τ_1 and τ_2 determine curvature positions. The values of these parameters as determined from the optimization are shown in

Table 1. The continuous term structure of the spot and interest rates, calculated from formulas (Equation 8 and Equation 10) using the parameters from Table 2 are graphed in Figure 9.

Following the algorithm described in the previous subsection, we simulated $N_p = 500$, the number of

paths for both the electricity prices and fuel prices for a period of 15 years. At each path, we calculated the NPV and the IRR of the cash-flows for positive price spread. Assuming an initial investment (hook-up cost) of $C_0 = \$25,000$ and no extra maintenance cost, the final estimates are: $NPV = \$605,686$ and $NPV = 248\%$. The standard error in the two estimates are respectively $SE(NPV) = 27\%$, and $SE(IRR) = 29\%$. One comment about the high value of IRR is that the calculation does not take into account the cost of the generation unit, only the part related to its connection to the grid. As noted before, such generators would already be in place as part of EPS. In addition, the value of the NPV is somewhat inflated and can only be used as a benchmark. The reason for this is the discount factor, which we assumed to be the risk free rate. For any other required rate of return this value would be smaller.

Table 2: Estimated extended Nelson-Siegel parameters used to calculate the continuous term structure interest rates curve

β_0	β_1	β_2	β_3	τ_1	τ_2
0.0454	-0.0044	-0.0149	-0.0145	2.5232	25.6014

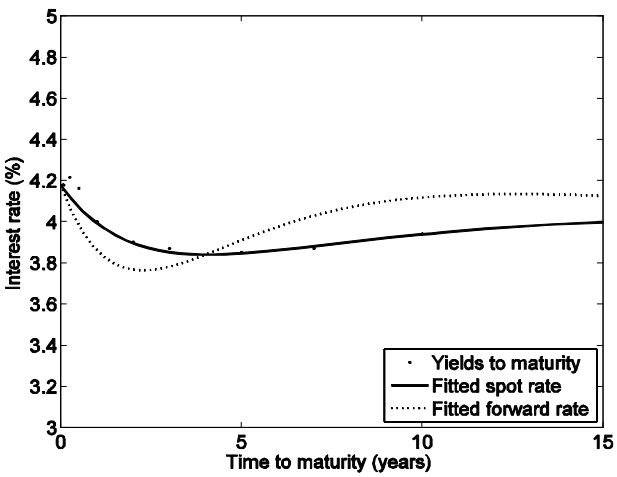


Figure 9: Extended Nelson-Siegel fit for the estimated duration of the project

7. CONCLUSION

In order to estimate the potential profitability of independent power producers, we performed the Monte Carlo simulations calculating two commonly accepted measures in the capital budgeting profitability: the net present value and the internal rate of return. These two measures were calculated on the basis of the cash flows generated by the power asset. A cash flow would be generated when the spread between the prices of the energy sold and the price of the fuel to produce this energy would be positive.

Our approach differs from most valuations of power generating assets in several aspects. First, we modeled the prices of electricity and fuel separately. This allowed for the incorporation of realistic sampling of fuel prices, subject to fuel tank capacity constraints. The tank would be "refilled" at the current fuel prices when emptied and the time between re-fills would be dependent on the time interval where the spark spread is positive.

Secondly, we paid close attention to the distribution of the spikes in electricity prices. Small power generating units with high production cost have high sensitivity to the extreme values as a result of their high production cost. We used the 'peaks over threshold' method to fit the tail distribution of the electricity prices and model more accurately the price spikes.

Finally, in order to account for the changes of the interest rates over time and their effect on the profitability measures we modeled the term structure of the interest rates by the extended Nelson-Siegel form.

Our results for the parameters of the chosen representative power generation unit show the viability using units that are currently part of the emergency power systems for peak hour power producing.

REFERENCES

Alberta Electric System Operator (AES0), 2006. Independent System Operator Operating Policies and Procedures, October 2006.

Balkema, A.A., and de Haan, L., 1974. Residual life time at great age. *Annals of Probability*, 2(5):792–804.

Bjork, T., 2004. *Arbitrage Theory in Continuous Time*. 2nd edition, Oxford University Press.

Burger, M., Klar, B., Müller, A., and Schindlmayr, G., 2004. A spot market model for pricing derivatives in electricity markets. *Quantitative Finance*, 4:109–122.

Choulakian, V., and Stephens, M.A., 2001. Goodness-of-fit tests for the generalized Pareto distribution. *Technometrics*, 43(4):478–484.

Dixit, A.K., and Pindyck, R.S., 1994. *Investment under uncertainty*. Princeton University Press.

Embrechts, P., Klüppelberg, C., and Mikosch, T., 2004. *Modelling Extremal Events for Insurance and Finance*. Springer.

Escribano, A., Peña, J.I., and Villaplana, P., 2002. Modeling electricity prices: international evidence. Working paper 02-27, Departamento de Economía, Universidad Carlos III de Madrid.

Eydeland, A., and Geman, H., 1999. Fundamentals of electricity derivatives. In: *Energy Modelling and the Management of Uncertainty*, Risk Books. 35–43.

Geman, H., 2005. *Commodities and commodity derivatives: modeling and pricing for agriculturals, metals, and energy*. John Wiley & Sons,

González, A.M., San Roque, A.M., and García-González, J., 2005. Modeling and forecasting electricity prices with input/output Hidden Markov Models. *IEEE Transactions on Power Systems*, 20(1):13–24.

Hines, W.W., Montgomery, D.C., Goldsman, D.M., and Borror, C.M., 2003. *Probability and Statistics in Engineering*, 4th edition, Wiley.

Hull, J. and White, A., 1990. Pricing interest-rate-derivative securities. *The Review of Financial Studies*, 3(4):573–592.

Kusko, A., 1989. *Emergency Standby Power Systems*. McGraw-Hill.

Lafferty, R., Hunger, D., Ballard, J., Mahrenholz, G., Mead, D., and Bandera, D., 2001. Demand responsiveness in electricity markets. Technical report, NASEO, Office of Markets, Tariffs and Rates, Jan 2001.

Lucia, J.J., and Schwartz, E.S., 2002. Electricity prices and power derivatives: evidence from the Nordic power exchange. *Rev. Deriv. Res.*, 5:5–50.

Martellini, L., Priaulet, P., and Priaulet, S., 2003. *Fixed-Income Securities: Valuation, Risk Management and Portfolio Strategies*. John Wiley & Sons,.

National Fire Protection Association (NFPA), 1998.. Standard for the Installation and Use of Stationary Combustion Engines and Gas Turbines (37). NFPA,.

National Fire Protection Association (NFPA) and American National Standards Institute, 2005. National Electrical Code.

Ourdev, I., and AbouRizk, S.M., 2006. Internal rate of return for frequent cash flows. Working Paper.

Roncoroni, A., and Geman, H., 2003. *A class of marked point processes for modeling electricity prices*. Preprint ESSEC Business School, Cergy–Pontoise,

Ronn, E.I. (ed.), 2003. *Real Options and Energy Management: Using Options Methodology to Enhance Capital Budgeting Decisions*. Risk Books.

Söderlind, P., and Svensson, L., 1997. New techniques to extract market expectations from financial instruments. *Journal of Monetary Economics*, 40(2):383–429.

Svensson, L.E.O., 1994. Estimating and interpreting forward interest rates: Sweden 1992–1994. NBER Working Papers 4871, National Bureau of Economic Research, Inc.

Vasicek, O., 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188.

Zempléni, A., 2004. Goodness-of-fit test in extreme value applications. Discussion paper No. 383, SFB. 386, Statistische Analyse Diskreter Strukturen, TU Mn&&chen.

AUTHORS BIOGRAPHY

Ivan Ourdev, PhD, MBA, is a Research Engineer with the Hole School of Engineering at the University of Alberta. His recent work has included investigation of intelligent agents for simulation and advanced computational methods.

Simaan M. AbouRizk, PhD, PEng, is the NSERC Industrial Research Chair in Construction Engineering and Management and the Canada Research Chair in Operation Simulation at the University of Alberta. He is renowned in the academic construction community for his research in computer simulation and its applications in construction planning, productivity improvement, constructability reviews and risk analysis and was recently recognized with the ASCE Peurifoy Award.

Mohammed Al-Bataineh, PhD, is a Senior Analyst with S.M.A. Consulting in Edmonton, Alberta, Canada, where he is implementing simulation for risk review and project planning analysis. His recent research work has included scenario-based planning for tunnel construction.

MULTI AGENT/HLA ENTERPRISE INTEROPERABILITY (SHORT-LIVED ONTOLOGY BASED)

G. Zacharewicz(a), O. Labarthe(b), D. Chen(c), B. Vallespir(d)

[(a)(c)(d)] IMS-LAPS / GRAI, Université de Bordeaux – CNRS, 351, cours de la Libération, 33405 TALENCE cedex, France
[(b)] CIRRELT, Université Laval, 2325 rue de la Terrasse, bureau 2535, Québec city, Québec, Canada

[(a)]gregory.zacharewicz@ims-bordeaux.fr, [(b)]olivier.labarthe@cirrelt.ca, [(c)]david.chen@ims-bordeaux.fr, [(d)]bruno.vallespir@ims-bordeaux.fr

ABSTRACT
This paper aims at proposing an implementation of the Federation oriented Enterprise Interoperability concept, using Multi Agent / HLA paradigm and the rising notion of Short-Lived Ontology. We give first, a review of ongoing researches on Enterprise Interoperability. Then, we recall on Artificial Agent Concept and HLA Standard that appear to be adequate to support execution of the studied concept. Indeed, on the one hand Agent dialogue fits the concept of information exchange in a federated enterprise interoperability approach, on the other hand the HLA standard, initially designed for military M&S purpose, can be transposed for enterprise interoperability at the implementation level, reusing the years of experiences in distributed systems. From these postulates, we propose the first Agent/HLA framework Short-Lived Ontology based to implement distributed enterprise models from the conceptual level of federated enterprise interoperability approach.

Keywords: Enterprise Interoperability, Multi-Agent-Systems, HLA, Ontology

1. INTRODUCTION

In the globalised economic context, the competitiveness of an enterprise depends not only on its internal productivity and performance, but also on its skill to collaborate with others. This necessity led to the development of a new concept called interoperability that allows improving collaborations between enterprises. No doubt, in such context where more and more networked enterprises are developed; enterprise interoperability is seen as a more suitable solution to total enterprise integration. Since the beginning of 2000, several European research projects have been launched to develop enterprise interoperability (IDEAS, ATHENA, INTEROP). Three main research themes or domains that address interoperability issues were identified, namely: (1) Enterprise Modeling (EM) dealing with the representation of the internetworked organization to establish interoperability requirements; (2) Architecture & Platform (A&P) defining the implementation solution to achieve interoperability; (3) Ontologies (ON) addressing semantics necessary to assure interoperability (IDEAS 02).

This paper proposes a new contribution of Information Technology (IT) architecture and platform to implement Enterprise Interoperability. In the first part, we present the various approaches of interoperability and the current consideration of interoperability stated as conclusion of the Interop Network of Excellence (FP6, 508011) (Chen 07). Then, we recall the concepts of software Agent and the High Level Architecture (HLA), i.e. a standard for distributed simulation.

Next, from our experience, we propose to investigate three aspects of interoperability. The first concerns time management in Enterprise Interoperability; the dynamic aspect has to be tackled with sound techniques. The second aspect concerns the definition of Enterprise Ontologies; federated approach in interoperability requires a new definition of high-level standard (i.e. Ontology) for exchanged data. The last aspect concerns the privacy of data. Indeed, enterprises must manage confidentiality of data shared between entities; levels of rights on enterprise data must be defined.

Computer science Ontologies, Artificial Agent language, and Object/Interaction in HLA can give keys to two first considerations. As well, the experience coming from Information Systems (IS) and M&S programming can be studied to keep data safe to address third point. From these postulates, we specify a platform implementation using HLA and Software Agents' autonomous dialogue concepts, to the concern of distributed federated Enterprise Interoperability models.

2. BASIC CONCEPT OF INTEROPERABILITY

Enterprise Interoperability refers to the ability of interactions between enterprise systems. The interoperability is considered as significant if the interactions can take place at least at the three different levels: data, services and process, with a semantics defined in a given business context (IDEAS 02).

Interoperability extends beyond the boundaries of any single system, and involves at least two entities. Consequently establishing interoperability means to relate two systems together and remove incompatibilities. Incompatibility is the fundamental concept of interoperability. It is the obstacle to establish seamless interoperation. The concept 'incompatibility' has a broad sense and is not only limited to 'technical' aspect as usually

considered in software engineering, but also 'information' and 'organization', and concerns all levels of the enterprise (Chen 07).

Our goal is to tackle interoperability problems through the identification of barriers (incompatibilities) which prevent interoperability to happen. Basic concepts relating to Enterprise Interoperability are categorized into three main dimensions as described below.

2.1. Dimension of interoperability barriers

This dimension takes into account three categories of interoperability problems.

Conceptual barriers are related to the problems of syntactic and semantic of information to be exchanged. This category of barriers concerns the modeling at high levels of abstraction as well as the level of coding.

Organizational barriers are related to the definition of responsibilities and authority so that interoperability can take place under good conditions.

Technological barriers are related to the problem of use of information technologies. This category of barriers concerns the standards that are used to present, store, exchange, process, and communicate data through the use of computers.

2.2. Dimension of interoperability concerns

This dimension identifies various levels of enterprise where interoperability takes place. These levels are based on ATHENA Architecture.

The business level refers to working in a harmonized way at the levels of organization and company in spite of for example, the different modes of decision-making, methods of work, legislations, culture of the company and commercial approaches etc. so that business can be developed and shared between companies.

The process level aims at making various processes working together. A process defines a sequence of services according to a specific need of a considered company. Commonly, in a company, several processes run in interactions (serial or parallel). In the case of a networked enterprise, internal processes of two companies must be connected to create a common process.

The service level is concerned with identifying, composing, and making function together with various applications (designed and implemented independently) by solving the syntactic and semantic differences, as well as finding connections to various heterogeneous databases. The term `service' is not limited to computer-based applications but manual ones as well.

The data level refers to making different data models (hierarchical, relational, etc.) and different query languages working together. The interoperability of data is related to find and share information coming from heterogeneous bases, which can moreover reside on different machines with different operating systems and databases management systems.

2.3. Dimension of interoperability approaches

This dimension takes into consideration the three admitted approaches to develop interoperability as illustrated in figure 1.

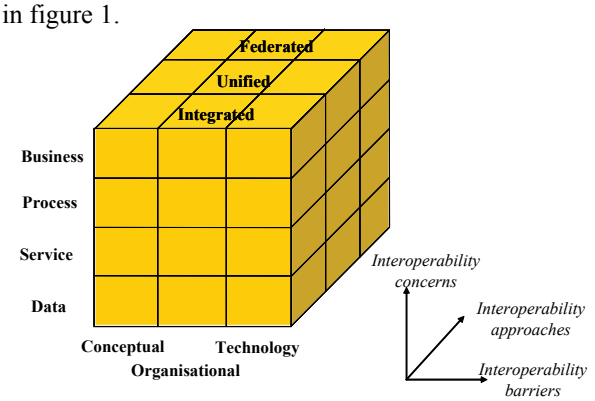


Figure 1: Framework for enterprise interoperability

Integrated approach: there exists a common format for all models. This format must be as detail as models. The common format is not necessarily a standard but must be agreed by all parties to elaborate models and build systems.

Unified approach: there exists a common format but only at a meta-level. This meta-model is not an executable entity as it is in the integrated approach but provides a mean for semantic equivalence to allow mapping between models.

Federated approach: there is no common format. To establish interoperability, parties must accommodate on the fly. Using federated approach implies that no partner imposes their models, languages and methods of work.

Today, most of the approaches developed are unified ones such as for example in the domain of enterprise modeling, we can mention UEML (Unified Enterprise Modeling Language UEML (2003) and PSL (Process Specification Language) which aim at supporting the interoperability between enterprise models and tools.

Using the federated approach to develop Enterprise Interoperability is most challenging and few activities have been performed in this direction. The federated approach aims to develop full interoperability and is particularly suitable for an inter-organizational environment (such as networked enterprises, virtual enterprises, etc.). In the Enterprise Interoperability roadmap published by the European Commission (IST 06), developing federated approach for interoperability is considered as a research challenge for the years to come.

3. HLA RECALLS

The High Level Architecture (HLA) is a software architecture specification that defines how to create a global software execution composed of distributed simulations and software applications. This standard was originally introduced by the Defense Modeling and Simulation Office (DMSO) of the US Department Of Defense (DOD). The original goal was reuse and interoperability of military applications, simulations and sensors.

3.1. HLA concepts

In HLA, every participating application is called federate. A federate interacts with other federates within a HLA federation, which is in fact a group of federates. The HLA set of definitions brought about the creation of the standard 1.3 in 1996, which evolved to HLA 1516 in 2000 (IEEE 00).

The interface specification of HLA describes how to communicate within the federation through the implementation of HLA specification: the Run Time Infrastructure (RTI).

Federates interact using services proposed by the RTI. They can notably "Publish" to inform about an intention to send information to the federation and "Subscribe" to reflect some information created and updated by other federates. The information exchanged in HLA is represented in the form of classical object class oriented programming. The two kinds of object exchanged in HLA are Object Class and Interaction Class. Object class contains object-oriented data shared in the federation that persists during the run time; Interaction class data are just sent and received information between federates. These objects are implemented within XML format. More details on RTI services and information distributed in HLA are presented in (IEEE 00).

In order to respect the temporal causality relations in the execution of distributed computerized applications; HLA proposes to use classical conservative or optimistic synchronization mechanisms (Fujimoto 00).

3.2. HLA Implementation Components

An HLA federation is composed of federates and a Run Time Infrastructure (RTI) (IEEE 2000), figure 2.

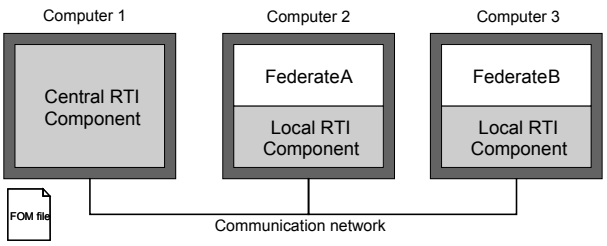


Figure 2: HLA Component Framework

A federate is a HLA compliant program, the code of that federate keeps its original features but must be extended by other functions to communicate with other members of the federation. These functions, contained in the HLA specified class code of *FederateAmbassador*, make interpretable by a local process the information received resulting from the federation. Therefore, the federate program code must inherit of *FederateAmbassador* to complete abstract methods defined in this class used to receive information from the RTI.

The RTI supplies services required by distributed executions, it routes messages exchanged between federates. It is composed of two parts.

The "Local RTI Components code" (LRC, e.g. in figure 2) supplies external features to the federate for using RTI call back services such as the handle of objects and the time management. The implementation is the class *RTIAmbassador*, this class transforms the data coming from the federate in an intelligible format for the federation. The federate program calls the functions of *RTIAmbassador* to send data to the federation or to ask information to the RTI. Each LRC contains two queues, a FIFO queue and a time stamp queue to store data before delivering to the federate.

Finally, the "Central RTI Component" (CRC, e.g. in figure 2) manages the federation notably by using the information supplied by the Federation Object Model (FOM) (IEEE 2000) to define Objects and Interactions classes participating in the federation.

A federate can, through the services proposed by the RTI, "*Publish*" and "*Subscribe*" to a class of shared data. "*Publish*" allows diffusing the creation of object instances and the update of the attributes of these instances. "*Subscribe*" is the intention of a federate to reflect attributes of certain classes published by other federates.

4. AGENT-BASED DISTRIBUTED SIMULATIONS

The Multi-Agents System (MAS) concentrates on the study of the collective behavior which results from the organization and interactions of agents for the resolution of problems. A MAS is a distributed system in which there is generally no centralized control or global point of view. A MAS is composed of agents which act in an autonomous way but do not locally have the knowledge, the resources or the information required to ensure the coherence of the concerted actions in a MAS. This section is dedicated to the presentation of the agent definition and the introduction of distributed simulations of Agent-Based Systems.

4.1. Agent Definition

Actually there is no consensus in the scientific literature on the definition of an agent. Disciplines in which reference is made are numerous and various authors have proposed different definitions as for example Ferber (1999) and Huang and Nof (2000). However, the definition proposed in Jennings *et al.* (1998) is commonly used within the MAS community: "*an agent is a computer system, situated in some environment that is capable of flexible autonomous action in order to meet its design objectives...*». In Wooldridge *et al.* (1995) the authors define the concept of an agent according to the following properties:

autonomy: an agent operates (task selection, decision-making, etc.) without human or other direct intervention and neither the actions it realizes nor its internal state are submitted to any intervention;

reactivity: an agent perceives its environment and reacts in an appropriate way to the environment changes;

pro-activeness: agents are able to act by taking initiatives driven by their goals;

social ability: agents are able to interact with other agents through communication language or social rules.

The importance accorded to the properties expressed above depends on the application needs and nature. Agents and MAS constitute an active research field in which numerous applications are developed. In Nwana (1996) and Wooldridge (2002) the authors propose surveys of the agents according to various application domains (cognitive agents, software agents, mobile agents, etc.). Agents perceive the modifications of their environment and perform actions on it. Among the possible actions, agents have to determine the most suited decisions that can reach their objectives. In addition to the application domain, environment, interaction and organization influence the design of the agent.

4.2. Distributed simulations of Agent-Based Systems

The properties which characterize MAS are particularly adapted for the modeling and simulation of distributed and dynamic systems. For the distributed simulation of multi-agent systems several tools and methodologies are defined in the literature. The following sub-section is dedicated to the presentation of research works focused on the integration of multi-agent systems through a HLA federation.

Lees *et al.* (2007) presented a tool named HLA_AGENT which integrates the SimAgent toolkit and the HLA. The SimAgent toolkit is used for the development of alternative agent architectures which can evolve in various environments. Concerning the organization of HLA_AGENT, a federate is associated to a SimAgent which is in relation with a Federate Ambassador (handles callbacks from the RTI) and a RTI ambassador (handles calls to the RTI).

Song *et al.* (2007) proposed a multi-agent data collection system to improve the performance of existing mechanism in HLA simulation systems. Each agent is connected to the RTI and is composed of four functional modules: Communication, Cooperative Decision, Data Logging and Data Processing. Furthermore, each agent is responsible of a Database composed of a subset of the simulation data.

Minson and Theodoropoulos (2008) introduced a middleware layer named HLA_REPAST to facilitate the integration between HLA and the sequential MAS simulation toolkit Repast. The authors supplied a complete description of the necessary steps associated to the creation of a federation of interacting instances of RePast models within HLA.

Chen *et al.* (2008) was interested of the data accessing problem to optimize the execution of distributed simulation of agent-based systems. In this context two routing approaches was proposed. The first named range-based attribute locate data according to a set of attribute value range, and the address-based approach locate a particular state variable from a given identifier. In order to study the dynamic of the two approaches the authors used the PDES-MAS framework proposed in Logan and Theodoropoulos (2001).

Cicirelli *et al.* (2009) proposed to interface a set of agents (grouped under the name of Theatre) to the HLA/RTI via two components. The TransportLayer component is connected to an RTI Ambassador for transmission and reception of messages. The ControlMachine component is connected to a Theater Ambassador for the management of the Theatre logical time and the external received simulation messages.

5. PERSPECTIVES TO INTEROPERABILITY

5.1. Removing Barriers to Interoperability

From the state-of-the-art of federated enterprise interoperability and implementations experiences presented in § 2, we can define several directions for, almost natural, interoperability barrier removal with Agent and HLA concepts in the following domains.

The first direction concerns the definition of commonly recognized paradigms and data structure able to evolve during run time.

The second not addressed requirement at the enterprise modeling level is the data synchronization. The data exchanged order is crucial; ignoring this can lead to not desired indeterminist model behavior.

Finally the enterprise modeling must consider the confidentiality management of data. The interoperability can be considered between concurrent enterprises in that context, a strategy of data sharing/not sharing between these must be defined. We present, in the following, propositions to address these requirements.

5.2. Enterprise Model Transformation Methodology for Distributed Execution

From the postulate that different enterprise models implementation using HLA, each of them follows its own development cycle. (Zacharewicz *et al* 09) introduced a common methodology by converging HLA FEDEP (IEEE 2000), MDA (OMG 2003) and MDI (Bourey 2007) steps, to clarify and rationalize the implementation method and the models. This life cycle proposed to standardize the steps to reach simulation or implementation from a conceptual enterprise model. This formalization will help reuse of development knowledge and will give a common metric to compare solution developments.

Phase 1: In first step, the objectives of the federation of enterprises have to be defined. Basically, the common goal of all federation created by this methodology consists in defining federation of interoperating enterprise models. This representation can naturally use the typical enterprise relation model elaborated in the enterprise model interoperability field (e.g. UEML models (2003)).

Phase 2: As second step, the mapping of enterprise models into HLA federates is realized. In detail, the way models handle received information and how they send information to the federation is addressed, these mechanisms can conform to synchronization algorithm proposed in (Zacharewicz 2008a). We pay here attention to reuse already existing enterprise models. Not existing enterprise models federates are created. In addi-

tion, we define information to exchange, i.e. what is the structure of distributed ontology and messages.

Phase 3: In the third step, the methodology maps enterprise interoperating connections between models into HLA interactions and objects. Then, these data are structured to generate the associated FOM. The strategy concerning confidentiality of data is also explicitly addressed in that step. Besides, to respect time constraints, objects and interactions among enterprise federates are time stamped related to local logical time of supplier enterprise federate to be handled and delivered right in time by RTI.

Phase 4: The federation is executed. The results obtained by simulation are used for validation of the models by test and analyze. In case it does not fulfill the specification, the methodology allows feedback correction as described in the FEDEP last step.

6. AGENT FEDERATED ENTERPRISE INTEROPERABILITY

6.1. Short-lived ontology concept

In the federated Enterprise Interoperability approach, no common persistent ontology is supposed to exist; the communication must be accommodated on-the-fly. In consequence the ontology that structures the messages exchanged must be short-lived, (i.e. non persistent). We state that the communication mechanism, in this approach, can be informally illustrated as follow in figure 3. We mainly distinguish two cases.

In case a., the enterprise 1 sends information and the ontology to decode it at the same time. This ontology is supposed to be only valid for this information.

In the case b., the enterprise 1 sends only the information to enterprise 2. Once enterprise 2 receives the information, it checks within its local ontology if it is able to decode the information. If not, it asks for the ontology associated to the message to the sender of the message. The new received ontology can be conserved to be reused with further data sent by the same emitter.

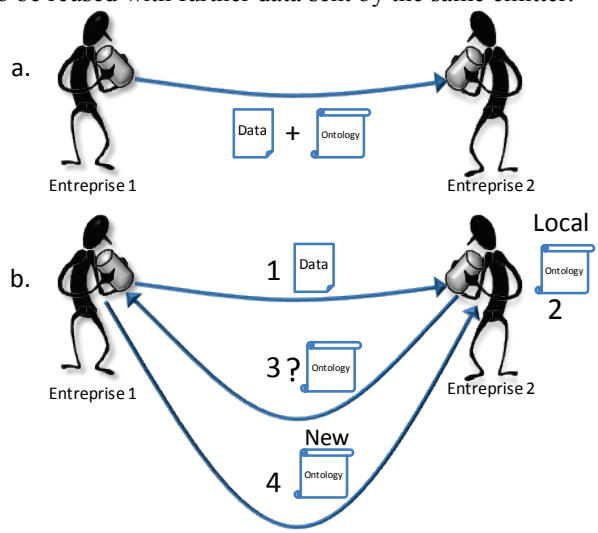


Figure 3: Federated Interoperability Data Exchange

In the first solution, the information size exchanged is more important, it can be intercepted and the confidentiality can be broken. In the second case the confidentiality is enforced but it can require more exchanges between the two partners and consequently overlapping the communication duration. Nevertheless, for confidentiality (i.e. §4.3) and accuracy to the definition of §2.3, we choose to focus in this article, on the second solution. From that postulate we introduce the concept of "short lived" ontology (our ontology definition can be based on Gruber (1995) definition), where ontology can be, in some case, suppressed after use or have finite duration validity. It maps the on the fly accommodation requirement of federated interoperability.

6.2. Agent for Short-lived ontology concept

From the concept presented in the preceding point, we state that the autonomous dialogue between Agents, from Multi Agent System (MAS) (Ferber 95) and Agent Based Simulation (Huang and Nof 2000), can map properly the "on the fly" concept of federated interoperability at process level. We propose to use the dialogue mechanism algorithm of Agent programming, introduced in (Ferber 1995) and (Huang and Nof 2000), to solve at computerized level the problem of federated Enterprise Interoperability. This dialogue between Agents, aims at establishing communication (e.g. two enterprises that discuss to agree on domain ontology), it is based on Agents cooperation behavior settings, and messages exchanged language (ontology).

On the one hand, (Ferber 1995) specifies the communication behavior algorithms of Agents' with Petri Nets (PN). On the other hand, (Zacharewicz *et al* 2008a), have tailored the use of DEVS/HLA (introduced by (Zeigler *et al 20*00) as a generalized M&S language (including PN) for distributed systems, gaining accuracy and flexibility (these models communicate within a distributed environment by message passing). Thus, we state that DEVS/HLA synchronized communication can support a unified, reusable and interoperating implementation of distributed Agents' dialogue.

On behalf of previous paragraphs propositions, we propose to develop a MAS simulator in the aim of validating Enterprise Interoperability concepts, studying the performance by simulation and implementing a concrete solution for Enterprise Interoperability ISs.

The research for developing MAS distributed Platforms is wide (as can denote for e.g. a repository of Agent-Based Simulation Platforms proposed in (Marietto *et al* 2002)), and actual MAS simulators are powerful (i.e. (Huang and Nof 2000)). Nevertheless, they mainly do not tackle the problem of interoperability and reuse of components at coding level (e.g. heterogeneity of syntax, semantic, time management, etc.). In consequence, to preserve Interoperability at all levels of Enterprise Modeling including execution level; we propose to implement an Enterprises Federated Agents System that will be HLA compliant (to guaranty also run time interoperability between heterogeneous software components).

7. IMPLEMENTATION OF FEDERATED INTEROPERABILITY COMMUNICATION

7.1. Framework Definition

The proposition starts from the statement on interoperability needs on interfacing enterprises IS in the context of project cooperation. The figure 4 depicts the requirement on exchanging data from heterogeneous Information Systems, including vendor tools such as SAP and other specific developed solutions. It is issued by generalizing study case of (Zacharewicz 2008a), various enterprises are involved in a common project and their heterogeneous components need to be interfaced. Existing interoperability between components is represented with plain arrows and in demand interoperability with dotted arrows. (Labarthe 2007) reports on solution to establish interoperability using MAS in the communication of enterprises IS (i.e. figure 4 long dotted set); they have implemented an agent communication mechanism using JADE platform facilities. Zacharewicz *et al.* (2008a) defined an interoperable Workflow using DEVS (Zeigler 2000) / HLA (i.e. figure 4 short dotted set). By joining these approaches, this paper introduces the basis for a generalized approach to realize interoperability between heterogeneous components. Two ways of research are envisaged.

The first requirement is solving the interoperability of data and services, HLA can be part of a solution.

For instance in the practical case of figure 4, a solution is to establish links to an "Interoperability Service Bus", referring to Enterprise Service Bus of Chapel (2004) concepts, to connect new features with already connected components, (e.g. DEVS/HLA Anylogic, HLA…). We detail in the next point how an HLA compliant platform can facilitate the integration of all required components.

The second point is trying to take into account the requirement of interoperability at each development process steps and how it can facilitate the interoperability at lower level of abstraction. This idea in MDI (Bourey 2007) is to define interoperability models at the Business level of modeling enterprise and to facilitate process level interoperability, and then to develop data interoperability services coherent to previous levels of abstraction. This point is not detailed in this paper but is still in our scope of studies (Zacharewicz et al. 2008b).

7.2. Definition of Ontology Level

Different levels of ontology are required in our approach. From low level with poor semantics associated to HLA objects to information transport level (HLA bus). Agent KQML (Finin 1994) will be used as an intermediate level able to match from low level description to high level description used in heterogeneous platforms, software or enterprise models involved in the system using reference to domain or application Ontologies. For instance, one challenge is to be able to transform and transport SAP, Anylogic or ARENA descriptions of a problem through a communication in a distributed network.

7.3. HLA Compliant Execution Model

As enounced in previous section, this new interoperability concept of Enterprise Federated Agent needs to be tackled at run time. Based on our experience in HLA support, we propose an innovative implementation of Enterprise Interoperability Federation (i.e. figure 5).

The "Interoperability" components layers can be added to IS of enterprise either they are ad hoc developed or vendor solutions. The idea is to add a component to code and decode information exchanged with the original IS, this component is considered as black box and no modification is realised on it. We present in detail in this section the components required for this global distributed platform.

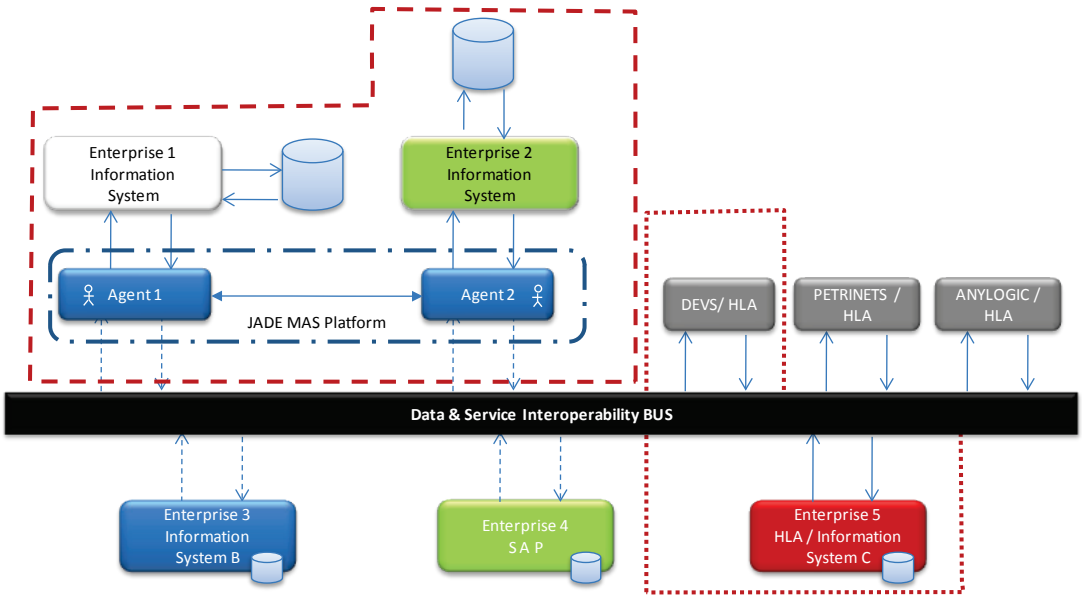


Figure 4: Federated Interoperability Data Exchange

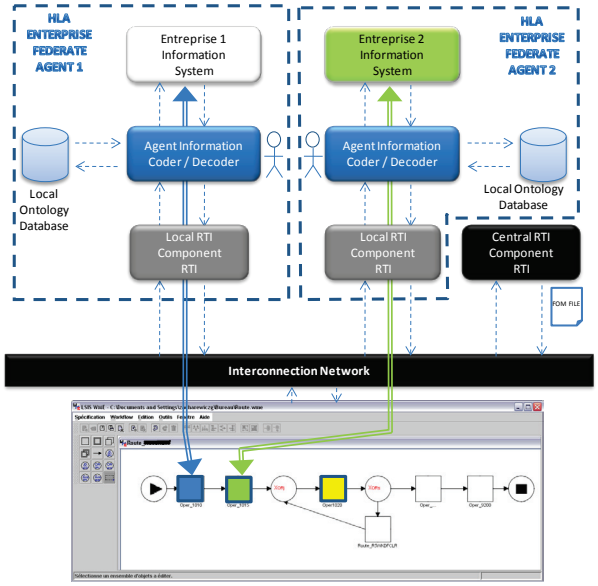


Figure 5: Federate Agent Data Exchange Process flow

7.3.1. Information System Services Layer

The distributed implementation requires the extension of two add-ons to the local enterprise IS to define HLA Enterprise Federate Agent (HEFA). We present in the figure 5, the elements of this new architecture. The respective local enterprise ISs remain unchanged, HLA only required to add components to interface with input output messages of the IS.

7.3.2. Multi Agent Coding / Decoding Layer

All agents involved in the data ontology matching are detailed:

Agent 1 Storing data: This agent is employed to store the received information and will check the capacity to decode information using a communication with ontology agent. Receiving the agreement to use the data, it sends data and ontology to the information system

Agent 2 Ontology: This agent is linked to a repository of local ontology; it checks the consistency of the information regarding the local ontology and decides if the data can be exploited. If yes, it sends back to the storing data agent the information and the ontology to use. If not, it asks to the communication agent to start dialog to obtain the appropriate ontology.

Agent 3 Communication: This agent will start a conversation with the respective agent of the data sender to deal on the modality to receive the appropriate ontology. We propose, in the following, that this dialog will be established using HLA message communication protocol but from a conceptual consideration it can be considered as a general approach where HLA is just one practical technical solution.

7.3.3. Local RTI Component Layer

This level is the lower level; it is the service and data level. It deals with technologies employed to exchange computerized information.

The first component is the Local RTI component (LRC), i.e. §3.2. This code is required to insure the reception and the emission of information within the transposing from the local description of data to the HLA format. This component manages also the time by delivering the message according to their timestamp and to the local time.

The second one is the Agent Information Coder Decoder (ICD). This new specific software component, introduced in this paper, will be able to analyze the information received by querying the local ontology data base to determine the capacity to interpret the received message according to the behavior introduced in § 6.1. If the local ontology is not able to interpret the received message, ICD will ask through the LRC for the associated ontology to the sender of the message. At the end, this is the technical transcription of the Agent communication behavior.

We illustrate in figure 5, the use of two instances of our structure connected to a Workflow monitoring tool. This tool is employed to run a simulation of a Supply chain or of a document exchanging process and is triggering right in time the information systems of the interoperating enterprises.

7.4. HLA Interaction and Object Class Model

The idea is to propose a new mapping for data to be exchanged in a HLA compliant distributed system between Interoperable Enterprise Agents. We propose in figure 4, a generic HLA FOM that will support the descriptions of the data required to insure the exchange of information in the figure 3. case b.

HLA object Root (N)	Information channel (S)	Information channel 1 (PS)
		Information channel 2 (PS)
		Information channel 3 (PS)
		Information channel 4 (PS)
		Information channel ... (PS)
		Information channel n (PS)
HLA Interaction Root (N)	Ontology (S)	Ontology 1 (PS)
		Ontology 2 (PS)
		Ontology 3 (PS)
		Ontology 4 (PS)
		Ontology … (PS)
		Ontology n (PS)

Figure 6: Federate Enterprise Interoperability FOM

Information to exchange: Information exchanged between enterprises information systems will be mapped with HLA Object class models (that handle persistent information in the distributed execution). The enterprise IS federates will publish and subscribe (PS) to these classes of information. An information channel Object (i.e. figure 6) represents the informational link between at least two enterprises. We notice that communication channels preserve confidentiality.

Ontology: Ontologies exchanged between enterprises are not persistent in the studied approach; they will be mapped into HLA Interactions (that are non persistent information exchanged). Enterprise IS Federate

Agents will publish and subscribe (PS) to these classes of information. One Ontology class (i.e. figure 6) is associated to each information channel; each Ontology definition can change during run time. Eventually, validity duration can be set for each Ontology.

The structure of the generic ontology and of the messages will be implemented in the FOM presented in figure 6. Because of generic concepts introduced in this paper, HLA interaction parameters and HLA object attributes are not fully specified; they will be in more detailed depending on the applications. It gives flexibility to the data structures exchanged.

7.5. Data Exchange Time Management

To respect causality, we recall in that section, an HLA time management specification of exchanged messages sequence between HEFAs introduced in (Zacharewicz 2009).

At first point, each HEFA sets its channels of communication that link it with other HEFA it is interested to interoperate. For that purpose, it defines Publish/Subscribe (PS) participation to HLA objects defined in the FOM (see figure 6). At the same time, the HEFA can set the ontology's it is able to provide (P) and the ones it is interested to receive (S) at the beginning; (this status will evolve during run time because of creation and destruction of non persistent ontology).

In figure 7, we describe the time management mechanism to exchange information between two interoperating HEFAs. The services and call backs provided by the RTI are here mentioned by directly referring to HLA 1516 standard reference book (IEEE 2000).

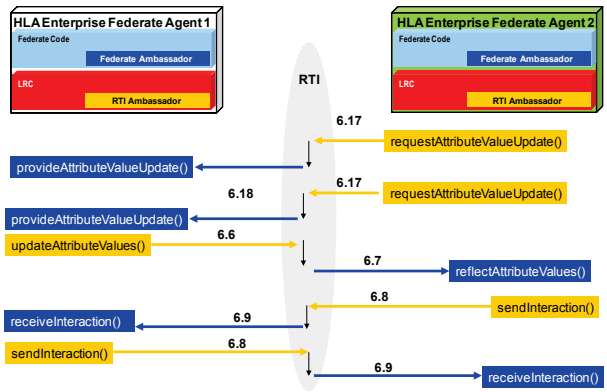


Figure 7: Federate Data Exchange Time Management

Assuming an HEFA2 is interested in information emitted by HEFA1. HEFA2 must subscribe to the information published by an HEFA1. Then, the steps of information exchange during run time are described in figure 7. In first step, HEFA 2 asks for information from HEFA1 by calling *requestAttibuteValueUpdate()* RTI service (e.g. in figure 7, HEFA 2 asks for two data to HEFA 1). The RTI transmits these demands to EF1 with *provideAttibuteValueUpdate()* callback. When HEFA1 is ready to distribute these information or their new values (regarding to time constraints) it will deliver them to HEFA2 with *updateAttibuteValues()*. HEFA2 possesses now the information. It transmits it to its information coder / decoder component (i.e. figure 4). This component checks the capacity to decode the message with its local ontology. If it is not able, it asks for the associated ontology of the message with a non persistent questioning to HEFA1, i.e. by using the service *sendInteraction()*. We notice that these actions are transcript from Agent behaviour specification. Once HEFA1 receive*Interaction()* call back, It will supply the ontology with the same communication process (*sendInteraction()* to HEFA2). In fine, the ontology is just sent and received; no common semantic information persists between HEFA1 and HEFA2. Yet, the ontology received can be locally stored to simplify and speedup next data exchange between HEFA1/HEFA2.

8. CONCLUSION AND FUTURE WORK

This article has given a state of the art of Enterprise Interoperability concepts and illustrated the use of Agent concepts and HLA standard for the implementation of enterprise applications federations.

From the new concept of short-lived enterprise ontology for federated Enterprise Interoperability, we proposed a specific implementation of distributed enterprise models for simulation or real time information exchange. At the end, the keys for implementation given by Agent dialogue mechanism has helped to bridge the gap from Enterprise Interoperability concepts to HLA compliant distributed implementation in the field of Enterprise Modeling by following a new standardized and systematic approach.

Currently, the environment is at specification and conception time. Meantime of verification and validation of our approach, the development of a beta version is beginning in the aim to test, in particular, the implementation of the short-lived Ontologies. The next step of detailed conception and implementation will be initiated in the perspective of the rising partnership between IMS Bordeaux and CIRRELT Québec labs.

9. REFERENCES

Bourey, J.P., Grangel Seguer, R., Doumeingts, G., and Berre, A.J., 2007. *Report on Model Driven Interoperability*, Deliverable DTG 2.3, *INTEROP NoE*, April, pp. 91. Available from: http://www.interop-vlab.eu/ [accessed in 15 June 2009]

Chappell D. A., 2004. *Enterprise Service Bus*, Portland: O'Reilly.

Chen, D., Dassisti, M., and Elvesæter, B., 2007. *Enterprise Interoperability Framework and Knowledge Corpus - Final report Annex: Knowledge Pieces,* Contract no.: IST508 011, Deliverable DI.3 Annex, Available from: http://www.interop-vlab.eu/ [accessed in 15 June 2009]

Chen, D., Ewald, R., Theodoropoulos, G. K., Minson, R., Oguara, T., Lees, M., Logan, B. and Uhrmacher, A. M., 2008. Data access in distributed simulations of multi-agent systems. *Journal of Systems and Software*, 81(12), 2345-2360.

Cicirelli, F., Furfaro, A. and Nigro, L., 2009. An Agent Infrastructure over HLA for Distributed Simulation of Reconfigurable Systems and its Application to UAV Coordination. *Simulation*, 85(1), 17-32.

Ferber, J., 1999. *Multi-agents Systems. An Introduction to Distributed Artificial Intelligence*, London: Addison Wesley.

Finin, T., Fritzson, R., McKay, D., and McEntire, R., 1994. KQML as an agent communication language. *Proceedings of the Third international Conference on information and Knowledge Management*, November 29 - December 02, Gaithersburg, (Maryland, USA).

Fujimoto, R.M., 2000. *Parallel and Distributed Simulation System*, New York: Wiley Interscience.

Huang, C. Y. and Nof S. Y., 2000. Autonomy and viability measures for Agent-based manufacturing systems. *International Journal of Production Research*, 38(17), 4129–4148.

IDEAS Consortium, 2002. *Thematic Network, IDEAS Interoperability Development for Enterprise Application and Software Roadmaps,* Annex 1 – Description of work. Available from: http://www.interop-vlab.eu/ [accessed 15 June 2009]

IST. 2006. *Enterprise Interoperability Research Roadmap*, Final Version, July, Available from: http://cordis.europa.eu/home_en.html [accessed 15 September 2007]

IEEE std 1516.22000, 2000. *IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) - Federate Interface Specificatio*n, New York: Institute of Electrical and Electronic Engineers.

Jennings, N.R., Sycara, K. and Wooldridge, M., 1998. A Roadmap of Agent Research and Development", *Autonomous Agents and Multi-agents Systems,* 1(1), 7-38.

Labarthe O., Espinasse, B., Ferrarini, A. and Montreuil B. 2007. Toward a Methodological Framework for Agent-Based Modelling and Simulation of Supply Chains in a Mass Customization Context. *Simulation Modelling Practice and Theory*, 15(2), 113–136.

Lees, M., Logan, B. and Theodoropoulos, G. K., 2007. Distributed simulation of agent-based systems with HLA. *ACM Transactions on Modeling and Computer Simulation*, 17(1).

Logan, B. and Theodoropoulos, G. K., 2001. The distributed simulation of multi-agent systems. *Proceedings of the IEEE - Special Issue on Agent-oriented Software Approaches in Distributed Modelling and Simulation,* 89(2), 174–185.

Marietto, M. B., David, N., Sichman J.S. and Coelho, H., 2002. Requirements Analysis and Agent-Based Simulation Platforms: State of the Art and New Prospects. *Lecture Notes in Artificial Intelligence,* Berlin: Springer-Verlag, 2581, 89–104.

Minson, R. and Theodoropoulos, G. K., 2008. Distributing RePast agent-based simulations with HLA. *Concurrency and Computation: Practice & Experience*, 20(10), 1225-1256.

Nwana, H.S., 1996. Software agents: An overview. *Knowledge Engineering Review*, 11(3), 205–244.

OMG, 2003. *MDA Guide Version 1.0.1*. Object Management Group, Document number: OMG / 20030601 Available from: www.omg.org/docs/-omg/03-06-01.pdf [accessed 15 June 2009]

Gruber T. R., 1993. Towards Principles for the Design of Ontologies Used for Knowledge Sharing, In: *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Dordrecht: Kluwer Academic Publishers.

Wooldridge M., 2002. *An Introduction to Multi-agents Systems*, NY: John Wiley & Sons.

Song, H. J., Shen, Z.Q., Miao, C.Y., Tan, A.H. and Zhao G. P., 2007. The Multi-Agent Data Collection in HLA-Based Simulation System. *Proceedings of the 21st International Workshop on Principles of Advanced and Distributed Simulation*, pp 61-69, June 12-15, San Diego, (CA, USA).

UEML, 2003. *Public Web for UEML*, EU Project IST–2001–34229, Available from http://www.uemlwiki.org [accessed 15 June 2009]

Wooldridge, M. and Jennings N.R., 1995. Intelligent Agent: Theory and practice. *Knowledge Engineering Review*, 10(2), 115–142.

Zacharewicz, G., Frydman, C., and Giambiasi, N., 2008a. G-DEVS/HLA Environment for Distributed Simulations of Workflows. *Simulation*, 84(5), 197–213.

Zacharewicz, G., Chen, D., Vallespir, B., 2008b. HLA Supported, Federation Oriented Enterprise Interoperability, Application to Aerospace Enterprises. 08E-SIW-074. *Proceeedings of 2008 International Simulation Multiconference EuroSISO,* June 19-19, Edinburgh (Scotland).

Zacharewicz, G., Chen, D., Vallespir, B., 2009. Short-Lived Ontology Approach for Agent/HLA Federated Enterprise Interoperability*, Proceedings IEEE of International Conference I-ESA China 2009 Interoperability for Enterprise Software and Applications*, pp. , April 22-23, Beijing (China)

Zeigler, B., Kim, T.G., Praehofer, H., 2000. *Theory of Modeling and Simulation*, New York: Academic Press.

AUTHORS BIOGRAPHY

Gregory ZACHAREWICZ is Associate Professor in Bordeaux 1 University (IUT MP). His research interests include Discrete Event Modelling (e.g. DEVS), Distributed Simulation, Distributed Synchronization Algorithms, HLA and Workflow. He recently focused on Enterprise Modeling and Interoperability. He has published more than 20 papers in Conference and International Journals. He is Reviewer in Conferences (Summer SCS, Winter-Sim...) and SCS Simulation journals.

Olivier LABARTHE is Associate Professor of operations and decisions systems in the Faculty of Administrative Sciences at Laval University in Québec, Canada. He received a PhD in Administrative Sciences and Computer Science, from Laval University and Aix-Marseille III University respectively. His research interests focused on the multi-agent modeling and simulation for the design and the management of manufacturing and logistics networks.

David CHEN is Professor at University Bordeaux 1. His research interests focus on enterprise modelling and interoperability. He has been actively participating in EU program (FP1-FP6), several cooperation programs between EU and China. David CHEN is member of CEN TC 310/WG1 and ISO TC184/SC5/WG1 (Modelling and architecture) and involved in IFIP WG5.12 and IFAC WG5.3 (Enterprise Integration and Networking). He has published more than 80 papers in international journals and conferences.

Bruno VALLESPIR is Professor at University Bordeaux I and member of IMS. Within this framework, he is today responsible of the research group about production systems. Bruno VALLESPIR's research topics are production management, performance evaluation and enterprise modelling including modelling tools and design methodologies. On these topics, he published more than 60 papers in conferences, journals and books, was involved in around 20 international, European (from FP1 to FP6, mainly ESPRIT) or French projects and working groups. Recently, he was involved in INTEROP NOE. He is member of several working groups and steering committees of IFIP and IFAC.

SIMULATION OF THE EMERSION PROCEDURE FOR A NEW UNDERWATER GENERATOR

Amable López Piñeiro[(a)], José Andrés Somolinos[(b)], Luis Ramón Núñez Rivas[(c)] and Eva María Novoa Rojas[(d)]

[(a)(b)(c)]Departamento de Sistemas Oceánicos y Navales. ETSIN UPM Madrid

Arco de la victoria s/n 28040 Madrid

[(d)]Fundación Instituto Tecnológico SOERMAR

Cardenal Herrera Oria 57, 28034 Madrid

[(a)]amable.lopez@upm.es, [(b)]joseandres.Somolinos@upm.es, [(c)]luisramon.nunez@upm.es, [(d)] eva.novoa@soermar.com

ABSTRACT

During the last years, a great amount of electrical underwater generator designs have been proposed for exploiting the sea energy resources. Multidisciplinary engineering is needed to conceive, to design and to construct any of these devices and a large amount of concepts have to be taken into account when an underwater generator is designed. SOERMAR and the Technical University of Madrid are developing a new concept for marine current energy use: The GESMEY project, whose main goals and conceptual design are briefly described. It is based on a Y structure with additional elements which can handle its own floatability by controlling the amount of water ballast inside these elements.

This paper analyses one of the methods proposed to move the generator from the operation depth to the sea surface in order to prepare it for the maintenance position in a controlled way. The obtained simulated results of the emersion procedure are presented in this paper.

Keywords: Underwater generators, Simulation, Non-linear systems

1. INTRODUCTION

The growing energy demand of the world, the variable cost of oil and the Kyoto agreement to decrease the CO2 emissions, have made that several technicians and scientific to study alternative ways of obtaining energy. By considering the ocean as one of the most important renewable energy resource, different technologies have appeared for extracting energy from it, some of these ones stand out for its importance: wind offshore, waves energy and tidal and current energy. (Bedard, Roger, 2008), (Savage, Anne et al. 2007)

The last one has a lot of locations with high energy density, better mean/max power ratio that wind and waves systems, lower environmental impact, higher expected reliability and, long term predictable velocity and energy due that thermal, salinity and these seasonal currents are well known and nearly constant, on the other side, tidal currents can be predicted a lot of years in advance with a high accuracy. (Delgado Cabello, Javier et al. 2006)

There are a lot of concepts proposed for sea wind, tidal and wave energy extraction. Same decades ago, tidal energy was obtained by means of artificial barriers (like La Rance facility), but at present most of the prototypes that have been proposed use techniques, that are following similar principles of those for extracting energy from the wind. If the layout of underwater generators is generally based on a marine current turbine, the main differences between some prototypes are related to the rotor size and orientation, the location of the generators and the devices of anchoring to the sea bottom. (Eddine Ben Elghali, Seif et al. 2007), (Fraenkel, P.L. 2007).

The Ocean Current Turbines (OCT) for extracting electrical energy from the marine currents have the advantage (in comparison with the wind energy generators) of the greater density of the water, which implies smaller size of the propeller.

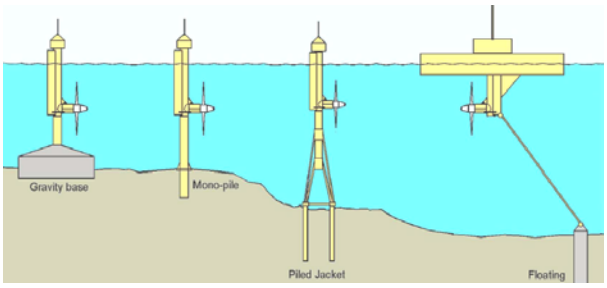


Figure 1: Conventional anchoring systems

The main limitations of these prototypes can be considered as follow: Important civil work on bottom and special equipment for deploying with big elevation mechanism to maintenance (See figures 1 and 2). In order to minimize these disadvantages, UPM and SOERMAR are developing an underwater system, with an improved technology, capable to produce energy from the ocean currents: the GESMEY generator.

Figure 2: Existing devices

Grouping all the generators of different technologies, as function of their field of application; the evolution of the OCTs can be classified attending to this technologies integration in the following way: (See figure 2)

- 1st generation: Designed for the utilization of the tides, closing estuaries with preys and using similar turbines than in the fluvial flows of water.
- 2nd generation: Systems placed over the bottom of the sea. There exists a great set of devices. The main limitation is the civil work for installing the generator and the cost of the maintenance tasks, which requires special machinery, or extracting the generator to the surface of the sea with external machines… Another limitation is the relative small depth of the bottom of the sea for this generation of devices
- 3rd generation: Designs for high depth bottoms (more than 40 m), or floating devices, and the GESMEY design which is briefly described below.

From a strategically point of view of the Spanish coasts where marine currents can be exploited, the zone with the biggest currents in Spain is the Strait of Gibraltar (GS). This strait is the natural zone of communication between the Mediterranean Sea and the Atlantic Ocean. It has a variable width from 44 km in the west input until the minimum of 14 km in the narrowest zone between *Tarifa* and *Punta Cires*.

There a great exchange of water exists between both seas. In general, the currents can be considered as bi-directional ones. In this particular area, the main causes of this exchange can be summarized attending to the following natural considerations:

- Different density of water of both seas. Because of their not equal salinity and temperature.
- Difference of levels between both seas due to the tides and due to the evaporation of water in the Mediterranean Sea side of the strait.
- Difference between the atmospheric pressure from the Alboran Sea (east zone) and the Gulf of Cadiz (west zone).

More precise values and studies of the properties of the marine currents in the zone of the Strait of Gibraltar can be seen at (Vargas, J.M. et al. 2006),(López Piñeiro, A. et al. 2008), (Juanes González J.M,2007).

2. SYSTEM DESCRIPTION

The GESMEY system (Figure 3) has a propeller with three blades that there moves an electrical generator, which is located in the central dome or pod. Three arms joins the dome with each of its tip float. The ensemble is joined by a device with ropes that anchor to the sea bed.

The main function of the central dome is to locate the elements for converting the mechanical energy of the propeller into the electric power that can be extracted to an electric power station of the land using a submarine cable that includes the electrical line of power and the link of communication with the control remote station.

Inside the central dome, the auxiliary devices needed for cooling, protection, measurement and control are located. An important part of the inner volume of the floats is used as water ballast stores. The control of the quantity of water ballast allows handling the position and/or the orientation of the generator. Different control schemes can be used for this purpose, because one, or several, or all the water ballast stores can be managed individually.

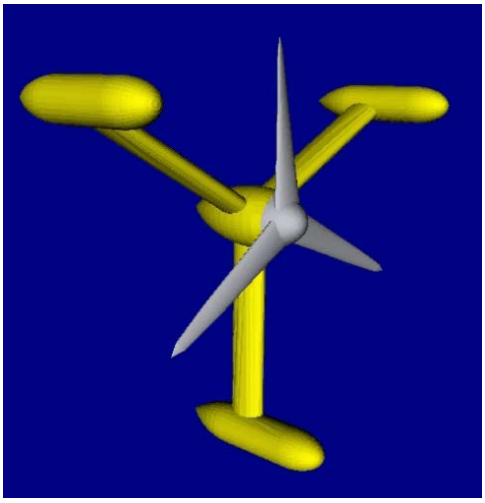


Figure 3: Basic GESMEY design

The weights and forces distribution lets the system, compensate the torques applied to the generator in the sense that it can line up with the own direction of the marine currents. The applied torque to the propeller makes the generator be orientated as it is shown in figure 4.

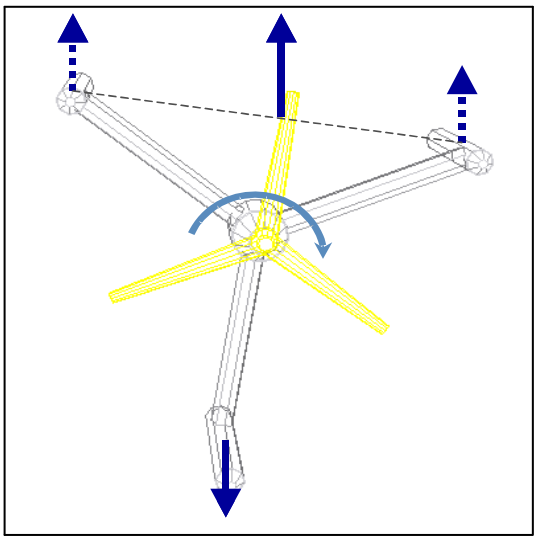


Figure 4: Main Forces and torque acting over the GESMEY

The proposed system will have a few economic costs lower than those of other existing generators due to its major stability, flexibility and because the operations of maintenance have been highly simplified.

The modes of operation of GESMEY can be summarized in two ones: the normal operation mode -submerged-, and the state of operation for maintenance -floating in the sea surface-. One of the principal aims in the design of the generator was to facilitate the maintenance work from a point of view of the access to the elements of the system with a self-operated system to move from one operation mode to the other.

Another of the most evident improvements of the proposed design is that the use of external elements or external machinery for the emersion or sunk of the generator is avoided.

It is possible to control the depth of the generator handling the quantity of water ballast inside it and/or controlling the tension of the ropes that anchor the generator to the sea bed. Both procedures of immersion and emersion seem to be symmetrical. Nevertheless, the emersion procedure turns out to be critical and is for it that needs a special control to move from the depth of operation to the surface, with uncontrolled buoyancy lost, without sudden changes, and in a smooth way.

3. MODES OF OPERATION

The modes of operation of GESMEY can be summarized in two ones:

- The normal operation mode -submerged-, (as Figure 3) with a vertical orientation. In this mode of operation, the generator can be of any of these states: in a blocked state (ordered from the control unit), with the propeller in a stopped state, because of the small current velocity and in normal operation state, that is to say, generating electrical energy.
- The state of operation for maintenance -floating in the sea surface- with an horizontal orientation. One of the principal aims in the design of the generator was to facilitate the maintenance work from a point of view of the access to the elements of the system with a self-operated system to move from one operation mode to the other.

Using an automated procedure of emersion, another one of the evident improvements of the proposed design is that the use of external elements or external machinery for the emersion or sunk of the generator, can be avoided, since it is the own generator which is capable of generating the control signals to go to the surface of the sea without any kind of external intervention.

The change from a mode operation to another has to be ordered by a human operator from the SCADA system, and has to be executed by the own generator in an completely automated way. Two basic procedures represent the changes of modes of operation of the generator:

- The immersion procedure lets the generator go from the surface of the sea to the normal position of producing energy (vertical and submerged).
- The emersion procedure makes the generator become to an horizontal posture over the sea, starting from the position of normal operation. Figure 5 shows the different initial-intermediate-final stages during the emersion procedure.

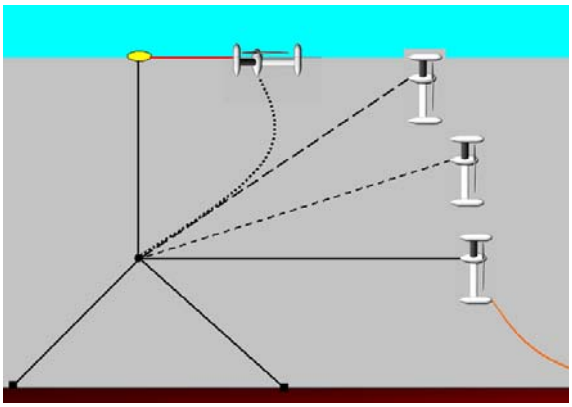


Figure 5: Schematics view of the emersion procedure of the generator.

4. CUALITATIVE DYNAMIC MODEL OF THE SYSTEM

The dynamic equations that govern the movements of the generator are obtained from a model of only one point mass located in the central dome. The applied forces (it is a strongly non linear system) that act over

the generator are of different nature, and can be grouped as:

- The vertical forces: Composed by the own (proper) weight, the floating force (which is controllable), the vertical hydrodynamic friction force and the vertical components of the reaction forces over the ropes
- The horizontal forces: The hydro-dynamical forces due to the marine currents and the horizontal components of the reaction forces over the ropes.

Simulation of the quasi-static position and orientation of the generator under the normal mode of operation requires the use of distributed masses along the structure and a more precise forces and torques computation. For the dynamical response of the system during transient motions, (immersion and emersion procedures), the dynamical equations are obtained from the mentioned forces applied to the main mass. All forces are obtained as function of the values of the estimated current speed in the Gibraltar Strait environment.

Taking into account all the equations implied in the system, a dynamical model with next properties is obtained:

- It is an strongly non linear system
- The coefficients of the equations are time dependant
- There exists a great coupling among many of the equations.

These are the main causes because it is necessary the simulation of the system into a computer in order to obtain the solution (temporal response of the system) during the emersion procedure. These equations have been programmed in the SIMULINK-MATLAB environment using double integration and variable step discretization.

Because of the simplicity of the design, the only control signal which lets handle GESMEY is the amount of water ballast inside the generator. Based on this limitation, controllability of the system, time response signals, levels of saturation of the actuators, and some others results have to be analyzed.

5. SIMULATION RESULTS

Figure 6 shows the blocks diagram implemented in the SIMULINK environment. It clearly can be seen the non linearity of the system and the strong coupling among equations mentioned in the previous section.

The values and coefficients used in this simulation are corresponding estimations from some of the Gibraltar Strait zones. The dimensions of the generator and lengths of the ropes have been optimized for each of the concrete areas.

The obtained results confirm that it is necessary the generation of reference signals for controlling the management of the ballasts of the generator, so that they allow us to obtain a smooth response, with no strong oscillations.

Without an adequate reference of the water ballast of the underwater generator, or without any closed-loop control system in the emersion procedure, it is possible to put in risk the own generator, as well as the devices,

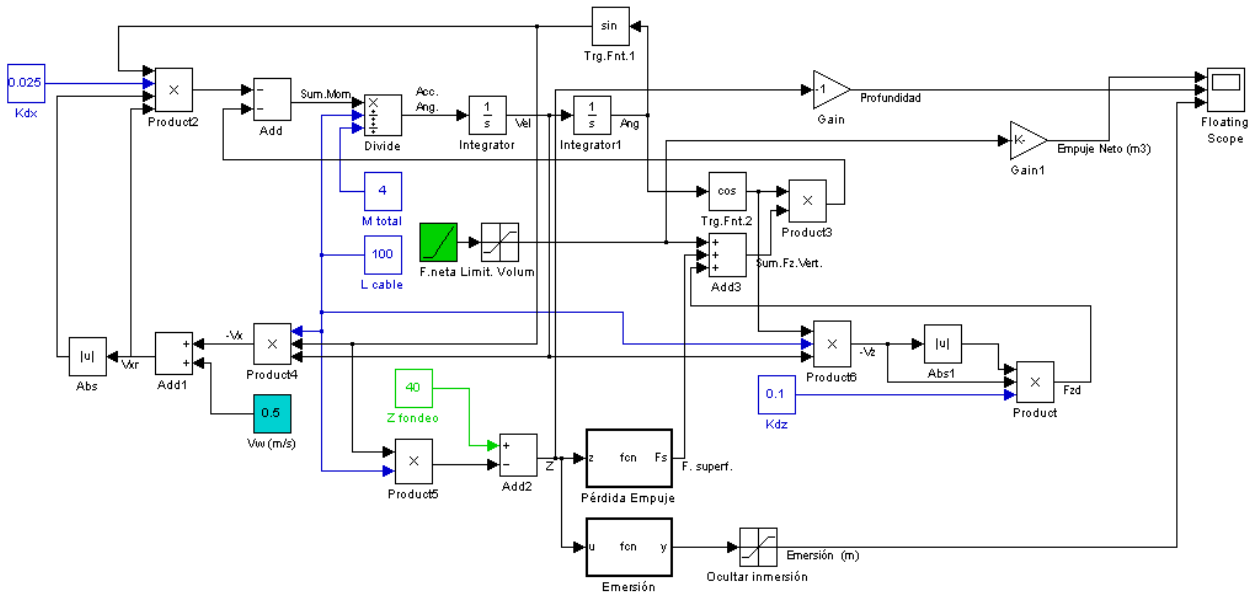


Figure 6. Blocks diagram in the SIMULINK environment

components and even personnel of the team of maintenance, when it emerges from depth.

With an adequate reference of control of water ballast, it is possible to obtain a good response of the system during the emersion procedure.

The used reference signal can be seen from figure 7. It represents the evolution of the input signal in m^3 of removed water ballast. At the beginning of the operation, the water tanks are practically full, and a very small resulting (floating) force is applied to the generator. The procedure beginning instant occurs when the ropes that hold the generator are given up (t = 0 seconds). If a bigger floating force is used as reference signal, the response of the system can become an inadmissible overshooting response.

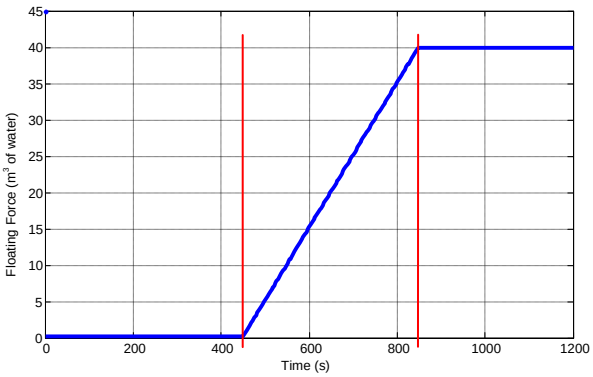


Figure 7. Signal of reference.

Then, when the generator has reached the surface of the sea, a ramp input signal is used as reference to empty the tanks of water in order to the generator emerges so much as it could. The desired depth position of the generator is defined by its own weight with the tanks of water ballast fully empty (minimum).

One of the main ropes is joined with the generator in a permanent mode. This link prevents the generator takes an undesirable path far from the nominal position. This main rope and the interface between water and the air let the generator turn from a vertical to an horizontal orientation without any controlled external effort.

The simulated response of the depth of the generator (see figure 8) shows that the emersion will be carried out in approximately 500 seconds. It shows a temporary response without any overshoot and with an oscillating superposed signal of very small amplitude and small damping. The vertical solid lines denote the time interval when the water ballast stores are being emptied completely.

In order to characterize the mentioned small oscillating signal in the steady state, this one exhibit an amplitude of over 0.6 m (when a 100 m length cable and a generator of about 30 m diameter is used), a natural frequency of oscillation of approximately 2.4 cycles per minute and a small damping ratio. From a practical point of view, this signal will be able not to be taken into account if a simple comparison between this signal and the amplitude of sea waves during the maintenance period (calmness) is made.

A few seconds after the transient emersion of the generator, the steady state error of the response can be considered cero if small oscillating signal is neglected.

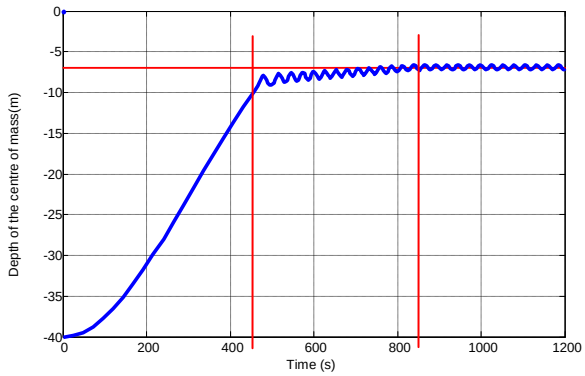


Figure 8. Depth of the generator during the emersion procedure

The difference between the depth of the center of mass of the generator from the beginning instant of the ramp until the end of the ramp time can be easily appreciate in the same figure 8. For this simulation, the obtained value of this distance is about 1.58 m. This is considered of a great importance to let a human worker be able to access to the interior of the dome with major comfort, without the need of a diver not of special facilities, as soon as the generator has reached the final position.

6. CONCLUSIONS

A new submarine generator for the exploitation of the energy of the marine currents has been briefly described in this study. Its special configuration allows these generators to be installed in deeper waters than some of the existing generators up to now.

The costs of the necessary civil work for its installation, as well as the economic costs of the maintenance tasks are minimized due to the fact that the generator is able to emerge and it can also be immersed simply controlling the water ballasts inside the own generator and the ropes of anchorage to the sea bottom.

These operations of immersion and emersion can be done in a completely automatic way, either from the order given by a human operator or a remote control system. It can be easily observed that the emersion procedure results the most critical one in the sense of the risk of the execution, the risks of persons, devices, and the own generator. In other words, the immersion results a simpler and less complex procedure than the emersion one.

The equations that govern the dynamic behaviour of the generator in these transitions, turn out to be not linear, but they are strongly connected and with coefficients dependent on time. These circumstances

increase the big risks that exist in case the procedure of emersion is not carried out in a controlled way.

The simulation of the obtained equations in a computer is necessary in order to study the dynamic behaviour of the system and to design the appropriate input signal of reference.

Results show that it is possible to execute the emersion procedure in a completely automated way, by simply controlling the ballast water of the generator.

In order to improve the system response during the same emersion procedure, a sensor of depth and an inclinometer are currently being used to control the depth and the orientation of the generator respectively, with a closed-loop control scheme.

ACKNOWLEDGMENTS

This work was partially supported by the Spanish "Ministerio de Ciencia e Investigación" under grant IAP-590000-2008-6 in collaboration with the SOERMAR Foundation and the Universidad Politécnica de Madrid for the development of the patent number P200700987.

REFERENCES

Bedard, Roger, 2008 "Technology Characterization Ocean Wave and Tidal Energy". *Global Marine Energy Conference*. New York.

Delgado Cabello, Javier et al. 2006. "Two-dimensional structure of the tidal flor at Tarifa Narrows, Strait of Gibraltar". *5th Asamblea Hispano-Portuguesa de Geodesia y Geofísica* Sevilla

Eddine Ben Elghali, Seif et al. 2007. " A Simulation Model for the Evaluation of the Electrical Power Potential Harnessed by a Marine Current Turbine". *IEEE Journal of Oceanic Engineering*, Vol 32, No. 4.

Fraenkel, P.L. 2007 "Marine current turbines: pioneering the development of marine kinetic energy converters". *IMechE 2007 Proceedings. IMechE* Vol. 221 Part A: J. Power and Energy

Juanes González J.M. PH. Doctor Thesis: "*El potencial energético útil, de las corrientes marinas en el Estrecho de Gibraltar*". E.T.S. Ingenieros Navales U.P.M. Madrid, 2007.

López Piñeiro, A. et al. 2008. "Tecnologías para el aprovechamiento de la Energía de las Corrientes Marinas". *Proceedings of the 47th Congreso de Ingeniería Naval e Industria Marítima*. Palma de Mallorca

Oficina Española de Patentes y Marcas. 2008 "Título de Patente de Invención nº P200700987" a la UPM. OEPM. Madrid.

Savage, Anne et al. 2007. "Tidal Technologies Overview". *Sustainable Development Commision. ENTEC UK Limited*. Bristol.

SOERMAR – UPM Memorias parciales del Proyecto GESMEY. l VI PLAN NACIONAL I+D+i Subprograma de investigación aplicada industrial IAP-590000-2008-6 Madrid (Restricted distribution)

Vargas, J.M. et al. 2006. "Fortnightly and montly vairability of the exchange through the Strait of Gibraltar" *Progress in Oceanograpy*. Vol 70. Issues 2-4. Pages 466-485

AUTHORS BIOGRAPHY

Dr. López Piñeiro, A. received the Master degree in Naval Architecture and Marine Engineering in 1978, the Ph. D. degree in 1980 and the Professor degree in 1988 from ETSIN at Universidad Politécnica de Madrid. From this date, he has been the head of the R&D group on electric and electronic systems of the ETSIN. From 2007 is working on marine current turbines and at this moment is the head of the ETSIN R&D group of GESMEY project

Dr. Somolinos, J.A. was born in Madrid, Spain, on 1972. He received the M.S. degree in industrial engineering (Honors) and the Ph.D. degree (PremioExtraordinario) from the Universidad Nacional de Educación a Distancia, Spain, in 1995 and 1999, respectively. He was with the Electrical Engineering Department, Universidad de Castilla-La Mancha from 1995 to 2006 where he became an Associate Professor in 2002. From 2006 to 2008 he was a Visiting Professor at the Control and Automation Department, Universidad Carlos III. At present, he is with the Department of Oceanic Systems, Universidad Politécnica de Madrid, Spain. His research interests include modeling and control of robotic systems, instrumentation and industrial electronics.

Dr. Nuñez Rivas L.R . is presently Professor in Marine Renewable Energy Sources at the ETS Ingenieros Navales (Polytechnic University of Madrid),he´s joined at this Institution since 1984 and had several positions there, first as Academic Secretary and late as Director during four years. He has also worked for the Spanish Shipbuilding Industry and now he is a consultant in several Official Spanish Research Institutions on the maritime sector"

Novoa Rojas, E.M. received the Master degree in Naval Architecture and Marine Engineering from the ETSIN at Universidad Politecnica de Madrid. Currently she is the general manager of the Fundación Instituto Tecnológico SOERMAR.

AN INTEGRATED SIMULATION AND OPERATION COSTING APPROACH TO ASSESS ORDERS PRODUCTION COSTS: A REAL APPLICATION

Antonio Avai[(a)], Marzio Sorlini[(b)],

[(a)] Technology Transfer System s.r.l., Milano, Italy
[(b)] SUPSI-ICIMSI, Manno, Switzerland

[(a)] avai@ttsnetwork.com, [(b)]marzio.sorlini@supsi.ch,

ABSTRACT
This paper describes a new and integrated approach for providing an accurate and reliable estimation of work orders production costs meant to ease customers' bid preparation. For many small and medium third-party service providers, facing almost homogeneous product portfolio and a finite number of production lines, is really worthwhile to provide a precise production costs estimation as a basis for further decisions to accept or reject an order. Because of information on processing times are tough to be estimated in such companies, the idea was to integrate the discrete event simulation technique, providing accurate and reliable process evaluations, with the operation costing technique to assess orders production costs. A new tool called "*cost simulator*" was developed based on this integrated approach and it was validated on a real industrial case in the wooden painting sector.

Keywords: discrete event simulation, operation costing, order cost, production management

1. INTRODUCTION

Discrete event simulation is widely used to model production lines (Roser et al. 2003) and to analyze and optimize their overall performances as well as their behaviour (Boer et al. 1993), (Voorhorst et al. 2007).

Cost analysis integrated with simulation technique has been mainly investigated to support decisions during the process **design phase** for instances to assess the net present value (Murray et al. 2000), to coordinate manufacturing investment with marketing and product design to jointly achieve optimal product line solutions (Michalek et al. 2005), to find an optimal production control (Gahagan et al. 2005) or to assess the efficiency of business process reengineering (Wang et al. 2001). Such an approach has been applied also to the **logistic area**, for example to compare transportation costs in a cooperative and a non-cooperative context (Diaz et al. 2003) or to evaluate the operational costs and performances associated with liner shipping (McLean 2008). While Von Beck has linked Activity Based Costing (ABC) and discrete event simulation to provide cost estimation in manufacturing environments to support more informed operational and strategic decisions (Von Beck et al. 2000).

The operation costing technique is more suitable for third-party services because it represents a good compromise between reliability of forecasted data, inputs monitoring and computation complexity, especially for companies with almost homogeneous product portfolio and a finite number of production lines (Howell 1987), (Drury 2007).

This paper describes a new integrated approach combining the discrete event simulation with operation costing to assess work order production costs for bid preparation. This approach was also applied on a real case in the wooden painting sector

The next two paragraphs (§2 and §3) shortly report the problem and the objectives of the work. After, a description of the addressed production system (§4) is given, while paragraph 5 presents the overall approach and the resulting tool. In paragraph 6 results and sensitivity analysis are described.

2. PROBLEM DESCRIPTION

The main challenge of the work was to provide a good and reliable estimation of work orders production costs meant to easily and precisely prepare customers' bids.

The small company facing such a problem is a third-party service provider running its business in painting wooden components mainly for furniture, such as doorways for kitchen, bathroom and living room, drawers, window frames, etc.. It receives materials (un-painted wooden boards) from its customers for workmanship only. The finished wooden products are then returned to the contractor.

Fig. 1 Some finished products

This company can paint a broad range of wooden products changing in sizes, geometry and typology, also starting 15-20 new products each year.

Each order, coming from the customers, is formed by different products, varying in:

- Product typology: doorways, drawers, window frames, etc..
- Product sizes: the length as well as the width can range between 50 mm up to 1300 mm;
- Quantity of each product.
- Processing types to be performed to obtain products complying with customer requirements. 9 processing types are offered by the company, each one with some variants amounting to 30 different process sequences.

This wide multiplicity of components, processing types and orders makes the bid preparation a real challenge. The main problems are the order estimation of: indirect costs, lead time (always influenced by many factors), time and number of personnel involved. To get a good estimation of time and resources utilization is fundamental to prepare a reliable and accurate customer bid.

Looking at the last balance sheet of this company, 46% of costs, leaving aside raw materials (that weight for over 40%), are due to manpower involved in the production department, making its correct ascription a *conditio sine qua non* for a good bid estimation.

3. OBJECTIVE OF THE STUDY

The objective was to identify a new approach able to provide an accurate and reliable estimation of work orders production costs meant to ease customers' bid preparation and/or to verify their feasibility. Price is a critical success factor for small third-party service providers: high prices discourage customers, while too low prices abate company profitability. It's really worthwhile to be able to provide a precise production costs estimation as a basis for further decisions to accept or reject an order.

This approach, called *cost simulator,* integrates discrete event simulation with the operation costing, coupling accurate in-depth information on production times and resources with a precise estimation of production costs. Combining these two techniques, it is possible:

- To get accurate cost estimation, even concerning new products.
- To highlight production system bottlenecks and to suggest ways to manage them, thanks to simulation.
- To support awareness creation: both the preliminary analyses performed in order to design the tool and its constant usage and update drive the company towards a formalized representation of internal capabilities and resources.

This approach is useful both for salesmen during the bid preparation as well as for the production responsible to improve production performances.

4. THE PRODUCTION SYSTEM

The production flow of the analyzed small company starts from the warehouse, where components (wooden boards), received from customers, are stocked, goes through the four production departments and finishes in the packaging department, where products are prepared to be delivered.

Each department is equipped with a plant performing peculiar processes. Thanks to this kind of organization, the company is able to work a broad spectrum of wooden products. A more detailed description of the five departments follows:

1. The dyeing department: it's the starting point of the working process. Raw component is dyed and prepared for next processing by means of a dyeing plant with 2 rotary machines with forced dyeing and a warm air ventilation oven for drying the wooden elements.
2. The lacquering department: this plant is able to lacquer flat wooden elements of various dimensions and types as well as to apply transparent finishing on thick pieces.
3. The painting department: elements, showing particular shapes (curved doors, bases and boards), are painted manually using the spraying "Airmix" technology.
4. The finishing department: this automatic plant can produce several finishing in high quantity and quality. It is equipped with an innovative spraying system with piece-dimension reading (meant to reduce over-spraying) and with a product recovery carpet (meant to reduce the impact of emissions in atmosphere).
5. The packaging and tinning department: the finished components are firstly checked for quality control and then packaged with a packaging unit equipped with an automatic paper-box and labelling device.

The working sequence depends on customer requirements and component materials, while component shape has an influence on plant selection. For instance, the painting process can be performed in two departments: flat components are painted in the finishing department while curved ones are painted manually in the painting department. Generally speaking, a component is dyed one or two times, then it's finished and packaged.

Currently, 28 workers are employed full time in the five departments and 5 foremen staff the 5 plants. The main tasks of each foreman are to continuously monitor the plant running operations, to prepare the raw materials for painting devices and, sometimes, to help other workers in making their job on the plant.

The other 23 employees can turn between the different plants based on the production needs performing the following tasks:

- Loading components on the conveyor system and unload finished ones.
- Performing manual operations such as brushing ash doors.
- Cleaning the plant at least once a day.

The number of workers involved in the loading and unloading processes depends on the components weight and length. The heavier and/or longer is a component, the higher is the number of involved workers. Anyway, due to space constraints, a maximum of 3 workers can load or unload at the same time. Furthermore, the load and unload processes in the same plant happen in a spot way: workers perform loading until the plant is full of components and unloading just at the end of the process. In the meantime the plant works automatically and components go across the plant 2 or 3 times in order to undergo many hands of painting, or front and back painting. During this period, the loading team can be allocated to another plant. When the first worked component is available to be unloaded, some workers have to be re-allocated to this plat to start the unloading process and, at the same time, another team re-starts loading components to be painted.

Some other manual processes are made in the same way. In this case the total number of involved workers can range between 2 and 7.

It's easily understandable how the estimation of personnel saturation is very tough but, at the same time, how it heavily influences the production costs. Furthermore, an efficient resource allocation (for example re-allocation between loading and unloading on the same plant) can influence production costs.

5. THE COST SIMULATOR

The cost simulator goal is to calculate reliable and accurate cost estimation of work orders. It's based on the integration between discrete event simulation and operation costing techniques. It is composed by two different models:

- A Cost Model containing the cost information needed to get production cost estimations.
- A Simulation Model, representing all the production processes.

The Simulation Model simulates a work order production flow. Resulting data, dealing with timing and resources utilization, are inputs of the Cost Model in order to estimate costs.

The next two paragraphs report a detailed description of the cost and Simulation Models.

5.1. The Cost Model

Providing reliable cost estimations for assessing a work order production cost is the main goal of the Cost Model.

In order to perform an estimation of production costs, traditional cost accounting procedures (Howell 1987), (Drury 2007) suggest a distinction between direct and indirect costs. The overall production cost for a work order is equal to the sum between costs directly and unambiguously chargeable to the given order and a percentage of costs that are not directly accountable to the particular order. In the here described scenario, given that the supplier holds the ownership of wooden boards, dye, solvents and similar products are the only **direct costs**. The amount of dye for each work order depends on the overall dyed area.

On the other hand, indirect costs estimation appears to be difficult due to the above described peculiarities.

An operation costing approach has been adopted: the only direct (dye) costs are univocally charged to a given work order, while indirect costs are (almost causally, but proportionally) allocated using suitable dimensions. Third party wooden painting production characteristics are consistent with main applicability indications of operation costing (Zuk 1990), (Drury 2007): products with strong similarities yet differentiated in some forms from each other, batch production and variable but discrete production systems.

The Simulation Model gives an estimation of production times and workforce time per work order; therefore indirect costs were split into two groups: **workforce costs** and **other production costs**, with the goal to reach an estimation of unitary (hourly) values for each of the two categories. In fact, each work order is processed in different departments and, on the other hand, each worker is able to operate in more than one department.

Average hourly workforce cost was valued as a ratio between the overall annual production workforce cost (derived from the last final balance) and the annual amount of production hours. Only the 23 ordinary workers salaries were taken into account for this calculation, while the salaries of (5) foremen appointed to supervise each department were left aside.

Average hourly production cost for each department was calculated in a more complex way. Two typologies of costs were defined:

- IFO_DFP (Indirect for the Order, Direct for the Production unit): costs directly accountable for the 5 production departments. The identified classes are:
  (a) The 5 foremen salaries.
  (b) Equipment (and pertinent components) depreciation.
  (c) Equipment leasing/rental and maintenance.
  (d) Equipment leasing/rental.

- IFO_IFP (Indirect for the Order, Indirect for the Production unit). They were:
  (a) Water, energy, gas, compressed air and other costs clearly ascribable to existing equipment and calculated thanks to a pro-quota distribution among the 5 departments.
  (b) Production shed rental, shed heating, training costs, consumables, and other costs partially ascribable to the production activity.
  (c) Overheads.

Going from (a) to (c), the causality of the ascription of the costs to each equipment decreases.

Each cost, belonging to IFO_IFP(a) category, was ascribed to a given department/equipment using a proper ascription driver. For example, the energy cost per hour of equipment was calculated as the annual cost for all the equipments energy multiplied by an "equipment energy weight". It is the ratio between the equipment maximum power multiplied by its annual running time and the sum of the maximum power of each plant multiplied by each plant annual running time.

The overall amount of workforce working hours was selected as a driver for all the IFO_IFP(b) costs. The amount obtained dividing the overall IFO_IFP(b) costs for the overall amount of workforce working hours was added to the workforce hourly cost. This ascription derives from the observation that IFO_IFP(b) costs are strictly related to the number of working hours. This choice drove the authors to consider IFO_IFP(b) unitary costs as an additional charge for the workforce hourly cost.

The overall number of equipment running hours is, finally, the driver adopted for IFO_IFP(c) costs.

Summing up, the cost for a given work order derives from:

$$C(WO(i)) = DC(i) + WHC * WWH(i) + \sum_{j=1}^{5} PHC(j) * PRH(j,i)$$

$$C(WO(i)) = Work\ Order\ (i)\ Overall\ cost$$

$$DC(i) = Work\ Order\ (i)\ Direct\ Costs =$$

$$= Required\ Dye\ Volumes * Dye\ cost\ per\ volume\ unit$$

$$WHC = Workforce\ Hourly\ Cost =$$

$$= \frac{(Workforce\ Annual\ Cost + IFO\_IFP(b))}{Workforce\ Annual\ Working\ Hours}$$

$$WWH(i) = Workfoce\ Working\ Hours\ for\ Work\ Order\ (i)$$

$$as\ calculated\ by\ the\ simulation\ model$$

$$j = 1...5 = equipments - departments$$

$$PHC(j) = Production\ Hourly\ Cost\ for\ equip.\ j =$$

$$= IFO\_DFP(j) + IFO\_IFP(a)(j) + IFO\_IFP(c)(j)$$

$$PRH(j,i) = Production\ Running\ Hours\ of\ equip.(j)\ and$$

$$for\ Work\ Order\ (i)\ as\ calculated\ by\ the\ simulation\ model$$

The Cost Model is MS-Excel-based and interacts with an already existent internal accounting platform adopted by the company, automatically importing and re-grouping required inputs. The company account manager validated the above described Cost Model.

5.2. The Simulation Model

The main goal of the Simulation Model is to assess processing times needed to work a customer order. So it has to reproduce all the material production flow.

For this real application, it was decided to simulate only 3 departments:

1. The dyeing department.
2. The finishing department.
3. The packaging department.

In fact, examining all the work orders of the last 6 months, and interviewing the company team manager, it comes up that about 70%-80% of orders are worked only in these two departments:

The first step was to analyse and to map all the production processes involved in working an order, from its reception to its delivery, considering also the reworked process. However, the defectiveness rate is very low, so the rework process was not included in the model.

In contemporary, a deep analysis on the orders was carried on with the objective to identify a set of pilot orders that will be used to test and validate the Simulation Model as well as the cost simulator.

Afterwards, the two departments along with the packaging one were analyzed in a deeper way. All the technical and operational data were gathered and a particular attention was kept to collect information on work orders management. A list of all the feasible operations for every plant was defined and, for each one, some key parameters were fixed such as: conveyor belt velocity, number of workers, number of complete turns in a plant, etc. All these data were coded and put in an Excel file which is the input file driving the Simulation Model.

All the examined plants are loaded manually and, because of its impacts on the system performances, the loading process was analyzed deeply. Workers involved in this process have to comply with some rules and constraints but, at the same time, they should put on the conveyor belt as many components as possible in order to maximize the system throughput. For instance, it's very fundamental to leave some room between close components loaded on to the conveyor belt to assure a uniform paint layout on each part and reduce the risk of future rework operations. Currently the loading process is made by workers without any kind of support: neither defined and written procedures nor software tools to optimize components placement are available. Workers decide the loading sequence time by time based on their past experiences, on how all the components have been piled and avoiding to mix different orders.

In the analysis phase, the loading process was formalized and coded in a computer model with the help of Excel. The goal was not to find out the optimum loading solution considering a work order, but a feasible

sequence complying with the actual rules and technical constraints. In the Excel model, all the main constraints were defined and, by means of some formulas reflecting the actual rules, a loading sequence is calculated starting from any kind of work order. Furthermore it was defined a key indicator called "area efficiency" meant to measure the ratio between the square meters taken by components loaded on the conveyor belt (the brown area in Fig. 2) and the available square meters (the yellow area in Fig. 2). Likely, the system throughput should increase, raising the value of this indicator. So a first level of optimization was implemented in this model selecting the solution with the best key indicator value.

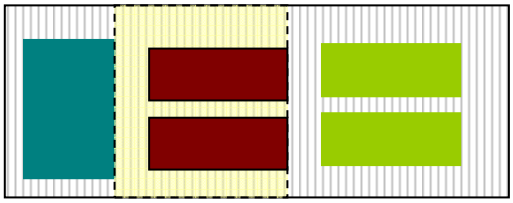


Fig. 2 The component placement in the loading process

The Simulation Model was developed in Arena® (Kelton et al. 2007). At the beginning of each run, and for every department, the model reads the input file, in Excel format, with the following information:

1. The components loading sequence.
2. The machining type to be performed.
3. The operational data.
4. The number of workers needed for each manual job, especially for the loading and unloading task.

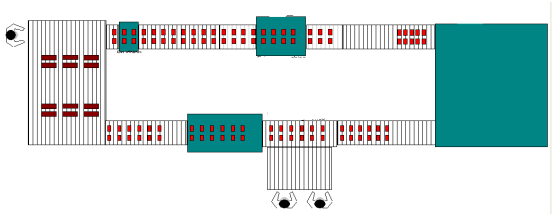


Fig. 3 The finishing department layout

The model simulates all the working processes needed to paint, finish and packaging customer orders recording the time frame in which workers have been busy. At the end of the simulation, the model records the following data in an output file in Excel format:

1. The total number of man hours taken to perform each manual task in each work center.
2. The total number of man hours taken to work the order in each plant.

The Simulation Model also keeps into consideration the set up time of each plant when the responsible prepares the raw materials (dyes, solvents, etc.) to paint or finish the components and makes all the procedures to clean the spraying machines and/or change some tools.

From the Cost Model, the Workforce Hourly Cost (WHC) and each Production Hourly Cost for every equipment (PHC(j)) are imported in the output file in order to calculate the production costs of the simulated order (Work Order Overall Cost).

Furthermore, the model generates an output report with some key indicators such as:

1. The average hourly throughput of each plant measured in square meters per hour.
2. The drying oven saturation measured as the ratio between the square meters occupied by components and the total square meters available in the oven.
3. The average lead time measured in second.
4. The average Work-In-Progress measured in square meters.

Such a report provides the foundation to make result analysis to achieve some system improvements. In addition, using the information generated by the simulation about the time in which human resources have been busy, an allocation plan can be identified in advance. As mentioned above, an efficient allocation plan can help in saving money and time.

6. VALIDATION AND RESULTS ANALYSIS

The validation is the process of determining whether a simulation model is an accurate representation of the system for the particular objectives of the study (Law 2001). This process usually takes a long time and it's not as simple as it might appear, thus it was decided to carry on the validity assessment only with the help of process experts from company.

The validation method was as following: first the simulation model was tested comparing the flow time, (i.e. the total time taken by a work order in a department to be completed) against some historical data; then the manpower time was validated.

During the first test, the simulation model was refined few times to add minor details influencing the results, and, in the end, it reached a high level of accuracy. The difference between simulation and historical data was about 5%, which is adequate for this kind of project.

The manpower time comparison was more complicated because simulation results were not so close to the historical data as it was expected. Further investigations were made by direct observations: the simulation model estimations were accurate. The actual method to assess workers saturations, based only on data recorded by the same workers, was often not correct, thus generating bad estimations.

6.1. Sensitivity analysis

Once the validation phase was ended up, the Simulation Model was also used to estimate the impact of some key factors on the system throughput. Such estimation was

performed only for the finishing department because it has the highest annual utilization rate.

A sensitivity analysis was carried on the following two key factors:

1. Area efficiency during the component loading process.
2. Plant saturation.

The close relationship between these two factors and the throughput is already known, but the scope was to quantify the effect. The results are described in the next two subparagraphs.

6.1.1. Area Efficiency analysis

The main goal of this analysis was to evaluate the impact of the components placement on the plant performances.

As already mentioned, components have to be loaded onto the plant complying with some rules, constraints and, at the same time, trying to fill in the available loading area as much as possible. The more square meters are painted per hour, the higher is the productivity.

This sensitivity analysis was performed on a set of hypothetical orders composed by only one product with different dimensions, in order to get different values for the key indicator "area efficiency". As already said, it measures the ratio between the painted square meters and the ones available on the conveyor belt. The considered set is composed by 50 samples. They are all worked with the same process and the total number of square meters to be painted is similar and great enough to avoid the effect of the warm up period in the simulation. Furthermore, it was assumed that the loading time is not related to the number of components to be placed on the conveyor belt each time, but only to the components length. It means loading 2 or 3 components with the same length takes the same time. It's not such a stringent assumption, also because it actually happens in the real production process, where the number of people charged of loading components, depends on components length and weight.

From a first analysis, the area efficiency indicator value is directly related with the component width, assuming its length as fixed (see Fig. 4). In other words, increasing the components square meters, the area efficiency indicator and the system throughput raise up.

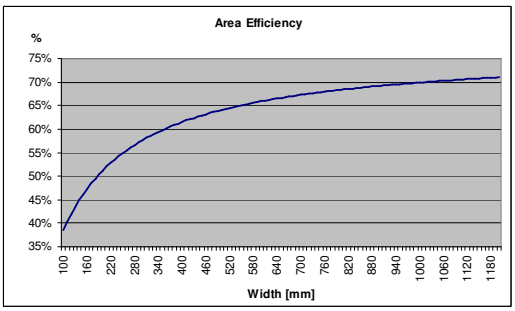


Fig. 4 Area efficiency vs. component width

But this direct relation is not valid any longer when considering components where both dimensions are different. The chart in Fig. 5 shows the average area efficiency value for each class of components (a class is a collection of components with the same length but with different widths).

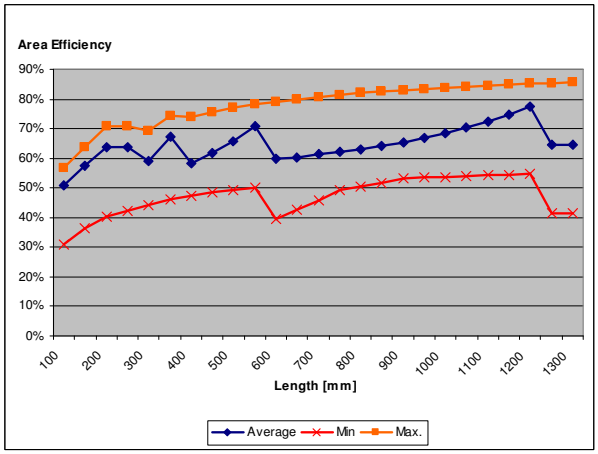


Fig. 5 Area efficiency vs. components classes

The simulation results confirm the close relationship between the area efficiency and the system throughput. The correlation value is about 0.95, a quite high value. The linear regression is as follows:

$$\text{Throughput} = 219.36 * \text{area efficiency} - 11.78 \qquad (1)$$

However, carefully analyzing the results, for instance looking at indicator values close to 62% (see chart in Fig. 6), the productivity ranges from 119 to 138 square meters per hour. It is due to the different length of the worked components. A new linear regression was calculated keeping into consideration also this factor and the formula is as follows:

$$\text{Throughput} = 218.16 * ae + 0.008 * cl - 13.01 \qquad (2)$$

ae= area efficiency
cl =component length

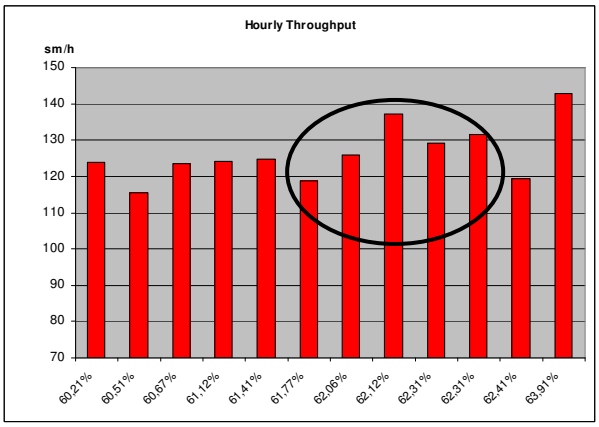


Fig. 6 Throughput vs. area efficiency

The graph in Fig. 7 compares the throughput generated by the simulation and the value calculated with the linear regression.

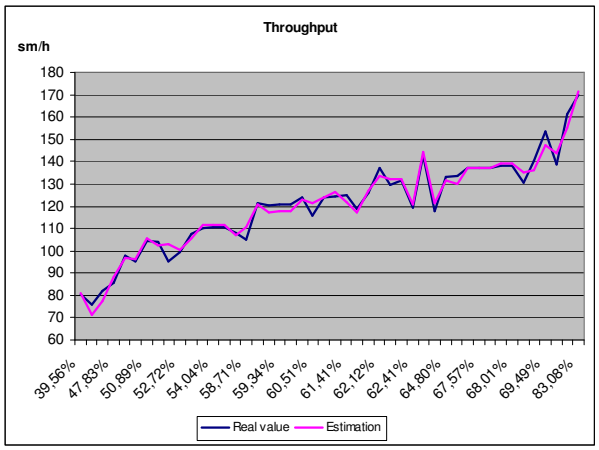


Fig. 7 Area efficiency vs. throughput

To be statistically sure about this result, a validation hypothesis test was performed with the null hypothesis being the two parameters (area efficiency and component length) have no effects on the throughput (Chung, 2004). For a two-sides test at a common level of significance α =0.05, the critical values from the *t* distribution are -2.017 and 2.017. The calculated *t* for area efficiency and component length are respectively 31.3 and 10.7. They both exceed the above values and so the null hypothesis is rejected. Area efficiency and component length have an impact on system performances.

This analysis highlights how an efficient loading sequence can increase the plant throughput: moving from 50% to 60% of area efficiency means to gain about 20% of productivity.

6.1.2. Plant saturation analysis

Generally, components go across the finishing plant 2 or 3 times. Clearly this plant has a fixed capacity and if the square meters of an order exceed its capacity, the work order has to be split into smaller batches whose capacity can fill up the plant. It means that while the first part is unloaded, the second part of the same order can be loaded into the plant.

Looking at the different components types and dimensions worked in this plant, its capacity, under normal working conditions, can range between 300 and 500 square meters.

The main goal of this analysis was to evaluate the impact of plant saturation on the hourly throughput. In order to keep the analysis as understandable as possible, it was assumed to work orders with only one component but the total square meters to be painted changes in each trial. The area efficiency value is close to 70% and the processing type requires working both front and back part, meaning to cross all the plant 2 times.

As expected, there is a close relationship between plant saturation and system productivity: the correlation value is near 1. The throughput is sensible to the order quantity that's the smaller the quantity to be worked, the lower the hourly productivity. It can cut down up to 19% considering a small order (see Fig. 8). But also considering very big orders, with 2 days of lead time, the productivity value can decrease from 3% to 6%.

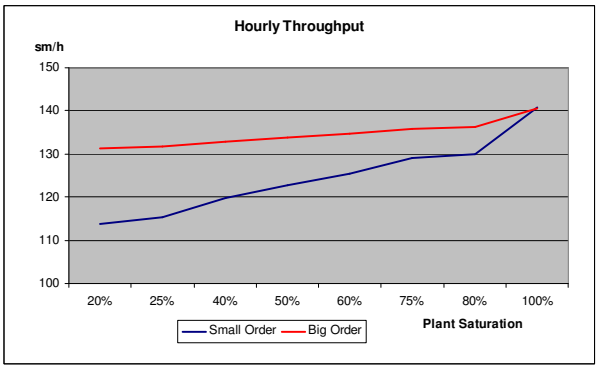


Fig. 8 Throughput vs. plant saturation

These results suggest the idea to try working with the plant always full of pieces, joining different work orders, when possible, in order to keep the plant as saturated as possible.

7. CONCLUSION

The cost simulator, based on the integration between the simulation and the operation costing models, was able to provide more accurate production costs for historical orders than the existent method merely based on past experiences. Furthermore, this calculation takes few minutes, even for big orders.

The Simulation Model can be used also to define feasible resource allocation plans as well as to find out bottlenecks or critical aspects. For the analyzed company, the loading process is very critical because it has a strong impact on the throughput.

Lesson learnt is that the cost simulator adds real values when it's very hard to estimate orders costs due to their big variability and the evaluation of production times and resources utilization is difficult. These topics are quite common in many small and medium third-party services.

Next steps will be the completion of the Simulation Model to represent all the departments in the shop floor and to improve the integration between the cost simulator and the existing software tools meant to make easier for the account manager to add new costs and correctly map them in the Cost Model.

REFERENCES

Boër, C.R., Avai, A., Imperio, E., El-Chaar, J., 1993. Computer simulation for the design and planning of flexible assembly systems. *International workshop on "Application and development of modeling and simulation of manufacturing systems"*, pp.84-98. Beijing.

Chung, C.A, 2004. *Simulation modelling handbook*, Crc Press.

Diaz, M.L., Buxuman,n P., 2003. The value of cooperative planning in supply chains: a simulative approach. *Proceeding of the 11th European conference on information systems* (*ECIS 2003*). June 16-21, Naples (Italy).

Drury, C., 2007. *Management and Cost Accounting*, Edition: 7, illustrated, *Cengage Learning* EMEA, 2007.

Gahagan, S.M., Hermann, J.W., 2005. Finding the optimal production control policy using the production control framework. *In Proceedings of the 2005 Winter Simulation Conference,* pp. 1418-1427. December 4-7, Orlando (Florida, USA).

Howell, R.A., Soucy, S.R., 1987. Cost Accounting in the New Manufacturing Environment *Management Accounting*, p 42. August 1987

Kelton, W.D., Sadowski, R.P., and Sturrock, D.T., 2007. *Simulation with Arena, 4th Edition*, WCB/McGraw-Hill, New York, NY.

Law, A.M., McComas M.G, 2001. How to built valid and credible simulation models. *In Proceedings of the 2001 Winter Simulation Conference,* pp. 2577-2584. December 9-12, Arlington (Virginia, USA).

McLean, A.A., Biles, W.E., 2008. A simulation approach to the evaluation of operational costs and performances in liner shipping operations. *In Proceedings of the 2008 Winter Simulation Conference.* pp. 22-29. December 7-10, Miami (Florida, USA).

Michalek, J.J., Ceryan, O., Papalambros, P.Y., Koren, Y., 2005. Manufacturing investment and allocation in product line design decision making. *In proceedings of IDETC/CIE 2005*. September 24-28, Long Beach (California, USA).

Murray, S., Mackulak, G.T., Fowler, J.W., Colvin, T., 2000. A simulation-based cost modelling methodology for evaluation of interbay material handling in a semiconductor wafer fab. *In Proceedings of the 2000 Winter Simulation Conference.* pp. 1510-1517. December 10-13, Orlando (Florida, USA).

Roser, C., Nakano, M., and Tanaka, M., 2003. Buffer allocation model based on a single simulation. *In Proceedings of the 2003 Winter Simulation Conference*. pp. 1230-1246. December 7-10, New Orleans (Louisiana, USA).

Von Beck, U., Nowak, J.W., 2000. The merger of discrete event simulation with activity based costing for cost estimation in manufacturing environments. *In Proceedings of the 2001 Winter Simulation Conference.* pp. 2048-2053. December 9-12, Arlington (Virginia, USA).

Voorhorst, F.A., Avai, A., and Boër, C.R., 2007, "Optimizing a highly flexible shoe production plant using simulation", *Proceedings of 7th international workshop on modelling & applied simulation,* pp. 233-239. September 17-19, Amantea (Italy).

Wang, Y., Kao, Y. Su, C. 2001. Simulation analysis of an integrated cross-factory process: a case study. *In proceedings of European Simulation Symposium*, October 18-20, Marseille (France).

Zuk, J.S., 1990. Effective cost modeling on the factory floor: taking simulation to the bottom line (Panel). *Proceedings of the 1990 Winter Simulation Conference*. pp 590-594.Dece,ber 9-12, New Orleans (Louisiana, USA).

AUTHORS BIOGRAPHY

Antonio Avai is a partner and technical director of Technology Transfer System, an IT company located in Milan. He has managed research programs with focus on methodologies and leading information technologies to support all the life cycle of manufacturing processes. He has more than 15 years of experience in discrete event simulation and its integration with other software tools and technologies, and has authored several papers on these topics.

Marzio Sorlini graduated in Management Engineering at Politecnico di Milano with major in strategy and innovation management, he worked as junior researcher at ITIA (Institute of Industrial Technologies and Automation) being engaged in (5th and 6th FP) EU-funded projects on innovation management, decision support systems and innovative cost accounting methods for integrated product-service solutions development. He's currently employed at SUSPI-DTI (University of Applied Sciences and Arts of Southern Switzerland – Department of Innovative Technologies) and co-operates with Technology Transfer System.

APPLICATION OF RIGID BODY DYNAMICS TO 3D PLANT SIMULATION

I.F. Ceruti[(a)], G. Dal Maso[(b)], D. Rovere[(c)]

[(a)] Technology Transfer System s.r.l. - TTS
[(b) (c)] University of Applied Science of Southern Switzerland - SUPSI

[(a)] ceruti@ttsnetwork.com; [(b)] giovanni.dalmaso@supsi.ch; [(c)] diego.rovere@supsi.ch

ABSTRACT
Nowadays the interest for simulation applied to both manufacturing systems and plants is rapidly growing. In this application the classic approach shows its limitation due to the inability to manage objects whose position is the result of interaction with other environment bodies.

The will to overcome these limitations without losing any of the benefits acquired with the introduction of simulation approach, leads to the introduction of rigid body dynamics into the simulation.

The limitations of the classic approach, a description of the novel approach and the results obtained on a real CNC system are the bricks composing this paper.

Keywords: manufacturing simulation, 3D simulation, physical simulation, CNC driven simulation

1. INTRODUCTION

The importance of computer animated simulations is ever increasing both in the design and in the production phase of manufacturing systems.

A comprehensive simulation, when applied to an entire manufacturing plant, must be capable to represent various aspects of the reality. Most common features, found in current 3D simulation engines, are direct-inverse kinematics solution and collision detection. More advanced features cover other process aspects like material removal and sensors simulation.

This kind of simulation is based on the assumption that relevant body positions are known as a result of defined motion paths, influenced only by kinematic relations (the object position is an input state). As mentioned, with this approach only the kinematics relationships are taken into account and no physical constraints like contacts or gravity influence the objects' position calculation.

This represents a limit especially in presence of free bodies whose motion is not governed by an analytic law but depends on the interaction with other geometries and on physical properties like friction. In kinematics simulation the problem is solved animating the part along theoretical paths (e.g. gravity based feeder).

It is to be noted that current vision in Digital Representation of manufacturing process moves towards an "adherent to reality" representation, as highlighted in the "Manu*future* Strategic Research Agenda" (European Commission 2004). The classic simulation clearly fails under this point of view.

This paper presents a new approach meant to integrate the rigid body dynamics aspects into the plant simulation. The purpose is a paradigm shift where the previous "animated" models in which the system state is known "a priori" is substituted by a real-time physics simulation model where interactions between geometrical entities and physical constraints influence the time evolution of their position.

The benefits of our approach have been verified developing the model of an automated CNC drilling and sawing system. More advanced topics, such as the integration of the simulation with the CNC logic has been explored.

The paper outline will be the follow: after a presentation of classic computer aided 3D simulation approach, the novel rigid body dynamics approach is presented. Section 5 presents a conclusive analysis of the application to a real industrial system.

2. PREVIOUS WORK

Virtual Reality (VR) and simulation in the manufacturing lifecycle as high value adding tools for cost-effective and rapid creation, management and use of the Next Generation Factory have been presented in (Pedrazzoli 2007).

These tools can be introduced as decoupled modules at both Product design and Factory design level of the product/process life cycle. The research done towards such tools has shown that these are a powerful way to gain flexibility in CNC machines as described in (Mancini 2004).

Virtual Reality simulation, both applied at machine and plant level, is a recent research area: despite this it is based on known fields such as real time collision detection (Kockara 2007) and physically based object behaviour (Baraff 2003).

Various applications can be found in literature describing the benefits of using collision detection techniques for simulation purposes. Besides application to generic 3D environments (Bergen 2003), there has been an extensive research on application of collision detection to robotics (Steinbach 2006; Kuffner 2002; Okada 2006). Recent studies show application of such engines to wider manufacturing applications (Ceruti 2008).

Attempts to take advantages from a physically based simulation can be found in (Loock 2001). Together with collision detection, physics simulation (in terms of gravity effects on objects and physical based cable simulation) is used to validate assembly tasks procedures.

In (Glencross 2001) a more integrated and complete framework has been proposed to simulate the interaction of the objects in a VR environment.

Other authors (Carpin 2007; Greggio 2006; Garber 2002) use physical based simulation for the motion planning of rigid and articulated robots.

In such cases benefits come from the ability to compute better results if compared to traditional methods, when there is an interaction in a complex environment with moving objects.

3. CLASSIC SIMULATION APPROACH

Computer aided simulation applied to manufacturing plants combines the typical approach of a three dimensional viewer with the peculiar information used to represent mechanical data of a manufacturing layout.

The simulated environment mainly deals with mechanical components, production items and production support structures.

3.1. Scene-graph

The classic simulation approach is based upon a hierarchical data structure called scene-graph. This structure is a directed acyclic graph (DAG) in which nodes represent simulation entities and edges represent their positional relationships. In fact, at a higher abstraction level, objects taking part to the simulation can be considered as pure reference systems (XYZ-O) with an associated collection of properties.

In modular software architectures, engines taking part to the simulation (Pedrazzoli 2001) populate this collection storing their own customized data structures for their purposes. Examples of properties are meshes used by the 3D visualization engine, collision structures used by the collision detection engine and *voxels* data used by the real-time material removal simulation engine.

Relationships between nodes of the DAG represent relative geometrical displacements between objects' reference systems and can be expressed in terms of homogeneous transform matrices. Therefore, given two nodes i and j representing two simulation entities, the edge connecting them is fully described by the matrix

$$H_j^i = \begin{bmatrix} R_j^j & \boldsymbol{t}_j^i \\ \mathbf{0} & 1 \end{bmatrix} \tag{1}$$

where R_j^i represents the rotation part of the transform and $\boldsymbol{t}_j^i$ the translational part.

In plant simulation these transform matrices can be controlled either *directly* or *indirectly*. The control is *direct* when the absolute object's reference system 3D position and rotation is set and then relative transform are re-calculated.

An *indirect* control is obtained using mechanical joint models (rigid, rotational, translational) which constraint the motion of the children relatively to the parent position according to some setup parameters (like rotation/translation axes) and the runtime value of their joint variables (rotation angle or translation). In this case transform matrices $H_j^i(\boldsymbol{x})$ become functions of a set of joint parameters.

3.2. Simulation loop for CNC driven system

In the classic approach, running a simulation actually means animating the scene-graph nodes modifying the current absolute position, orientation and states of the objects composing the environment model.

This animation can be directly controlled by a source capable to generate inputs for the simulation model. In a CNC-driven simulation, the CNC computes the values of the joint variables that are fed into the simulation as state inputs through a proper communication interface.

These inputs activate the typical simulation loop (Figure 1): each input modifies some environment variables, triggers the simulation engines and brings the simulation into a new state.

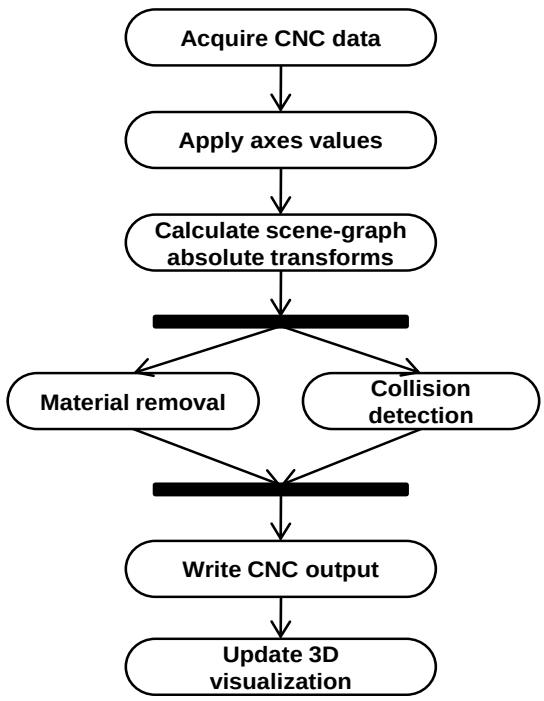


Figure 1 : Simulation loop for CNC driven system activity diagram.

Starting from the root node, walking through the scene-graph and taking into account the current type and value of the joints, the position and orientation of each node is updated.

Once the final state of each node is computed, simulation engines can perform their work accessing and modifying node's properties. As an example

consider the real time material removal engine that checks for geometry intersections between tool nodes and modifiable objects, and the collision detection engine that checks for contact points between objects provided with collision structures (Ceruti 2008).

The results produced by simulation engines can be directed both internally (e.g. other engines) and to the CNC (as feedback data) closing the communication loop. The collision detection engine, for example, produces a list of colliding nodes that can be sent both to the 3D visualization engine which could highlight the objects for easy contact identification, and to the CNC software to produce a warning message for the user and possibly debug the CNC logic.

3.3. Benefits and limitations in plant simulation

The aforementioned simulation brings many benefits to the end user along the plant/machine lifecycle. In the design phase, testing CNC logics with a simulated machine instead of using a real prototype is cost saving since:

- the risk of costly machine breakings is totally eliminated
- motions can be tested at real feed rates (while normally low speeds are needed in order to prevent unforeseen collisions between parts of the equipments), thus reducing the testing time
- simulation can be run at accelerated time with batch processes enabling a "what if" analysis on the CNC logics and providing the means for optimization

Moreover the 3D visualization of the plant part motions is unparalleled when used during the debug phase of new complex sequences of operations, like tool change procedures can be.

3D simulation can be also used during the production phase of the plant. In fact the developed models can be exploited for remote monitoring applications, providing that most of the modern CNC devices support network communication protocols. In this case the cost saving is due to the fact that it is possible to perform remote diagnosis of breakdowns thus reducing (when not totally avoiding) the duration of technical personnel interventions.

This classic approach to manufacturing plants simulation has however some limits which are direct consequences of the fact that the motion animation of each part of the system is directly under the control of the input source. In fact, though it is easy to handle well known and predictable movements, like a mechanical axis positioning, it is almost impossible to manage motions resulting from objects interaction, like contact forces, friction, and so on.

A clear example of this limitation is represented by the handling of the piece to be manufactured: this is not constrained by any mechanical joint to the plant structures and its position in the system is not under the direct control of the CNC; instead it is the result of the interaction with other objects in the scene like handling devices (grippers) and transportation elements (roller ways, chains, magnets). Moreover, when pieces are queued on the feeders it becomes essential to model the interaction between free objects.

The fact that classic simulation fails to manage these common situations highlights the need for the new approach described in the following section.

4. RIGID BODY DYNAMICS APPROACH

The simulation of the motion of a rigid body is based on the motion of a particle. The rigid body can be seen as a system of particles where the distance between any two given entities remains constant in time regardless of external applied forces. A rigid body is thus non deformable but this limitation is not a problem as we are not interested to model material deformations.

4.1. Motion of a particle

If we denote $x(t)$ as the particle location in world space at time t, $v(t) = \dot{x}(t) = \frac{d}{dt}x(t)$ gives its velocity. We can thus define the state vector for the particle as:

$$Y(t) = \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} \tag{2}$$

We want to know how the state of the particle changes over time, i.e. we want to calculate:

$$\frac{d}{dt}Y(t) = \frac{d}{dt}\begin{pmatrix} x(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} v(t) \\ F/m \end{pmatrix} \tag{3}$$

where the relation $\frac{d}{dt}v(t) = F/m$ is derived from Newton's second law. Thus we have to know the forces applied to the particle and its mass in order to solve the system.

4.2. Rigid body dynamics

Unlike mass particles, that have a finite mass but zero volume, rigid bodies have a considerable volume and have geometrical properties. At any instant the position of a point p_0, measured in body space, is given by:

$$p(t) = R(t)p_0 + x(t) \tag{4}$$

where $R(t)$ is the rotation and $x(t)$ is the translation of the body. To extend the definition of the state vector $Y(t)$ for a rigid body, we now need both $x(t)$ and $R(t)$. Similarly the motion of the body can be described by the linear momentum $P(t) = m\,v(t)$ and the angular momentum $L(t) = I\,\omega(t)$ as shown in (Shabana 2001).

To compute the evolution of the state $Y(t)$ it is no more sufficient to know the mass m, we also have to provide the values of the inertia tensor I. The solution is expressed in terms of the force $F(t)$ and of the torque $\tau(t)$ applied to the rigid body.

Let us consider a simple system composed of two rigid solid bodies. If we exclude the case of inter-

penetration, when two bodies are in contact their motion is constrained and we can have two situations:

- *Colliding contact*: when the velocity relative to each other is non zero. To prevent inter-penetration we must have an instantaneous change of velocity vector.
- *Resting contact:* when the relative velocity is zero.

We enforce these non-penetration constraints by computing appropriate contact forces between contacting bodies and then applying them to each body.

The methods for resolving these constraints, as well as extending the computation to the case of *n* contacts, can be found in (Bender 2006). To model the response of real world object to contact we have to assign at least the following coefficients to the body:

- Restitution coefficient: $0 \leq \epsilon \leq 1$, takes into account how the velocity changes after a collision. $\epsilon = 0$ means no bouncing at all while $\epsilon = 1$ means completely elastic contact.
- Coulomb friction coefficient: $\mu \geq 0$. It is relative to the sliding motion of two surfaces in contact. If $\mu = 0$ there is no friction, while if $\mu = \infty$ the contact is considered to be sticky.

4.3. Changes in the data model

The physics based approach to simulation involves a significant change in the modelling phase. In fact data required in classic simulation is still needed and it must be integrated with the information related to the physics collisions geometry, mass properties and contact parameters. It is possible to summarize the main changes in the data modelling as follows.

4.3.1. Creating the collision geometries

All the objects taking part to a physics based simulation have to be provided with a collision geometry that is used by the physics engine in order to compute contacts details (e.g. contact points and normals).

These geometry representations are usually far more simplified than the ones used for visualization purposes. This simplification is a required phase to lower the computational complexity: drawing N models is $O(N)$ while collision detection is $O(N^2)$.

4.3.2. Assigning physical properties

Modelling a physical simulation entity requires the definition of material related properties like mass values and distribution, restitution and friction coefficients and joints max forces and torques.

This aspect of the physical modelling requires a particular care because it is not always easy to find the real values of these parameters. Whenever possible they are extracted directly from the documentation of the machines and equipments as maintenance manuals and CAD drawings. This usually happens for mass properties and joints forces, while friction and restitution coefficients are set to the values available in literature and then tuned according to the simulations results.

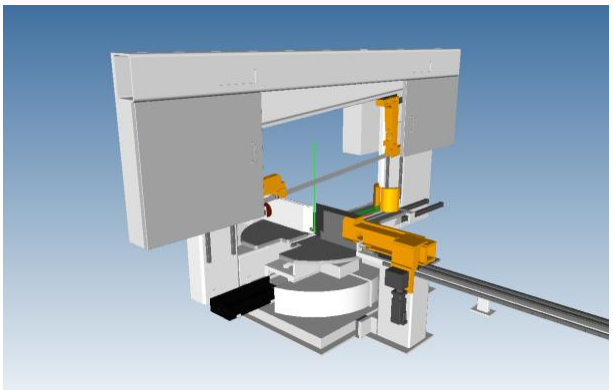


Figure 2 : A sawing machine visual model

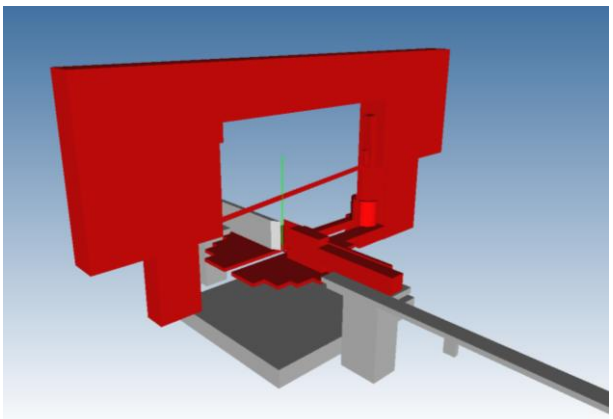


Figure 3 : Sawing machine physical model: a visual representation of its collision geometries

4.3.3. Tuning the simulation model

A big challenge in the modelling is represented by the simulation model tuning phase which is addressed to grant a high level of adherence to real system observed behaviour.

Since the model represents an approximation of the real system, it is often necessary to tweak the physical properties of the entities in relation to how they interact with each other in order to obtain meaningful results.

According to the target results it is always possible to increase the accuracy used to model some parts in order to locally improve the realism of the simulation while maintaining the quasi real-time responsiveness of the model.

4.4. Changes in the simulation loop

With the introduction of rigid body dynamics, the simulation loop described in section 3.2 has to be changed in order to take into account the physics engine presence.

The positional values coming from the CNC cannot be directly applied to joints; instead they are used to compute the velocity vectors of the bodies. These values are fed to the physics engine that computes the updated final state $Y(t)$ for all the objects in the model.

The translation and rotation part of the state is then applied to each node of the scene-graph and the simulation loop can continue with simulation engines triggering, like surface modification and collision detection.

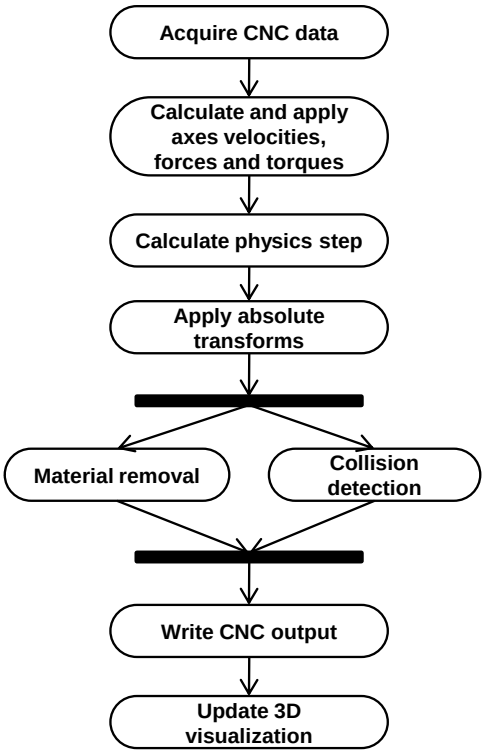


Figure 4 : Physic simulation loop for CNC driven system activity diagram

4.5. Benefits and limitations

The additional benefits of using physics based simulation are related to the possibility to manage the piece and to simulate a wide range of unpredictable behaviours. Moreover, the usage of a wrench based physics engine solver enables the simulation of complex closed links without the need to provide joints with closed form equations.

The limitations of the rigid body dynamics approach can be summarized as follows:

- *Increased modelling effort*: as section 4.3.1 shows, physics simulation needs simplified versions of the visual model geometries which are often replaced with a set of primitive volumes; this requires an extra modelling effort and time.
- *Risk of introducing model instabilities*: unbalanced physics properties (mass and forces) and constraints introduced while modelling the simulation environment can result in numerical instabilities of the physics step calculations.
- *Computational workload*: the introduction of the physics step (Figure 4) in the simulation loop requires a computational effort strictly related to the model complexity, which directly translates into higher memory and CPU requirements.

5. RESULTS

5.1. CNC drilling and sawing system

The case study layout is a heavy carpentry line for the manufacturing of steel beams and is composed by a drill unit and a saw unit connected by an automatic handling system.

The typical argument against having a drill and saw in tandem (side by side layout) has always been that if the drill is working, the saw is idle and vice versa. Even with this objection, the tandem approach has proven to be the ideal solution for many fabricators as the reduction in required plant space, the need for only one operator and the lower investment cost has proven to be the optimum solution for many facilities.

The CNC software engineers have developed a *multi tasking* solution that can be used in the majority of the applications to simultaneously drill and saw the profile. This multi tasking solution requires a more complex CNC logic to handle the sequence of operation needed to complete the tasks.

This layout is a suitable test case since the beam position is the result of complex geometrical interactions with both moving and steady elements of the handling system and not a CNC driven variable.

This is a typical case in which traditional simulation fails due to the high number of embedded logics required to model the behaviour of system. On the contrary, with the proposed approach only intuitive model parameters like friction coefficients, masses, forces and torques have been provided to the simulation. The obtained result is a simulation in which the CNC is reacting to unpredictable sensors signals, in the same way as it was connected to the real devices.

The interface between the application and CNC has been developed using a proprietary communication protocol based on TCP/IP.

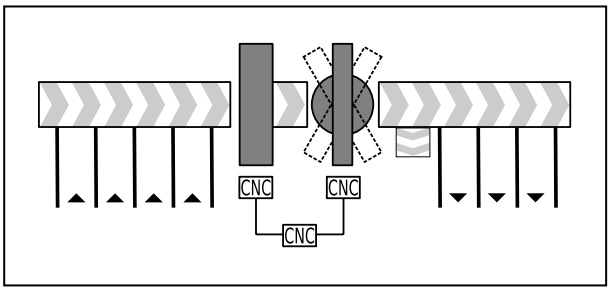


Figure 5 : The drilling and sawing system schema

5.2. Simulation Performances

The developed software has been based on the DDD Simulation Libraries (Ferrarini 2008) integrated with the ODE C++ physics library. Tests have been executed on a laptop equipped with an Intel® Core™ 2 Duo CPU T9300 - 2.50GHz, 3069 Mb RAM and a NVidia®

GeForce™ 8600M GT graphics card with 256 Mb of dedicated video memory.

The CNC used is a Mitrol® Minosse™ control unit for Ficep® lines integrating a PLC (IEC6-1131) at 2ms and 10ms and a machine axes control unit running both in asynchronous (Profile Position) and in synchronous mode (Interpolated Mode). Both 3D models of the line and CNC have been kindly provided by Ficep® S.p.A.

CPU, Memory and max FPS results have been collected using a beam and machines axes motion simulation. Model visual size is approximately 404K triangles. The physics calculations step has been set to 1ms, while the read/write interval from the CNC has been set to 10ms.

Table 1: CPU and Memory usage comparison between the two approaches (CPU usage is an estimated average value)

Simulation Results			
	CPU	RAM	FPS
Kinematics	15%	41 Mb	212
Physics	75%	53 Mb	33

The results show a significant impact of the computational workload of the physics engine on the CPU usage and a limited increment of occupied memory. Nevertheless, the resulting frame rate (FPS) value shows that the simulation remains interactive.

Figure 6 : Test system: laptop pc and CNC

REFERENCES

Baraff, D., Witkin, A., Anderson, J., Kass, M., 2003. Physical based modelling. *SIGGRAPH Course Notes*.

Bender, J., and Schmitt, A., 2006. Constraint-based collision and contact handling using impulses. *Proceedings of the 19th international conference on computer animation and social agents*, pp. 3-11, Geneva (Switzerland).

Bergen, G., 2003. Collision Detection. Interactive 3D Environments. *Interactive 3D Technology Series*, Morgan Kaufmann ed.

Carpin, S., Lewis, M, Wang, J., Balakirsky, S., Scrapper, C. 2007. USARSim: a Robot Simulator for Research and Education. *Proceedings of the IEEE 2007 International Conference on Robotics and Automation (ICRA)*, October.

Ceruti, I., Dal Maso, G., Rovere, D., Pedrazzoli, P., Boër, C.R., 2008. Real Time Collision Detection for Complex Simulations Based on Hybrid Multi Resolution Approximation of CAD Models. *The International Workshop on Modelling & Applied Simulation*, pp 225-232. September 17-19, Amantea, Italy

European Commission, MANUFUTURE – a vision for 2020, November 2004

Ferrarini, L., Veber, C., 2008. 3D graphic simulation of flexible manufacturing systems with Day Dream Daemon and 3DCreate. *Proceedings of 6th International IEEE Conference on Industrial Informatics (INDIN2008)*, pp. 1401-1406. July 13-16, Daejeon, Korea

Garber, M., 2002. Constraint-Based Motion Planning for Virtual Prototyping. *Proceedings of ACM Symposium on Solid Model and Applications*.

Glencross, M., and Howard, T., and Pettifer, S, 2001. Iota: An Approach to Physically-Based Modelling in Virtual Environment. *IEEE Virtual Reality Conference*; pp. 287-288.

Greggio, N., Silvestri, G., Antonello, S., Menegatti, E., Pagello, E., 2006. A 3D model of humanoid for usarsim simulator. *Proceedings of the First Workshop on Humanoid Soccer Robots*, Dec.

Kockara, S., Halic, T., Iqbal, K., Bayrak, C., Rowe, R., 2007. Collision detection: a survey. *IEEE International Conference on Systems, Man and Cybernetics*, pp 4046-4051

Kuffner, J., Nishiwaki, K., Kagami, S., Nishiwaki, K., Kagami, S., Kuniyoshi, Y., Inaba, M., Inoue, H., 2002. Self-Collision Detection and Prevention for Humanoid Robots. *Proceedings of the IEEE International Conference on Robotics and Automation*, pp 2265-2270

Loock, A. and Omer, E. S., and Stadtwald, I. 2001. A Virtual Environment for Interactive Assembly Simulation. Algorithms and Complexity Group.

Mancini, F., Lopez, M., Rinaldi, R., Boër, C.R., Pedrazzoli, P., Beccati, F., 2004. A new approach to programming flexibility for CNC machines. An example of application in woodworking machining. *Proceedings of the 37th CIRP Int. l Seminar on Manufacturing Systems (ISMS 2004)*, pp. 163-166, May 19-21, Budapest ,Hungary

Okada, K., Inaba, M., 2006. A Hybrid Approach to Practical Self Collision Detection System of Humanoid Robot. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3952-3957. October 9-15, Beijing, China

Pedrazzoli, P., Bathelt, J., Chryssolouris, G., Rovere, D., Pappas, M., Boër, C.R., Constantinescu, C., Dépincé, P., Westkämper, E. 2007. High Value Adding Virtual Reality Tools for Networked Customer-Driven Factory. *Proceedings of the 4th International Conference on Digital Enterprise*

Technology, pp. 347-352. September 19-21, Bath, United Kingdom. Maropoulos, Paul (Ed.), CIRP u.a

Pedrazzoli, P., Sacco, M.; Jönsson, A., Boër, C.R. 2006. Virtual Factory Framework: key enabler for future manufacturing. *Digital enterprise technology: perspectives and future challenges*, pp 83-90. Springer, 1

Shabana, A., 2001. Computational Dynamics. 2nd ed. Wiley-Interscience.

Steinbach, K. , Kuffner, J., Asfour, T., Dillmann, R., 2006. Efficient Collision and Self-Collision Detection for Humanoids Based on Sphere Trees Hierarchies. *Proceedings of the 6th IEEE-RAS International Conference on Humanoid Robots*; pp. 560-566. December 4-6, Genova, Italy

AUTHORS BIOGRAPHY

Dr. Ing. **Ilario Francesco Ceruti** is a software development consultant for Technology Transfer System s.r.l. He received his degree in Computer Engineering from Politecnico di Milano in 2005. Since then and he has done software development for engineering research, participating to EU founded research projects in FP6 on advanced 3D simulation of manufacturing environments (KoBaS). Since 2008 he has followed EU founded research project in FP7 concerning the development of tools to support design of mass customization shoes (Dorothy). His interests include computational geometry, computer graphics, databases and software engineering, and he is currently consultant for the development of a NC processes result 3D visualization.

Dr. Ing. **Giovanni Dal Maso** is a software development consultant. His interests include simulation and computer graphics. He has done software development for engineering research in EU founded projects. He received his degree in Electronic Engineering from Politecnico di Milano. He is now a researcher at SUPSI-ICIMSI.

Diego Rovere graduated in Mechanical Engineering (Robotics and Automation) at Polytechnic of Milan in 2003 with a thesis on an Automated System for the model and cost analysis in foundry. Researcher at the National Research Council (CNR) of Italy where he was responsible of the Java development group for the 3D simulation of manufacturing environments in the EU KoBaS project. Consultant for the development and integration of 3D simulation systems with CNC based on CANOpen protocol and with CAM systems for the woodworking sector .Consultant for the 3D simulation of WaterJet cutting processes in collaboration with Water Jet Sweden and Polytechnic of Zurich. Consultant of the Polytechnic of Milan for design and development of a software system for the stability analysis of milling operations performed on magnetic clamping planes. Consultant for the development of a control software for magnetic working surfaces with embedded force sensors. Responsible of the Java development and integration with a CAN based Virtual Cockpit Hardware of a simulation system of construction machines and agriculture vehicles. Since June 2009 researcher at SUPSI-ICIMSI.

FLEXIBLE PRODUCTION SIMULATION FOR APPLIED SCIENCES

Klaus Altendorfer[(a)], Josef Pilstl[(b)], Alexander Hübl[(c)], Herbert Jodlbauer[(d)]

Upper Austria University of Applied Sciences
Wehrgraben 1-3, A-4400 Steyr, Austria

[(a)]klaus.altendorfer@fh-steyr.at [(b)]josef.pilstl@fh-steyr-at

ABSTRACT

In this paper the structure of a flexible simulation model for production systems based on discrete event simulation is presented.

The basic idea of this flexible simulation model is to generate the shop floor structure of the studied production system based on work schedules and on the bill of materials. The production machines are not connected directly but the material flow is routed according to the information in the work schedule and the bill of material.

The model structure is based on different modules which use material flow entities and production order entities to communicate with each other and to simulate the production system.

As a result, the created simulation structure leads to a flexible simulation model generation which maintains a low effort for generating simulation models concerning different production system structures.

Keywords: Flexible Simulation, Production Simulation, Production Planning

1. INTRODUCTION

Manufacturing companies are often faced with challenging market situations concerning product complexity and changing demand. Their customers frequently focus on their logistic performance in additional to product quality. The logistical performance can, for example, be measured as the service level, delivery leadtime or average lateness. To influence this logistic performance, decisions have to be taken about capacity investment, production planning strategy and parameterization of the production planning method.

The influence of such decisions is discussed by analytical models in literature (e.g. Hopp/Spearman 1996 and Koh/Bulfin 2004). Since such models often lack the practical applicability, another method often applied in research papers as well as for applied research projects with companies is simulation. For example Jodlbauer/Huber (2008) compare different production planning strategies with simulation, Kutanoglu/Sabuncuoglu (1999) study different dispatching rules for production systems by applying simulation and Abdul-Kader/Gharbi (2002) provide a simulation model for capacity estimation of production systems. Furthermore, Huang et al. (1998) for example discuss different production planning strategies in an applied research project for a cold rolling plant by using a simulation model.

Based on these cited applications of simulation studies in the field of production research, it is clear that simulation can be used as a flexible tool to analyze production structures or production planning systems. Since such simulation studies are usually quite difficult to set up and are also quite specific concerning their application (see Thomson 1994), this paper presents a flexible production simulation approach which can be used both for industry and for research. This approach is based on the idea of object oriented simulation (OOS) which is argued to enable flexible discrete event simulation models by Joines/Roberts (1998), Borenstein (2000) and Anglani et al (2002).

2. LITERATURE REVIEW

Simulation is a widely used tool to model production systems as discussed in the *Introduction*. This short review of developments in simulation, especially in discrete event simulation, is focused on the process of generating a simulation model.

As discussed by Thomson (1994), creating a simulation model can be complex as well as resource and time intensive, but simulation models are often built for only one special case, the so-called throw away model. The need to analyze a new alternative requires again personnel and time resources. An alteration of the existing models is also usually very complicated and time intensive.

Simulation models have to become more flexible to reduce the development time. For this reason, Kronberger et al. (2008) developed a model generator that atomically generates a simulation model with data out of ERP (enterprise resource planning). They discuss several limitations of their automated model generator and their model depends on a certain data structure in the ERP system.

McNally/Heaevy (2004) successfully implemented simulation as a desktop resource in a medium sized facility as decision support for various manufacturing personnel. The simulation model is used both for strategic and short term planning and is linked to the

company's information systems. They report their main difficulties in model maintenance and data collection.

Son et al. (2003) developed a model generator for a real time shop floor control system which is based on simulation. They developed an automated simulation model generator which creates a simulation model for discrete part manufacturing based on a database and a certain structural information source about the shop level execution model.

Gravel et al. (1994) used simulation for a decision support system for production planning to support the production manager to in choosing the main parameters of a production planning system. The simulation model is again linked to the information system of the manufacturing company.

In the paper of Irizarry et al. (2001), a flexible simulation tool to discuss different manufacturing cell design patterns is introduced. The model is used to discuss different design patterns of a printed circuit board assembly cell.

Other models like the ones of Jodlbauer/Huber (2008), Kutanoglu/Sabuncuoglu (1999), Abdul-Kader/Gharbi (2002) and Huang et al. (1998) are all directly linked to one application or one structure discussed. Such models are, as already discussed in Thomson (1994), throw away models which often cannot be used for further applications.

Joines/Roberts (1998) show how flexible simulation can be realized with object oriented simulation. Different simulation modeling packages can be reused easily. Borenstein (2000) criticizes that major changes in simulation models cannot be performed easily and mentioned object oriented simulation as a method to increase the modifiability of simulation models. An application of OOS is delivered by Anglani et al. (2002) where it is used for flexible manufacturing systems simulation.

This literature review shows that some implementations of differently flexible simulation models already exist, but that there is still a need for further research in this field, especially if one is not discussing the link between simulation and the information system of manufacturing companies.

For this reason, this article describes a simulation approach applying flexible simulation modules that support applied research for fast and flexible production simulation.

3. FLEXIBLE SIMULATION MODEL

The simulation model described in this paper still relates to work in progress. The following section describes the planned features of the developed model which to date, are only partially completed.

It is based on the idea of creating different modules which can be used in the simulation software. These different modules are then combined to create a discrete event simulation model of different production systems. The software used for this model is AnyLogic 6.4.

In the first subsection, the targeted application of this flexible simulation model is described. The second subsection shows which data structure is needed to create a model. In the third subsection an overview of the functionality of the simulation model is given. The fourth subsection encompasses the description of the single modules used in the simulation model.

3.1. Application of the flexible model

This flexible simulation model is designed to be used for two purposes.

The first purpose is to enable applied research on the hierarchical planning process as conducted in many manufacturing firms. The hierarchical structure of the production planning in the simulation model is orientated at the MRPII (manufacturing resource planning) concept. The influence of different methods for the long term planning, the medium term planning and the short term control are addressed with that model. The methods used for these three levels of planning should be compared for different production system structures such as job shop or flow shop. Furthermore the influence of lotsizing policies, recursive production structures, scrap rates and processing time variation are discussed with the flexible model.

The second purpose of the model is to enable applied research projects with manufacturing companies through the ability to simulate their manufacturing sites with a considerable low effort.

Based on these two purposes the following concept for a flexible simulation model has been developed.

3.2. Data structure needed

Contrary to some of the literature cited above concerning flexible simulation systems, this simulation model has no direct link to an ERP system which a manufacturing company might have. According to the purposes the model is built for, it was decided to base it on a predefined data structure similar to the one available in ERP systems. The structure needed for the simulation model contains the following matrices:

Bill of material: Any kind of bill of material can be used for the model as long as it is available in a matrix structure indication how many pieces of different components are needed to produce one piece of a higher level component.

Work schedule: The work schedule matrix has to give the information concerning which machine or machine group is needed to transform a certain component into another component. Furthermore the processing time for each production step as well as the setup time for this combination component to machine or machine group is included into the work schedule. Additionally, the transportation lotsize between the single machines has to be defined in that matrix.

Production planning parameters: Up to now two different production planning methods are included in the flexible model, which are MRP (material requirements planning) and CONWIP(constant work in process). The planning parameters used in MRP are the lotsizing policy fixed order quantity, planned leadtime between production steps and safety stock per product

or component. For CONWIP the parameters are lotsizing policy, WIP-cap and work ahead window. Additionally some dispatching rules for production control are implemented.

Structural information: The structural information about the simulated production system is a matrix including the machine type, the number of machines per type and the machine availability (MTTR and MTBF).

Customer demand: To create the customer orders a matrix including the average order rate, the average customer order lotsize, the variance of customer order lotsize as well as the average customer required leadtime and the variance of customer required leadtime is needed for each finished good.

3.3. Simulation model generation and functionality

The model generation is based on a certain data structure discussed above. The model generation is, contrary to some of the cited literature above, not based on an automatic link between the simulation software and the ERP system which a manufacturing company might use.

To generate a model the single machines or machine groups have to be created in the AnyLogic 6.4 simulation software by using the module Machine. Furthermore, all the matrices from the data structure needed have to be filled with the corresponding data. To finish the model generation, the runtime of the model (and an appropriate warm up phase) and the experiments to be performed have to be defined.

For each simulation run a set of logistical performance indicators including service level, tardiness, production leadtime, overall utilization, WIP and FGI among others, are written to an MS Access database. So the result of each simulation experiment is a database file.

Figure 1 provides an overview of the simulation model functionality for a better understanding of the model.

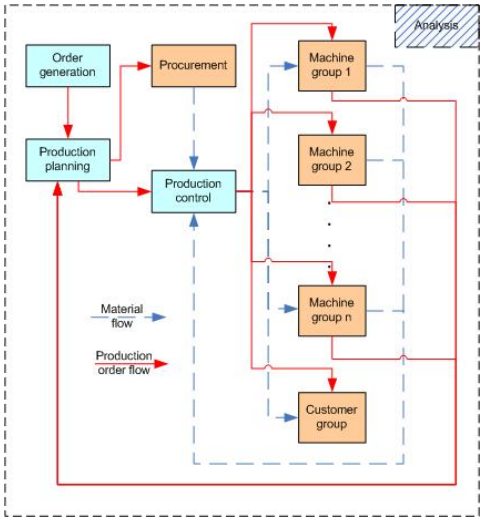


Figure 1: Functionality of the simulation model

As shown in Figure 1, the simulation model has two kinds of entity flows. The dashed line shows the material flow which is defined by the bill of material, the work schedule and the structural information. The bold line shows the production order flow which is defined by the bill of material, the work schedule and the production planning parameters.

The following steps are performed by a production order until the customer receives the products. The production order is generated via the parameters product type, amount and due date by the order generator and sent as an entity to the production planning module. Additionally, the production order is added to a list held in the analysis module. Based on the production planning strategy, production orders for the respective materials are created. These orders are sent to the production control module. The sorting of the production order according to the dispatching rule is performed in the machine group module. Based on the information in the production order, including the machine or machine group, the materials and the production order are routed to the corresponding machine or machine group. When a machine becomes idle, the next production order is produced. After the production order is finished, the materials and the production orders are routed to the production control module. In the production control module the availability of materials for the production orders is checked and the materials are sent to the corresponding machine group module whenever materials are available. All logistical figures needed are stored in the analysis module. Since no direct link between the single machines or machine groups is implemented in the model, but all flows are implemented with indirect links (AnyLogic implementation with Enter and Exit Blocks), the simulation model can quite easily be adapted to any production system structure.

3.4. Simulation modules

In this section the functionality of the seven modules as shown in Figure 1 is discussed in detail.

Order generation module: In this module the customer orders are generated with a quantity to deliver, the product type and the due date when to deliver.

Production planning module: The production planning module performs the complex task of creating and releasing production orders based on the customer orders. For the first setup of the model as discussed in this paper the two methods MRP and CONWIP are implemented in this module. For future research this module could be extended to other methods as Kanban, DBR (drum-buffer-rope) (see Schragenheim/Dettmer (2001)) or work load control (see Bechte (1998)).

Production control module: In this module the production orders and the materials are held and the material availability for the production order is checked. If the material needed for a production order is available, material and production order are released to the respective machine group. The module stores each of the finished goods, components and raw materials (one entity class called *material* in AnyLogic 6.4) used in the simulation model separately. Concerning the simulation model design, this storage is important for the flexibility of the production system structure which

can be modeled. Since each product is stored before and after a certain production step, in this storage the flexible flow of entities in the simulation according to the work schedule is assured. For simplicity reasons no transportation times between machines and the storage have been modeled yet.

Procurement module: In this module the materials are delivered for the first time to the production structure. This means the material entities are created. For simplicity reasons the procurement process is only modeled with a certain distribution of the delta between planned arrival date to real arrival date. The procurement is triggered by the production planning module with production orders for raw materials.

Machine group module: The machine group module can simulate planned maintenance, unplanned machine breakdown, machine set up time, scrap rate and assembly of components. A transport lot size can also be created after production in this module. Furthermore, the orders are sorted according their priority applying dispatching rules.

Analysis module: This module records statistics such as the utilization of machines, lead time, inventory, delivery reliability and other logistic key figures. All the values stored are independent of the production system structure. The only adjustment needed when the structure of the production system to model changes, is the addition or deletion of machine group dependent statistics.

Customer Group module: This module simulates the customer demand at the due date. The last generated production order out of a customer order is the customer demand at the due date and the materials are sent from the production control module to the customer group module.

3.5. Flexibility of the model

Based on these seven modules and on the structure as shown in Figure 1, the flexible simulation model can be applied to various production system structures.

As discussed in section 3.2, for each product a working plan and a bill of material have to be created for the simulation model. The working plan and the bill of material contain all the necessary information about the component which is needed for the machine group module and also for the routing through the production. Since the information about the routing is stored in the working plan and separated from the structure of the model, there is no more a need to construct fixed paths for the material flow in the simulation model, which makes the model more flexible concerning production system structure. Additionally, the structure of the products produced (e.g. sequential, convergent, divergent bill of materials or different processing and setup times) can easily be changed in the simulation model since again only the work schedule and Bill of material need to be changed.

4. OVERVIEW OF ANYLOGIC MODEL

The flexible simulation model described in this paper is still work in progress. The following section describes some already available and some planned features of the developed model.

4.1. Model architecture

Figure 2 gives an overview about the architecture of the developed simulation model in AnyLogic 6.4.

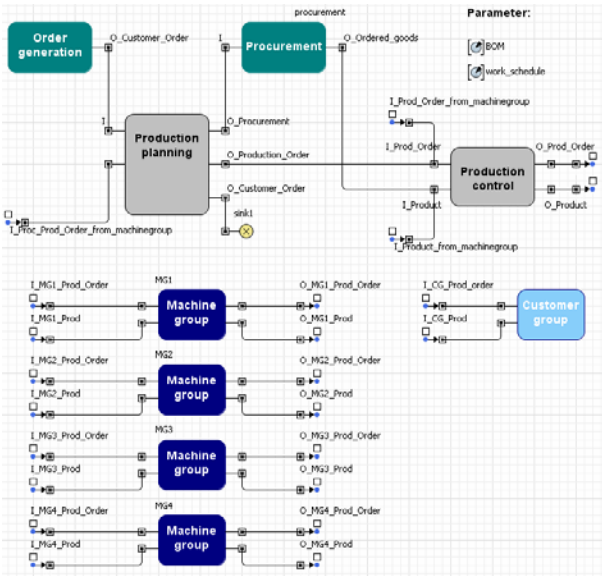


Figure 2: Simulation model in Anylogic 6.4

The Order generation module has for every good to finish a source which creates a customer order depending on a specific customer ordering rate as entity with the attributes due date, quantity to deliver and product type. All of the attributes have a specific variance depending on the predefined customer behavior. After creating the entity customer order the entity is sent to the production planning module.

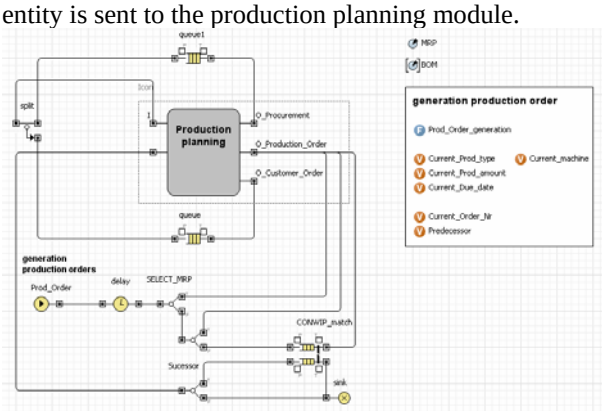


Figure 3: Production Planning module

In the Production planning module , as shown in Figure 3, the customer order is used for creating production orders. The function Prod_Order_generation controls the arriving entities customer order and checks the finished goods type. Each finished good has a parameter matrix working plan where the information about the routing of the components needed is also stored. The function gains the information about the routing for every finished goods type and creates a production order as an entity for every single machine group

module. The entity production order contains the attributes of the machine group to be sent, according to the finished goods type and the production planning method.

In the “delay” Block (from Figure 3) every generated production order entity is kept until its release to the module Production Control, depending on the production planning method.

As explained above, the input for the Procurement module is the entity production order (for raw materials). In this module a Source Block creates the needed raw materials as an entity. When the production order arrives, a function checks the material type and product amount and creates the necessary raw materials. To simulate a stochastic procurement leadtime, a stochastic delay time can be integrated in the Procurement module.

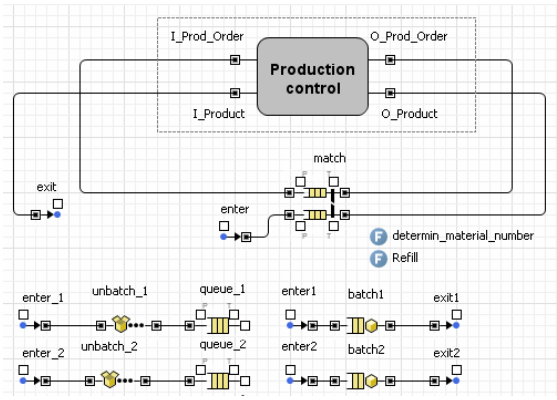


Figure 4: Production Control module

The Production Control module has two inputs: the production order entities for every single machine group from the Production Planning module and the material entities from the Procurement module or from the Machine Group modules. This module stores the material entities and checks the material availability for the released production order entities. The production order entities are stored in the upper queue of the “match” Block (see Figure 4) whereby the material availability check and the material release are triggered in the Function Block “Refill”. Each material entity has an own Queue Block within the Production Control module. Whenever a production order entity or a material entity arrives in this module, “Refill” checks the material availability. When material is available for a production order, the material entities get batched to the necessary production lot size and are sent to the second queue of “match”. In the “match” the materials and their respective production orders are released. After leaving the Production Control module, the material and production order entities are routed to the respective Machine Group module according to the work schedule.

In the Machine Group module the material entities (which are still batched to production lots) are sorted with a specific dispatching rule before they are released to the machines. Up to now, the EDD (earliest due date), FIFO (first in first out) and some other rules are implemented. Before the production of a lot starts, the batched material entity gets unbatched, then it is processed and after processing it is batched to the transportation lotsize again. Processing is interrupted by planned maintenance and unplanned machine failure. Scrap can also be simulated with a variable rate for each product. The material entities are sent back to the Production Control module immediately after a batching to transportation lots. The production order entities are sent to the Production Planning module after the last piece of a production lot is finished.

Additionally to the production orders for components, the Production Planning module creates one production order which triggers the release of the finished goods from the Production Control module to the Customer Group module.

The Analysis module consists of a set of variables and statistics which are continuously recorded.

4.2. Model application

The described simulation model can be applied to simulate a production with the production planning methods MRP and CONWIP. Since there are no necessary fixed routings between the machine groups, an easy and flexible application for different production structures is possible.

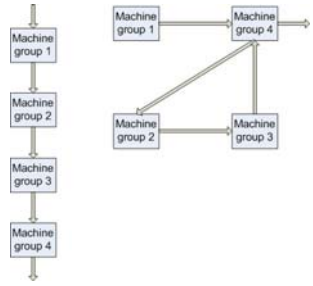


Figure 5: production structures

Both production structures shown in Figure 5 can be simulated without major changes in the model from Figure 2. Only the parameter matrix work schedule has to be adjusted.

5. LIMITATIONS OF THE MODEL

Based on the two proposed applications of the simulation model and on the implementation as described in sections 3 and 4, the following limitations occur.

The data included in the work schedule up to now does not include the variance of the processing time. The reason for this is, that the model is based on the assumption of exponential distributed processing times at this stage of development. In a further development, the model could be extended to other distributions.

No direct link between the simulation model and an ERP system of a manufacturing company is possible. Nevertheless, the work schedules and bill of material can be used in the model. Since no operational decision support is targeted with the model such a direct link is not necessary.

The simulation model is not generated automatically but there are still some manual tasks to

perform to create a simulation model concerning a certain production system. However, the structure of the simulation model is programmed in such a flexible way that different production systems can be simulated with minor model development time.

6. CONCLUSION

This paper presents a flexible design of a simulation model for production system simulation. The model is based on discrete event simulation and different hierarchical levels of production planning and control can be tested. Parameter sets for the different production planning methods can be compared and optimized with the simulation model as well. The flexibility of the simulation model concerns mainly the structure of the simulated production system. For the production planning two different methods are currently available and dispatching rules can be applied for the production control.

The concept presented is based on the data usually available in ERP systems like work schedules and bills of material. The simulation model is not generated by applying fixed material flow paths, but any path described in the work schedule is possible.

Concerning the two kinds of application this model is developed for, the application in applied research on the behavior of different hierarchical production planning methods can be fulfilled. The flexible structure seems to enable the applicability of the model to applied research projects with manufacturing companies, but such projects have not yet been conducted.

This method of creating a flexible simulation model as presented in this paper could in further research be extended by additional modules and can for this reason be the basis for many different simulation studies concerning production logistics.

REFERENCES

Abdul-Kader W., Gharbi A., Capacity estimation of multi-product unreliable production line, *Int. J. Prod. Res*, Vol. 40, 2002, p. 4815-4834

Anglani A., Grieco A., Tolio T., Object oriented modeling and simulation of flexible manufacturing systems: a rule based procedure, *Sim. Modelling Practice and Theory*, Vol. 10, 2002, p. 209-234

Bechte W., Theory and practice of load-oriented manufacturing control, *Int. J. Prod. Res.*, 1988, Vol. 26, p. 375-395

Borenstein D., Implementation of an object-oriented tool for the simulation of manufacturing systems and its application to study the effects of flexibility, *Int. J. Prod. Res.*, Vol. 38, 2000, p. 2125-2142

Gravel M., Kiss L., Martel J.M., Price W.L., A DSS for Production Planning, *Int. Trans. Pol. Res.*, Vol. 1, 1994, p. 363-373

Jodlbauer H., Produktionsoptimierung, 2. Auflage, 2008 SpringerWienNewYork

Jodlbauer H., Huber A., Service-level performance of MRP, kanban, CONWIP and DBR due to parameter stability and environmental robustness, *Int. J. Prod. Res.*, Vol. 46, 2008, p. 2179-2195

Joines J.A., Roberts S.D., Object-Oriented Simulation, in: *Handbook of Simulation (ed. Jerry Banks)*, 1998, New York

Huang M., Wang D., Ip W. H., A simulation and comparative study of the CONWIP, Kanban and MRP production control systems in a cold rolling plant, *Prod. Plan. & Control*, Vol. 9, 1998, p. 803-812

Kronberger G., Kerschbaumer B., Weidenhiller A., Jodlbauer H., Automated simulation model generation for Scheduler-benchmarking in manufacturing, *Proceedings I3M*, 2006, p. 45-50

Kutanoglu E., Sabuncuoglu I., An analysis of heuristics in a dynamic job shop with weighted tardiness objectives, *Int. J. Prod. Res.*, Vol. 37, 1999, p. 165-187

McNally P., Heavey C., Developing simulation as a desktop resource, *Int. J. Comp. int. manuf.*, Vol. 17, 2004, p. 435-450.

Son Y.J., Wysk R.A., Jones A.T., Simulation Based shop floor control: formal model, model generation and control interface, *IIE Trans.*, Vol. 35, 2003, p. 29-48.

Schragenheim E., Dettmer W.H., Manufaturing at Warp Speed, 2001, The St. Lucie Press/APICS Series on Constraints Management, Boca Raton

Thomson M.B., Expanding simulation beyond planning and design, *Ind. Eng.*, 26, 1994, p. 65-67

AUTHORS BIOGRAPHY

DI (FH) Klaus Altendorfer is working as an Assistant Professor in the field of Operations Management at the University of Applied Sciences, Steyr (Austria). Currently he conducts his PhD studies in the field of production logistics at the University of Vienna.

Josef Pilstl, BSc is working as a Research Associate in the field of production optimization at the University of Applied Sciences, Steyr (Austria). He is also currently studying the graduate school program “Operations Management”.

DI (FH) Alexander Hübl, MSc is working as a Research Associate in the field of Operations Management at the University of Applied Sciences, Steyr (Austria). Currently he conducts his PhD studies in the field of production logistics at the University of Vienna.

FH-Prof. Dr. Dipl.-Ing. Herbert Jodlbauer is director of studies of the degree course “Production and Management” and “Operations Management” at the University of Applied Sciences, Steyr (Austria). He received his PhD degree in engineering with distinction and has professional experience in process optimization and production planning and control.

INTRODUCTION OF SIMULATION METHOD AND POSSIBILITIES OF STANDARDISATION

József Perger[(a)], János Jósvai[(b)], András Pfeiffer[(c)], Botond Kádár[(c)]

[(a)] Audi Hungaria Motor Kft.
[(b)] Széchenyi István University, Professorship for Material Sciences and Automobile Production
[(c)] Computer and Automation Research Institute, Hungarian Academy of Sciences

[(a)] jozsef.perger@audi.hu, [(b)] josvai@sze.hu, [(c)] pfeiffer@sztaki.hu, [(c)] kadar@sztaki.hu

ABSTRACT
The paper focuses on the applicability of simulation technology in several hierarchical levels of a production oriented industrial firm. The topic of the paper addresses the discrete event simulation technology which is used to model the material flow and the manufacturing processes in the company. This paper would like to show some practice experience about the industrial introduction steps of manufacturing simulation, as well as our recent research result in the field of standardized simulation model building and data handling procedure.

Keywords: discrete event simulation, industrial management, standardization, PLM.

1. INTRODUCTION

In today's economy, product life cycle has been shortened. There is a new challenge for manufacturers: high quality products are to be produced, the introduction time is very short and the costs must be as low as possible. PLM is an engineering solution to address this challenge. It allows for users to analyze and manage products through the whole product lifecycle. One efficient tool and method is the modelling and simulation the manufacturing processes of the products.

Because of its great versatility, flexibility, and power, simulation is one of the most widely used evaluations and decision-support techniques (Law and Kelton 2001). While simulation, in theory, has great potential to assist in the understanding and efficient operation of manufacturing systems, several studies shows that there is a low usage of discrete event simulation by industry (Banks 1998).

The introduction of simulation into the manufacturing field of industry has its first steps. The different manufacturing areas have as well different data needs, and data gathering possibilities. The very first step of an industrial introduction is the harmonization of data handling, identification and gathering processes. Then the model elements could be defined for reuse, the elements have the ability to be configured for special behaviours, this is valid for equipment as well. The standardization of simulation modelling in the industrial field continuous with working and control methods, these logical systems are difficult to describe in standards, especially if there are many changes in the real logics and resource using methods.

The paper focuses on the possibilities of simulation standardization in the automotive industry, which is a recent issue in today's customer oriented production, as well as a methodology is presented regarding the introduction of simulation methods in the “every day use” in the same environment. The topic of the paper addresses the discrete event simulation technology which is used to model the material flow and the manufacturing processes in the company. The CAD/CAM related simulation e.g. finite element analysis and robotic simulation is out of the scope of this paper.

2. STATE OF THE ART

An extensive study of the penetration and use of discrete event simulation in the UK manufacturing industry identified only 11% of sites out of sample of 431 which were currently utilizing simulation as a decision support tool. This view of the penetration of simulation into industry is also supported by more recent surveys (Eriksson 1999, Hirschberg and Heitmann 1997, McLean et al 2003, McLean and Shao 2003). The literature on manufacturing systems simulation reinforces our conviction that simulation is a technique that still has a lot of underexploited potentialities.

When conducting a simulation study it is recommended that a structured systematic approach be carefully planned and rigidly adhere to. The 40-20-40 rule is a widely quoted rule in simulation related papers. The rule states that, in developing a model, an analyst's time should be divided as follows:

1. 40% to requirements gathering such as problem definition, project planning, system definition, conceptual model formulation, preliminary experiment design and input data preparation;
2. 20% to model translation
3. 40% to experimentation such as model validation and verification, final experimental

design, experimentation, analysis, interpretation, implementation and documentation.

The previous principle is confirmed in (Tecnomatix 2006), where the authors point out that collecting and preparing the data in order to use in the simulation study is one of the most important tasks, as it takes up about 35% of the project time. Creating the model takes up another huge amount of time (25%), while validating and correcting needs 15%, running the experiments 10%, finally analyzing and evaluating 15% of the project time (Pfeiffer 2007).

The key requirements of simulation influence the needs and expenditures of the realization process of a simulation. We specify the requirements, based on the challenges formulated above, and highlight the main directions to be followed in order to be able to fulfil the requirements. Thus, key requirements can be listed as follows.

- Data acquisition, preparation and modelling capability are key elements, while during the other phases, regarding a production simulation study the reduction of the expenditures is fairly not as promising as by the others.
- Consequently, improving model building techniques, applying reusable model elements, through modular software architecture and object oriented modelling.
- Integration to ERP, MES systems might results in a reasonable data acquisition platform.
- Reuse model components for different purposes in different life-cycle phases of the system modelled.

Regarding the national perspective, we think that there must be a solid base for industrial application in Hungary, regarding the numerous multinational, high-tech manufacturing enterprises. These companies often "import" their knowledge due to applying solutions and processes which are – so called – company-wide standards. Despite to this advantageous situation, conducting simulation studies, moreover, continuous use of simulation is nowadays not a key issue in Hungarian companies, and thus, the important simulation-related knowledge has not been transferred and transmitted until now in the industrial practice (Pfeiffer 2007).

A manufacturing system has usually a large investment, it consist mostly of capital equipment and software to operate them. The integration of these separate systems is time-consuming, and means high costs. To establish an efficient manufacturing system design the dynamic model of the organizations is useful. It makes possible to perform "what-if" analysis, but it needs experts in construction and also in analysis of results. The most of the industrial simulation models are addressed to a defined set of industrial issues. This means, the models are used for special tasks, reuse of models is difficult or only with large modifications possible. In our opinion these drawbacks could be reduced or even eliminated with the proper usage of simulation standards.

In the following space we will give a short overview about the standards and standardization activities which influence the simulation models and their implementation processes in an industrial firm.

2.1. Simulation related standards

2.1.1. NIST

National Institute of Standards and Technology (NIST) was founded in 1901 and is a non-regulatory federal agency within the U.S. Commerce Department's Technology Administration. NIST's mission is to develop and promote measurement, standards, and technology to enhance productivity, facilitate trade, and improve the quality of life. From automated teller machines and atomic clocks to mammograms and semiconductors, innumerable products and services rely in some way on technology, measurement, and standards provided by the National Institute of Standards and Technology.

Within NIST, the Manufacturing Simulation and Visualization (MS&V) efforts are focused on accelerating the development of simulation standards. The Manufacturing Systems Integration Program has a subproject called Simulation- based Manufacturing Interoperability Standards and Testing. This subproject contains several main fields, these are Frameworks and Architectures, Data Models and Standards (see below) CMSDIM, Simulation Prototypes and Testing Systems.

2.1.2. Core Manufacturing Simulation Data Information Model (CMSDIM)

The CMSD Information Model defines a data specification for efficient exchange of manufacturing data in a manufacturing simulation environment. The specification provides a neutral data format for integrating manufacturing application and simulation. The purpose of the CMSD Information Model is to:

- Enable data exchange between manufacturing simulation systems, other software applications, and databases.
- Support the construction of manufacturing simulators.
- Support testing and evaluation of manufacturing software.
- Support manufacturing software application interoperability.

This product defines a data interface specification for efficient exchange of manufacturing life cycle data in a simulation environment. The specification provides neutral data interfaces for integrating manufacturing software applications with simulation systems. The initial effort is focusing on machine shop data

definitions. The plan of the authors is to extend the data specification to include supply chain, aerospace assembly operations, automotive vehicle assembly operations, plant layout, and other relevant manufacturing and simulation information.

A NIST analysis shows, that several standards exist, but comparing these standards it can be stated that there are huge differences even in naming of conception. The analysis shows how much information discrete-event simulation requires, as well. Functionality of units must be described, such as: logical elements, product elements, process plan. Application objects are for example: arrival, breakdown, path, processor, schedule, etc. These attributes show how complex this area is, and the standardization of simulation tasks has lot of different influence types (NIST 2004.)

2.1.3. VDA Standard

This statement serves as an execution application and technology cross policy for internal and external simulation projects. It serves as the basis for the acceptance and performance including:

- the definition of general, organizational and computer technical guidelines,
- the standard procedure for simulation projects, and guidelines for implementation,
- the requirements for quality and its management,
- the global, non-project specific input data and requirements,
- the requirement for the data and its management and validation,
- the requirements for simulation models and their validation,
- the principles of modelling and programming,
- the requirements for experimental design and analysis. (VDA 2008)

Well-defined project structure and recommendations are given in this standard with specific points regarding areas and rules how to manage the development of a simulation model (Figure 1). On the base of the VDA the full simulation project can be controlled in structural way and this standard offers guidelines for this procedure. VDA also offers recommendations for model structure, verification and validation, documentation and evaluation. Furthermore it contains a description part for Plant Simulation, where software specific recommendations are assembled, such as use of class library, data handling, naming conventions, meanings of different colours, program head. There are some specific areas in the automation industry (e.g. chassis, surface handling, assembly, logistics and factory simulation), which are handled by VDA separately, giving exceptions for the modelling of these fields.

2.1.4. ANSI/ISA 95

ISA-95 is the international standard for the integration of enterprise and control systems and it consists of models and terminology. The information in the standard is structured in UML models, which are the basis for the development of standard interfaces between ERP and MES systems (Bradl 2008). The ISA-95 standard can be used to determine which information, has to be exchanged between business logistics systems and manufacturing operations systems. Another objective of the ISA 95 is to provide standardized models of activities in manufacturing operating systems.

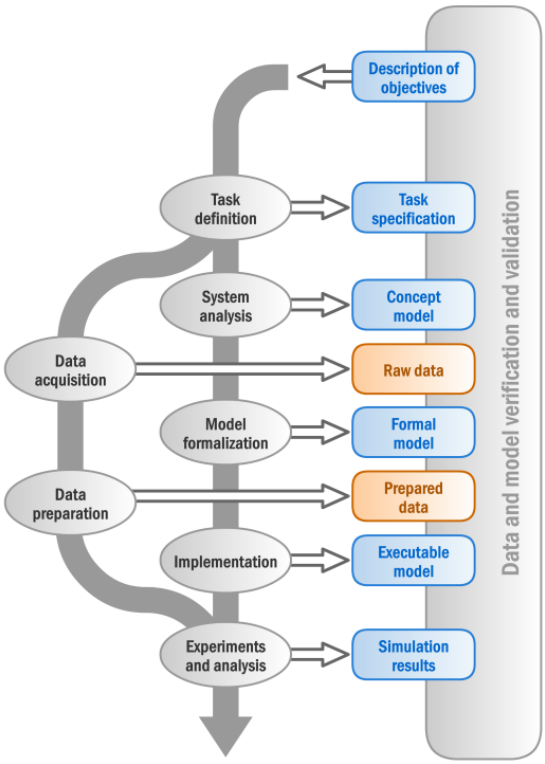


Figure 1: Process Model ASIM (2007)

There are no direct references for simulation in the ANSI/ISA 95 standard, however, the operational manufacturing activities and the functional components of the Manufacturing Execution Systems (MES) are given in detail and as a model of these material and informational systems, and the simulation models can follow the ANSI/ISA 95 models.

3. ENTERPRISE SIMULATION CUBES

In this section we introduce a new classification model called *Enterprise Simulation Cubes* in which, from the targeted simulation model point of view, we consider the enterprise decision hierarchy, stage of the modelled system in its own life-cycle and the functional divisions of the system in the enterprise. The dimensions above form a framework and according to these classification groups several different simulation cubes can be identified, each of which represents a specific sub-cube in the overall framework (Figure 2).

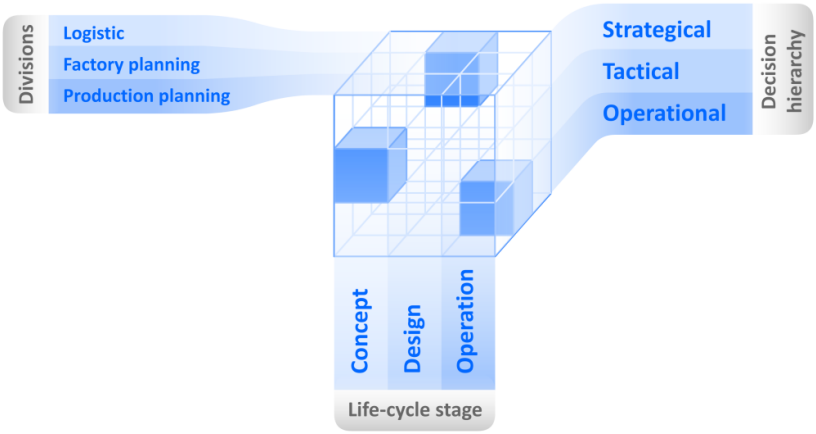


Figure 2: The concept of Enterprise Simulation Cubes

From the *decision hierarchy* point of view we consider the well-known

- Strategic,
- Tactical and
- Operational levels.

The Life-cycle projection includes the

- Conceptual,
- Design and
- Operational stages.

Last but not least we consider the functional divisions in the enterprise like:

- Logistics,
- Factory Planning,
- Production Planning, etc.

If a sub-cube is taken from above framework, which takes one specific value from each of the above categorizations, specific requirements, questions, the level of model's detail, the simulation time horizon, etc. can be considered.

The requirements in a sub-cube define the depth or resolution of the model. At one extreme, an entire production system can be modelled as a single "black box" operation with a random activity time (ProModel 2003). This solution is more relevant in conceptual life-cycle stages and/or during strategic decisions. At the other extreme, every detailed motion of a machine could be modelled with a one-to-one correspondence depicting the entire machine operation. This is more relevant in lower hierarchical levels e.g. design or operation life-cycle of the system in operational decisions.

Determining the appropriate level of detail is an important decision. Considerable high detail makes it difficult and time consuming to develop a valid model. Excessively low-level of detail makes the model unrealistic by excluding critical variables. Figure 3: illustrates how the time to develop a model is affected by the level of detail. The importance of including proper detail to meet the objectives of the study is also highlighted.

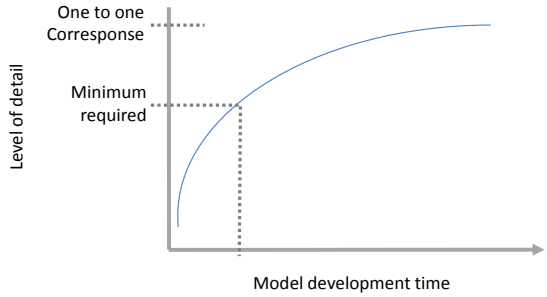


Figure 3: Effect of level of modelling detail on model development time (from ProModel 2003)

As for different life-cycle stages (orange arrow in Figure 4) the granularity of the simulation models differ. In the *conception phase* the simulation might be used for marketing a project to the management. The modeller should realize the simulation meta-model following the principle that the first-phase models usually do not require model components which are too detailed, i.e., the system itself to be modelled is very complex ((Pfeiffer 2007.), see Figure 4). At this work-phase simulation is not connected usually to the company information systems. Another constraint in this phase is to provide data mainly regarding investment cost and capacity, moreover, these results must be interpreted to managerial personnel, which usually mean high level graphical representation of the system modelled.

In the *design phase* simulation is used to find the best solution from a set of potential designs. The focus in this phase is the overall operating strategy. From the modelling point of view, the model-structure created in the preceding phase, is expanded with the *static* data

gathered from the DE, i.e., an interface to the company database has to be realised.

During the *operational phase* – where usually subsystem of the production system are built, delivered, and installed – the simulation is connected to the real control software to test the software implementation. The controllers use the emulation (refined simulation) model as a replacement for the physical equipment. In this way the control logic can be tested for the entire facility. If changes of the system are required the simulation model can be applied for improving the installed system or testing suggested modifications before implementing the changes.

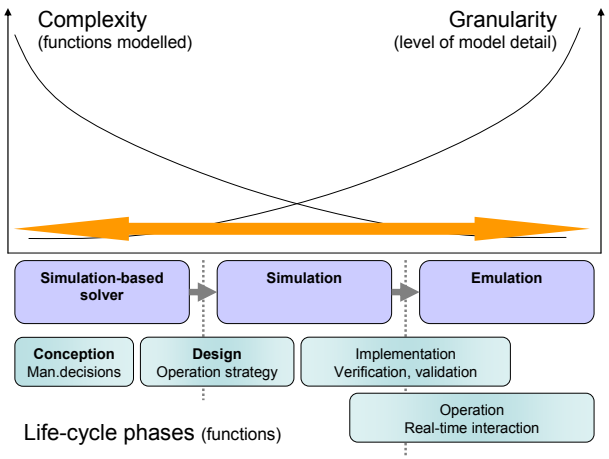


Figure 4: Proposed extension of simulation to different life-cycle phases of a production system

The complexity of model building should never be underestimated and it is always better to begin simple and add complexity rather than create an entire complex model at once (see different modelling details at the different phases in Figure 5). Building a model in phases (or stages) enables failures to be more readily identified and corrected as well. It is also easier to add detail to a model than it is to remove from it, furthermore, a model with excessive detail may be too expensive to program and to execute.

Our hypothesis is that if the level of modelling detail increases, the features and functions modelled must be reduced, also required by the limited computational efforts available.

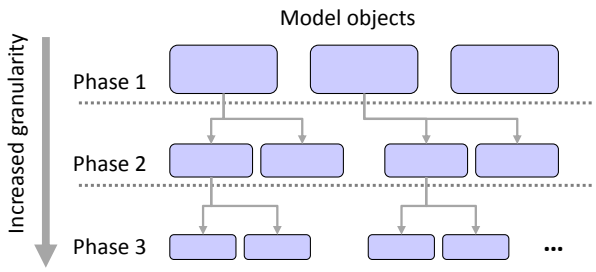


Figure 5: Granularity of the model objects at the different phases of the simulation models

The introduced new approach reflects a new conceptual view in simulation modelling of productions systems and may support better integrity to manufacturing ICT systems. The necessity and actuality of applying this new technique is proven through a literature review, furthermore, the proof of the concept is reinforced by two case-studies in the coming space.

4. CASE STUDIES

In this section we briefly introduce two case studies referring to separate sub-cubes of the framework presented in the previous section.

4.1. Tactical decision making at the design phase – case-study

The first case-study was carried out with the factory planning division of an industrial enterprise and mid-term tactical decision making was supported by a simulation model in the design of a new factory section. Before the real implementation of a budgeted project, with the static design and paper-based calculation in hand, a simulation model was created to test and validate the dynamic behaviour of the envisaged system.

Adhering to the Enterprise Simulation Cube, the target of the simulation model was to answer the question whether the static design is capable to offer the calculated throughput, but the model should integrate the supporting function of other divisions which was not possible to be included in the static design. During the simulation analysis, the main focus was given to a new unload station ("Unload 3" in Figure 6) at the end point of the main roller track system.

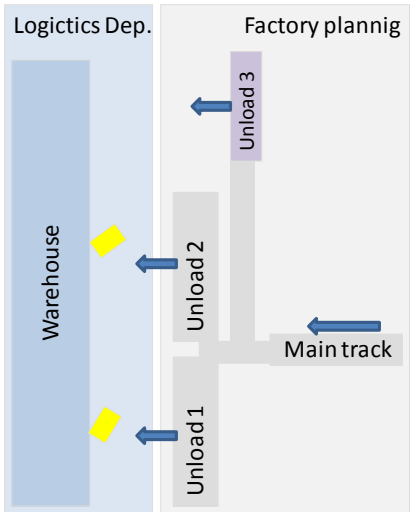


Figure 6: The layout and the main material flow of the selected section to be modelled

With the help of the standard, company wide modelling object libraries, the simulation model has been developed in Plant Simulation v8.1 and comprises the object classes as follows:

- MU: Product, pallet, elevator, logistics pallet, and forklift;
- RB: roller track component (consist of the track and the low level control system);
- Operator: operator of the system at the unload stations, driver of the forklift, etc.;
- Warehouse: store in and out the products;
- Control: control of the input and output data, the operators activity as well as the elevator;
- DataIF: data interface for easy data exchange;
- Statistics library and GUI.

Two main areas had to be analyzed in details with the resulted simulation models:

1. Define the sorting and mixing logic of the main elevator, i.e., in which order should the pallets be loaded onto the elevator.
2. Define the minimum cycle time for the forklifts necessary for the undisturbed operation of the unload station (the forklifts served as interfaces between the unload stations and the main warehouse).

In Figure 7 the results of different simulation scenarios are presented where the effect of the incoming product mix at the elevator is analyzed. In this dynamic analysis the simulation runs statistically and demonstrated the right-left-right logic RLR of the product mix dominate the other scenarios. The final control system, which is working today, applies this mixing rule.

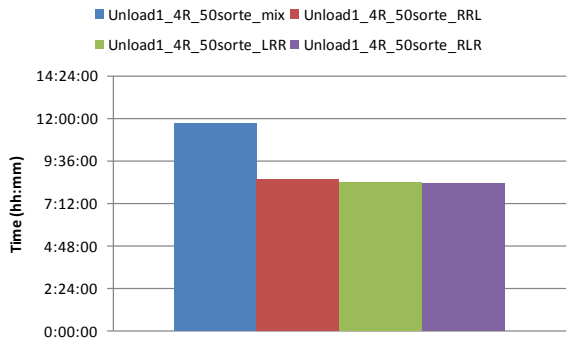


Figure 7: Time total required for unloading 1800 pallets by applying different mixing logics in front of the main elevator at the unload station 3

Figure 8 shows the results of the analysis that was carried out in order to identify the effect of the forklift's service time on the throughput of the system. According to the simulation runs we can state the forklift's service time effect comes out if and only if this service time is higher the 1 minute. From this point if the service time is higher the throughput of the system decreases linearly.

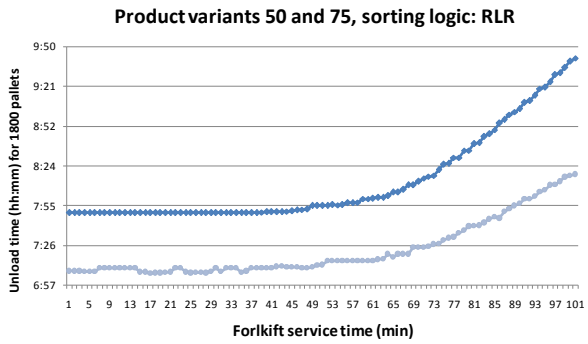


Figure 8: The total time required for unloading 1800 pallets at unload station 3, as the function of the forklift service cycle time

4.2. Operational decision making at the operation phase – case-study

The second case-study was carried out with the manufacturing planning division of an industrial enterprise. The simulation model developed in this case-study aimed at supporting the short-term production scheduling decisions of a manufacturing line.

The main target of this simulation model (Figure 9) was to answer the questions, how the production schedule and the line balancing affect the behaviour of the production line. As such, the main goals of the study were as follows:

- Determination of optimal production plan
- Balancing the utilization of workstations
- Stock reduction within the assembly cell

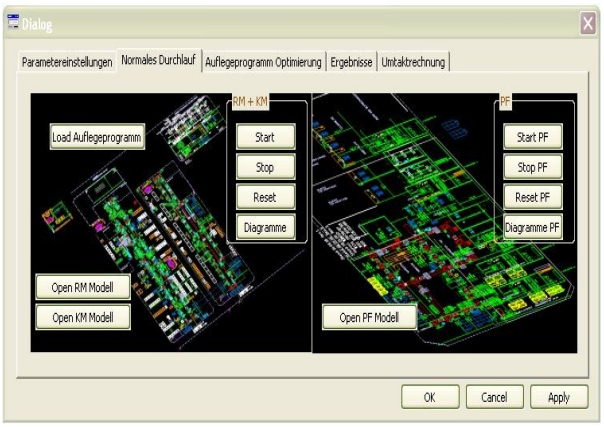


Figure 9: User interface of the production scheduling case-study model

The standard graphical analysis tools in the simulation model support the decision making, moreover, these diagrams are changing dynamically during the on-line simulation time strengthening both the verification and the real-time decision process (Figure 10).

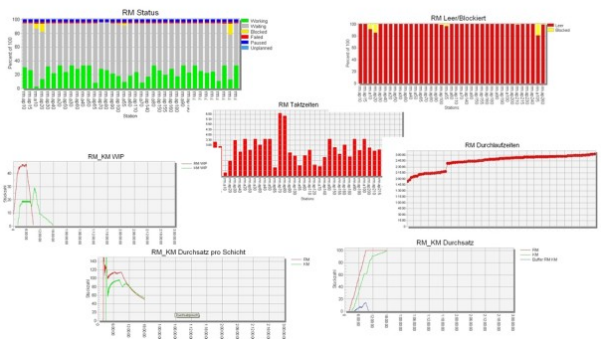


Figure 10: Graphical evaluation diagrams for operational decision making

The simulation model of the second case-study was further complemented with a Genetic Algorithm (GA) based optimizer. The same simulation model which was implemented for the analysis of the production line is also applied as the fitness function of the GA optimizer. The results of the genetic-based solution can be exported and used in outer application for further operations (see GUI of the GA-based optimizer in Figure 11).

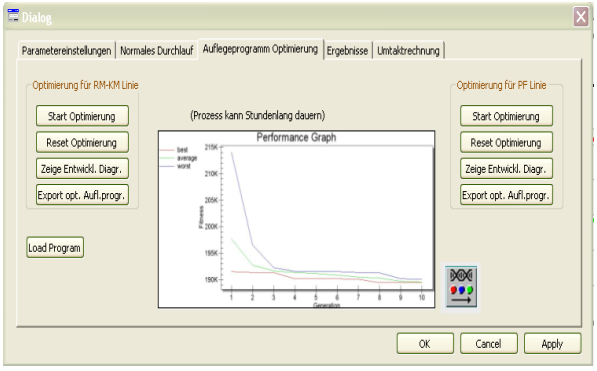


Figure 11: User interface of the GA-based optimizer

This operational production scheduling tool supports the decision making by applying real production data. In order to identify the aims of the model and the data gathering needs, as well as the definition of the level of detail, it was obvious to apply the Enterprise Simulation Cube. By this way the different areas of the life-cycle stages, the different decision hierarchy and divisions of the plant involved could be easily identified.

5. CONCLUSIONS

The paper focuses on the applicability of simulation technology in several hierarchical levels of a production oriented firm, furthermore on the possibilities of simulation standardization in the automotive industry, which is a recent issue in today's customer oriented production, as well as a methodology is presented regarding the introduction of simulation methods in the "every day use" in the same environment.

The paper discusses several process models designed for simulation steps, these models have the handling area from data gathering to experiment analysis. A new concept – Enterprise Simulation Cube - has been established to identify and classify simulation model views, aspects for detail level and model complexity.

Two case-studies were presented highlighting how to use the Enterprise Simulation Cube, and the practical advantage of it. One case-study concentrated on a tactical-level decision support for a factory design project while the second one demonstrated how the simulation tool can be used on the operational level in the daily shop-floor production planning process.

REFERENCES

Eriksson, U., 1999, Production Engineering, Chalmers University of Technology, Göteborg, Sweden. Working paper.

Banks, J., 1998, Handbook of simulation, Principles, Methodology, Advances, Application and Practice. JohnWiley & Sons Inc., ISBN 0-471-13403-1

Hirschberg, A.G., Heitmann, K., 1997, Simulation in German industry – a survey, in W. Hahn & A. Lehmann, eds, Simulation in Industry, -9th European Simulation Symposium, 1997, Passau, Germany, Society for Computer Simulation International, pp.: 429–433.

Law, A., Kelton, D.: Simulation modeling and analysis, McGraw-Hill, 2000.

Manufacturing Systems Engineering Group, Manufacturing Engineering Laboratory, Manufacturing Systems Integration Division, 2004. Production Systems Engineering: Requirements Analysis for Discrete-Event Simulation. Gaithersburg: National Institute of Standards and Technology.

Pfeiffer, A., 2007. Novel Methods for Decision Support in Production Planning and Control, Thesis (PhD). Budapest University of Technology and Economics.

Tecnomatix Technologies Ltd., 2006. Tecnomatix Plant Simulation Help.

McLean, C., Leong, S., Harrell, C., Zimmerman, P. M., Lu, R. F., 2003, Simulation standards: current status, needs, future directions, panel: simulation standards: current status, needs, and future directions, Proceedings of the 35th conference on Winter simulation: driving innovation, December 07-10, 2003, New Orleans, Louisiana, pp: 2019-2026.

ISA. http://www.isa.org/

ISA-95. http://www.isa-95.com/

SYSOSTDS http://www.sisostds.org/index.php?tg=articles&idx=More&topics=20&article=42

CMSD Product Development Group, Simulation Interoperability Standards Organization, Core Manufacturing Simulation Data Information Model (CMSDIM), PART 1: UML Model http://www.sisostds.org/index.php?tg=fileman&id

x=get&id=49&gr=Y&path=Specifications&file=CMSD2006-9-13.pdf
Bradl, D. What is ISA-95? Industrial Best Practices of Manufacturing Information 2008-05-19, http://www.sesam-world.dk/ _pdf/ m80/ What%20is%20ISA%2095.pdf
VDA, Ausführungsanweisung Ablaufsimulation in der Automobil- und Automobilzulieferindustrie, V. 2.0, 2008.11.06.
McLean, C., Riddick, F.: PC Modeling and Simulation Guidelines, Volume 1 – overwiev, NIST, 2004.
NIST. http://www.nist.gov/
MEL-NIST. http://www.mel.nist.gov/proj/msv.htm
McLean, C., Shao, G., 2003, Generic case studies for manufacturing simulation applications, in Proceedings of the 2003 Winter Simulation Conference, 2003.
ProModel User's Guide, 2003.

AUTHORS BIOGRAPHY

József Perger has been working for the Audi Hungaria Motor Kft since 1995. He has degrees from mechanical engineering, logistics and economy. Today he is working in the field of production planning. He is a PhD candidate now and he is doing research in the system of the local and regional conditions in order to improve the company. Also he would like to analyze the expenditures of the manufacturing process, using a digital manufacturing planning program.

János Jósvai is working at the Széchenyi István University, Győr, Hungary. He is a PhD candidate, his field is manufacturing planning and simulation methods. He has several year experiences in material flow simulation of manufacturing systems and in production process planning. In the field of research and development he spent significant time aboard with international cooperation in matter of digital factory.

András Pfeiffer, earned his PhD in 2008 at the Budapest University of Technology and Economics. Currently he is a senior research fellow at the Engineering and Management Intelligence Laboratory of 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.

Botond Kádár is a senior researcher at the Engineering and Management Intelligence Laboratory of the Computer and Automation Institute of the Hungarian Academy of Sciences (SZTAKI). He obtained his MSc and Ph.D. degrees at the Budapest University of Technology and Economics, Hungary, in 1993 and 2002, respectively. His current interest includes production control, simulation and multiagent approaches for production engineering and manufacturing systems and he is involved in several research and development projects from these fields. Dr. Botond Kádár is author or co-author of 70 publications with over 120 citations.

MODELLING AND SIMULATION OF A NATURAL ROOFING SLATES MANUFACTURING PLANT

David del Rio Vilas[(a)], Diego Crespo Pereira[(b)], Juan Luis Crespo Mariño[(c)], Alejandro Garcia del Valle[(d)]

[(a) (b) (c) (d)] Grupo Integrado de Ingenieria (GII), University of A Coruna, Spain

[(a)]daviddelrio@udc.es, [(b)]dcrespo@udc.es, [(c)]jcm@udc.es, [(d)]agv@udc.es

ABSTRACT

Producing natural roofing slates is a highly wasteful activity. Depending on commercial formats, only a 3% of material extracted from quarry becomes final product. Processing a block of slate weighing several tons into a tile of slate only a few millimetres thick requires great energy consumption as well as the use of an important level of resources, both human and material. The whole process is subject to the intrinsic variability of natural products which determines its reactive nature. In this paper, we present our work for the global improvement of a natural roofing slates manufacturing plant. To do so, a Modelling and Simulation approach has been adopted. Developing a Discrete Event Simulation (DES) model of such a high dynamic, variable and labour-intensive process has been proposed as a systematic way for its characterization and analysis.

Keywords: process simulation, flexible manufacturing systems, natural roofing slates.

1. INTRODUCTION

Europizarras S.L. is a Spanish company that produces natural roofing slate for institutional and residential buildings. More than 80% of its production is exported to other countries in Europe, especially France, where their slates have been awarded with the NF mark which sets the highest quality requirements in the industry.

The company is mainly devoted to the production of the highest value added roofing slates, that is to say, the thinnest commercial tiles. The thinner the tile is the harder and more wasteful the manufacturing process becomes. On the other hand, there is a quite constant demand of 3.5 mm thick tiles from France which provides a stable market.

In spite of the Spanish slates are the most employed in the world, the sector has scarcely benefited from technological transference from other industries. The level of automation is low as well as the application of lean manufacturing principles. The most arguably reason is perhaps the relative geographic isolation of slate production areas, mainly located in the northwest mountain region of Spain. Besides or as a result, it is labour-intensive and workers are exposed to very hard conditions both environmental and ergonomic. It is indeed difficult to find skilled workers or even convince youngsters to start this career so high salaries have to be offered. Accordingly, labour and operating expenses account for one third each of the total company costs set up.

In this context, the company has started a global improvement project comprising actions in the fields of production, quality, health and safety and environment. The purpose is to achieve a more efficient process in terms of productivity and the first step is to gain knowledge about the operations involved aiming at reducing uncertainty, defining capacities, and identifying both opportunities and limiting factors for a subsequent process optimization. These first steps are presented in this work.

2. THE PROCESS

For the extraction of slate from quarry light explosives are employed. The results are irregular and heavy blocks that are then loaded onto dumpers and transported to the manufacturing plant, located a few kilometres away. These blocks are then introduced in the Sawing Plant and stocked, so an adequate level of input is always assured. In this plant blocks are first cut into strips by means of circular saws and then a second group of saws cuts the strips into slabs which are then carried to the splitters on an automated conveyor belt.

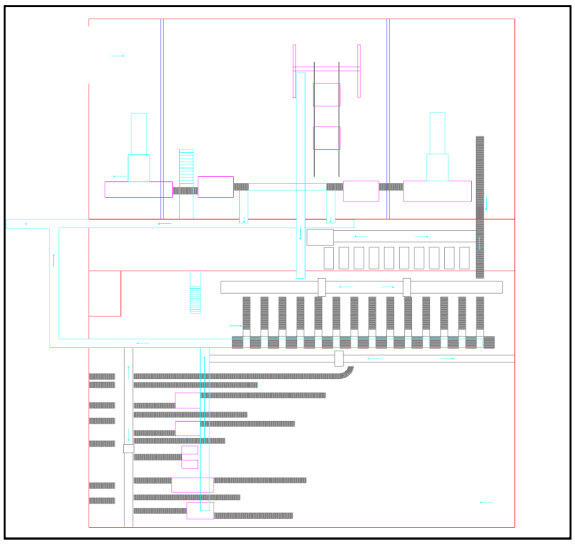


Figure 1: The CAD layout of the Manufacturing Plant

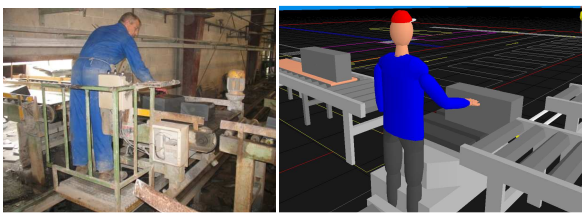


Figure 2: Slabs Arriving Process From Sawing. Real Process and QUEST Model

An operator on an electric rail mounted vehicle receives and distributes slabs among the splitters according to the specified format and their stock level (Figure 2).

Slabs are taken by the splitters one by one and cut in several pieces by means of a special type of chisel so they can handle them better and also determine its quality. Then, they change to a smaller chisel for cutting these parts into plates The chisel, placed in position against the edge of the block, is lightly tapped with a mallet; a crack appears in the direction of cleavage, and slight leverage with the chisel serves to split the block into two pieces with smooth and even surfaces. This is repeated until the original block is converted into a variable number of pieces. The resulting number of slates of different formats is variable, depending mostly on the quality of the slate rock from quarry as well as the splitters experience and skill.

Figure 3: A Splitter (left) and the Resulting Output: The Target Formats (regular lots in the left) and Secondary Less Quality Output Formats (the two series in the right).

A second operator collects the slates lots produced by the splitters on a second electric trolley and takes them to a third one who carries and distributes them amongst the cutting machines. Split stone is then mechanically cut according to the shape and size required. This operation is done both by manual and fully automated cutting machines.

Finally, slate presented is inspected one by one by classifiers with a trained eye prior to being placed in crate pallets. Slate that does not meet with quality requirements is set aside and recycled to be cut again into another shape until it complies with company standards. In case this is not possible, it is rejected. Slate pieces are packed until they are ready for final use. Slates are available under different sizes and grades. Quality is assessed in terms of roughness, colour homogeneity, thickness and presence and position of imperfections – mainly quartzite lines and waving-. Accordingly, the company offers three grades for every commercial size: Superior, First and Standard

Alternatively, the latter operator takes the recycled plates and transports them to their corresponding machines. A third task assigned to this labour is to stock material in buffers previously located to the machines' inputs whenever machines' utilization is full. So a triple flow is shared by one transportation system connecting a push system (lots coming from splitters) and a pull system (lots required by cutting machines). And even more, the assignation rules that the operator follow depend on his criterion, so it is easily comprehensible the complexity of modelling this system.

Figure 4: Distribution of Lots to Cutting machines.

A general process diagram is represented in Figure 5. Arrows in grey and triangles in red represent transportation and stocking operations respectively. These operations do not add value to the product whereas green circles and squares represent value-added operations, mainly transformations and/or inspections.

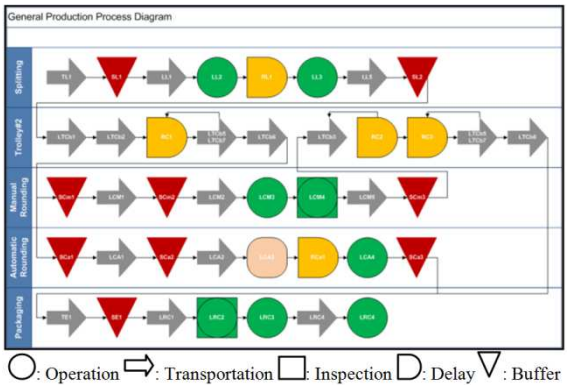


Figure 5: Complete Process Diagram.

The General Process Diagram offers at first glance a quick assessment of the abundance of these operations as well as the presence of feedback lines which also diminish the overall process performance. It becomes clear the necessity of reducing non value-added activities and rearranging the whole process in terms of layout design.

This process characterization has been employed in the definition of elements, processes and flows (both logical and physical) for the simulation model.

3. THE PROBLEM: VARIABILITY

Natural roofing slate manufacturing processes are subject to the intrinsic variability of natural slate. This vari-

ability corresponds with the possibility of variations both in mineral composition and morphology so that undesirable visual and structural effects in the final product may appear. It is the geological nature of the specific zone in the quarry that is eventually being exploited which determines this circumstance. Although there is certain knowledge about the quality of rock that is expected to extract in the quarry according to previous experience and/or mineral exploration operations, it is not possible to determine the real continuous mineral profile at a microscopic or visual level.

This uncertainty about the final quality has traditionally configured the whole manufacturing process resulting in a reactive system, that is, a system where there is no previously determined schedule and the assignment of operations to machines or labours is done according to the state of the system (Alfaro and Sepulveda 2005).

In our case, a foreman dynamically decides the formats to be cut as well as the number and identity of splitters, classifiers and machines assigned to each format according to his perception of process performance. Eventually, the functions performed and messages sent are allowed to adapt such that feedback paths in the process occur. Then, the overall system may exhibit emergent behaviours that cannot be produced by any simple subset of components alone, defining a complex system (Clymer 2009).

When proposing modifications in these systems special care has to be taken since even small changes in deterministic rules (SPT, FIFO, etc.) may result in a chaotic behaviour. Developing DES models of such processes has been proposed as a systematic way for its characterization and analysis (Alfaro and Sepulveda 2005). However, DES projects rely heavily on high input data quality (Leemis 2000). Accordingly, the input data phase constitutes on average one third of the total time of simulation projects (Skoogh and Johansson 2008). What it plays a negative role in the acceptation of simulation methods as a tool for improvement as long as it leads to unacceptable lead-times when dealing with well defined processes (Skoogh and Johansson 2007) can be turned into an advantage when facing a process in an early stage of statistical control.

In our plant, there is not such a thing as an organized and structured information system but heterogeneous and incomplete sources of data from which our input data management process may start. Only two data collection points had been implemented and they did not properly connect inputs and outputs between different operations, so traceability of products was not available. But above all, the dynamic nature of output formats according to a changing commercial strategy made really difficult to rely on a stable basis of data from which building useful information for an eventual simulation validation phase.

Adopting a Modelling and Simulation approach (M&S) for the slates process characterization provides an efficient and systematic procedure for this purpose as well as the necessary time for carrying out the definition of a data management system.

Thus, the modelling process has been employed as a means for determining and defining such sources of data and has finally lead to the proposal of a Data Acquisition System (Figure 6) consisting in the definition of control points and procedures by means of which subsequent stages of the classical MS methodology can be developed. The system involves the assignation of responsibilities, the implementation of control procedures and the computer management of data.

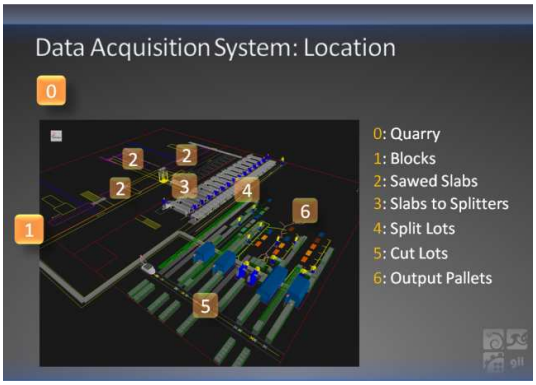


Figure 6: Location Points for Data Acquisition.

This constitutes a first attempt to distinguish between the two types of process variation (Deming 1986), i.e., Common Cause variation, which is intrinsic to the process and will be always present, and Special Cause variation, which stems from external sources and indicates that the process is out of statistical control.

In consonance with this purpose, we suggest the Product, Process and Resources (PPR) concept as an integrating approach to variability aiming at carrying out a complementary and parallel process of modelling and analysis. It is the Dassault Systèmes' integrated model that interlinks representations of the Product, the manufacturing Resources (tooling, factory, operators, etc.) and the production Process (DS 2009). In Figure 7 the identity and flow between sources of variability is depicted.

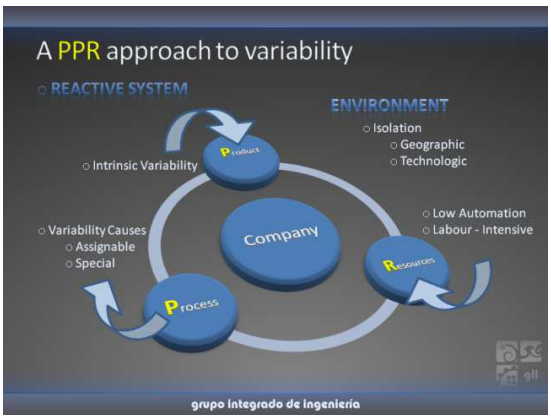


Figure 7: A PPR Approach to Variability

As it has already been explained, the product introduces its natural variability in the process, which is on its part affected by the environment. The utilization of human resources in physical and decision making processes relative to the product also involves a new source of variation. The resulting process is complex, reactive and out of statistical control.

Achieving statistical control will lead to improved levels of productivity and allow a proactive commercial management. An initial control system should permit the assessment of changes in transportation systems, layout design and other operational parameters. A final stage involves the implementation of a QC system by means of Artificial Vision Techniques for a more robust and powerful statistical control process. In the next diagram the proposed milestones in order of complexity and financial effort are resumed.

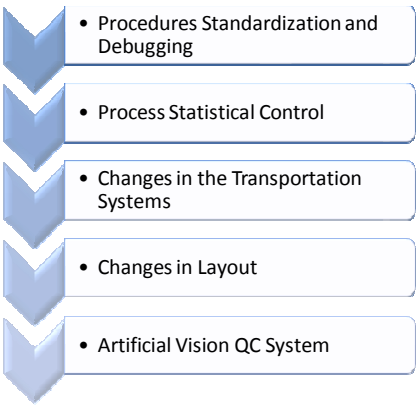


Figure 8: The Global Improvement Project: a Growing Set of Goals.

4. STATISTICAL ANALYIS

Once processes have been identified it was possible to start a data analysis phase. Data have been collected from two main sources. The first has been the analysis of recorded videos and other measurements performed during periodical visits to the plant. The second set of data was obtained from historical records provided by the company. Statistical tests of independence, distribution fit and goodness of fit of the available real data have been conducted. We have employed the MLE method for model's variables distribution fitting using the statistical software StatFit and R (R 2005). Distribution verification and selection is done according to the p-values obtained in the Kolmogorov-Smirnov test and by graphical inspection of the QQ and PP graphs of a wide range of typical distributions in the simulation field.

The process' core is the splitting operation. It is the task where product, resources and process circumstances converges in a less controllable way from a variability point of view. A piece of evidence is shown in Figure 9, where data collected during a period of three months of the number of lots produced per splitter is depicted. The presence of different splitting patterns corresponds with personal performance and the variability within a single pattern is an evidence of product variability.

Besides, this process is a constant in all slate manufacturing plants regardless process configurations so methodology and results are extensible.

A splitting process model has been made in order to assess upstream and downstream impacts on transportation processes and stock levels. To do so, we focused on its inputs and outputs, i.e. slabs arriving from Sawing and piles leaving from Splitting.

After an ABC classification of the number of slates produced, we could divide into two groups all the different formats that are produced in the plant for the sake of simplicity. One is the target format of tiles of 32x22x3.5 mm –the 80% of the total, that we named L32 (Slates Lots of 32) – and the other one gathers all other formats in a category named LN32 (Lots of Not 32). In addition, this two outputs model is in line with the assessing of changes in the simulation models focused on increasing levels of target format output whilst reducing levels of the rest of outputs.

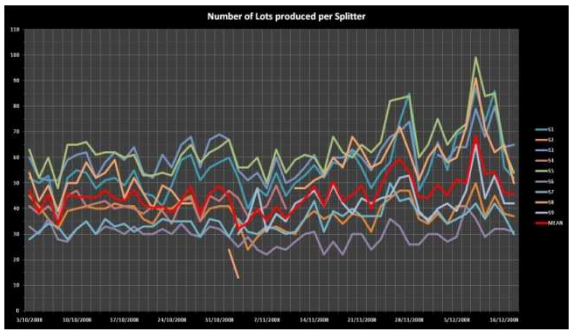


Figure 9: Evolution of Splitters' Production (three months of historical data)

We have modelled the generation of lot sizes and splitting times per slab by means of three random variables: the Splitting time per slab, the Deviation in the number of plates, and the Proportion of target plates.

These three variables are related. A higher splitting time can be associated with the processing of bigger slabs resulting in a higher number of plates. However, a perfect association between the splitting time and the number of plates is not acceptable since it is a common situation a bigger slab –and consequently a higher time– being partly dismissed due to its low quality and giving fewer plates than a smaller one but with better quality. The most suitable way to model this variability is generating such variables from its estimated joint probability distribution. Independence hypothesis was not accepted nor the perfect association one as we would be losing the relationship between slabs splitting rate and the number of plates produced or not considering that source of variability, respectively.

As simulation aims at providing information about the influence of changes in both sides of the splitting process it is necessary to build a model in which arriving slabs become slates lots in a realistic way.

An analysis of correlation between these variables was then made leading to a model for randomly generating processing times per slab and their corresponding lot sizes. As we had more data of processing slab times than of lot sizes, the random generation is done by obtaining first a splitting time from which a depending number of plates is generated together with the corresponding fraction of L32 plates according to the fitted distributions. Distribution fitting results for the splitting time, the estimation of total number of plates and distribution of residuals are summarized in the following figures. We could not find a good fit for the marginal distribution of L32 fraction so we use an empirical table from collected data.

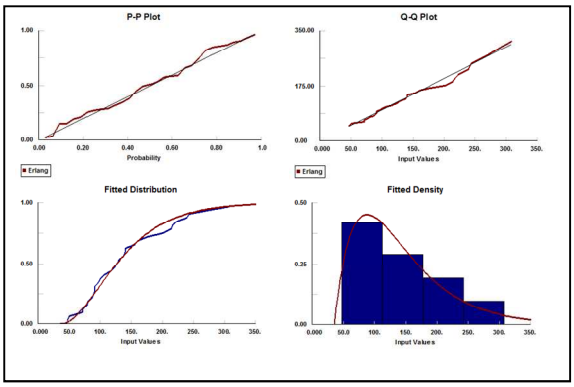


Figure 10: Erlang Distribution Fit of Splitting Time (k=2, min = 35s, beta = 53.18))

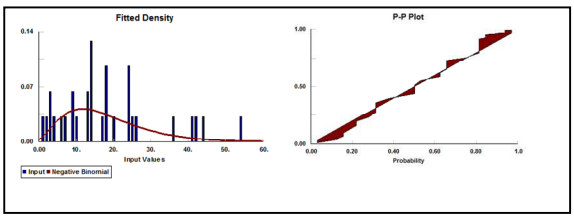


Figure 11: Negative Binomial Distribution Fit of Total Number of plates fit (k=3, p=0.141)

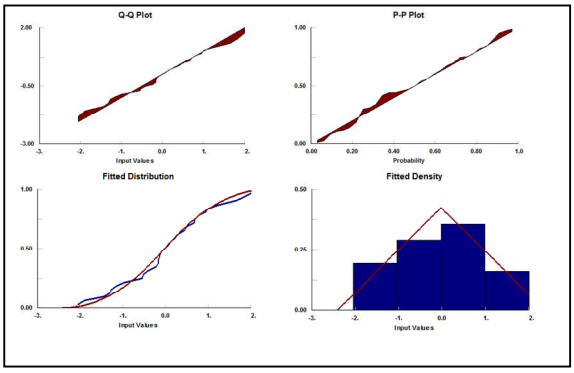


Figure 12 Triangular Distribution Fit of Residuals in Total Number of Plates Estimation (minimum =-2.41771, maximum = 2.37337, mode =-2.62736e-002)

At this point, the question about the quality of our input data and the worthiness of continuing data collection activities arises. One way to tackle with this uncertainty is by means of simulation. As explained in the next chapter, a splitting sensitive analysis model was proposed and results lead to an interesting and simplifying conclusion.

5. SIMULATION MODEL

We have built a simulation model of the manufacturing plant in DELMIA QUEST. DELMIA stands for Digital Enterprise Lean Manufacturing Interactive Application and it is the Dassault Systèmes' solution for engineering lean manufacturing processes. QUEST is a powerful three dimensional simulation environment by means of which a model can be comprehensively defined and assessed.

The classic model's definition paradigm establishes that a model that describes a system as network of system components and focuses on the physical details is called architectural. A model that describes a system as a network of functions and focuses on the behavioural details is called functional flow. The functional flow model is used during the conceptual system design phase of a system design project whereas the architectural model, after functions have been allocated to the system components or subsystems, is used during the system design phase (Clymer 2009).

Unlike this classic model's definition paradigm, the model's building process in QUEST is accomplished by an integrated development environment combining both the functional and architectural definition. The PPR approach is utilized as a way for the definition of model's elements, i.e., products (different types of parts, in QUEST terminology, representing different stages of slate transformation), resources (labours, transportation systems, and machines) and the process logical set up (controllers and logic recipients). This allows the progression of model building when data sources are variable or not well established and a direct implementation of the previous process depicting work.

In fact, since operational similarity is obtained by assuring geometric and kinematic similarity, model verification and assessment of changes in layout design can easily and quickly be noticed.

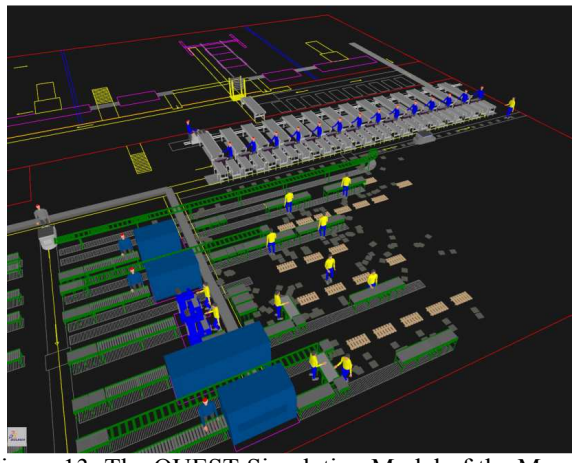


Figure 13: The QUEST Simulation Model of the Manufacturing Process.

Initial experimentation has included the assessment of the three different types of classifying arrangements that were actually being employed in the plant (Figure 15 and Figure 16). Every machine has two dedicated labours assigned for the quality inspection, classification, recirculation and packaging. Depending on their location, labours adopt a different position and split tasks in a different manner. There is not a certain reason why different operational modes are adopted.

The first mode consists of a confined labour performing classification tasks and the other one only performing packaging operations. The term confined refers to the fact that the labour is totally surrounded by rolling conveyors so he is not able to support his team mate whenever he is idle. In the second schema the classifier is not confined and thus can assist to his packer mate when idle. Finally, the third type is the one in which the two labours carry out both classification and packaging operations independently. The packaging process is made up by the loading, unloading, and transportation and return tasks.

These logical behaviours have been coded in SCL language and implemented in the model. Positions of labours and pallets as well as paths followed during their movements have been identically proposed as the real ones so a geometric assessment may be carried out. In Table 1 results obtained from a set of 50 simulations for every operational mode are summarized. Specifically, the number of lots processed, the time of tasks' completion, the total distance and the utilization ratio depending on mode of operation are presented.

Table 1: Time in hours and Distance in meters for every Task depending on Operational Mode

	Mode 1		Mode 2	
	Mean	Std.Dev	Mean	Std.Dev
	Packer			
Lots	363.4	2.82%	61.2	41.84%
Load	0.304	3.75%	0.186	152.81%
Unload	1.034	4.86%	2.552	3.50%
Transport.	0.768	2.97%	0.08	42.39%
Return.	0.672	3.39%	0.096	40.75%
Distance	3049	2.75%	638.6	40.15%
Util.(%)	100	0.00%	100	0.00%
	Classifier			
Lots			337.2	11.54%
Load			0.284	11.30%
Unload			1.102	13.03%
Transport.			0.718	8.89%
Return.			0.674	7.68%
Distance			2943.8	8.04%
Util.(%)	83.2	2.74%	100	0.00%

Mode 3	
Mean	Std.Dev
Packer & Classifier	
181.308	3.94%
0.174	3.15%
1.642	2.56%
0.312	4.18%
0.648	3.68%
2043	3.73%
100	0.00%

In Figure 14, the distribution of the time spent in the packaging's tasks is depicted. It comes out that transportation and return tasks consume on average half of the working time spent by the labours performing this job. Once again, these are non value-added operations and will have to be minimized by a convenient layout redesign.

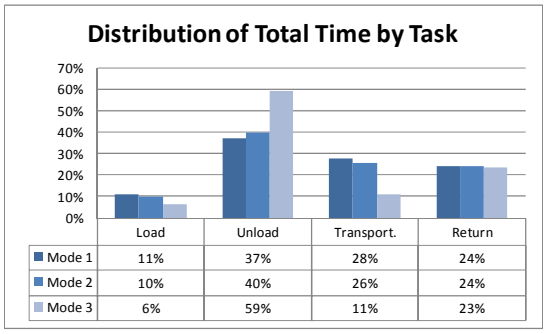


Figure 14: Time Distribution between Tasks depending on Operational Mode

Table 2: Results of the Assessing Classifying Modes Simulations

Mode	1	2	3
Production Rate (tiles/h)	1439.06	1577.66	1435.96
Standard Deviation	2.82%	3.37%	1.97%
Distance per working hour (km)	1.10	1.29	1.47

The results shown in Table 2 concluded that the best option is the intermediate one, that is to say, a classifier that may support packing tasks when idle. It implies an almost 10% improvement in productivity respect to the other two schemas and only requires a cheap and quick rearrangement of rolling conveyors at the machines' output area. This result is of special relevance since this final process stage actually becomes the bottleneck in those days when the slate coming from the quarry is good.

Figure 15: The Different Modes of Classifying Operation

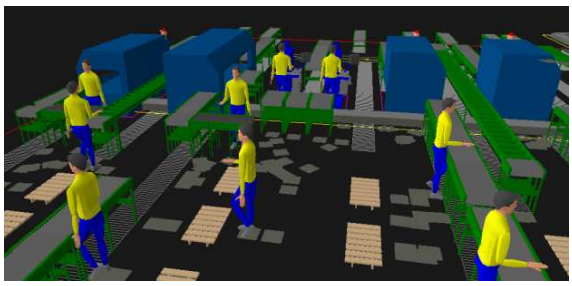


Figure 16: The Classifying Operation. The Three Different Arrangements Have Been Simulated

Another set of simulations was performed as a means of determining the worthiness of deepening in the statistical process analysis. This way, simulation has been employed to assess the influence of the assumption of splitting variable's independence versus different levels of correlation on queues and second transportation system utilization (splitters output).

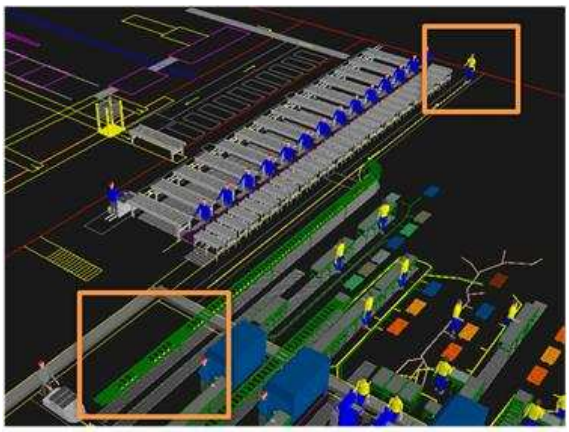


Figure 17: The Splitting Simulation Model. The Squared Areas Are the Points of Interest Where Assessing Results

Three different scenarios have been then proposed: the splitting linear model previously described resulting on a correlation of r=0.806, the absence of correlation (r=0) and a perfect correlation situation (r=1).

Conducted experiment consisted of the simulation of a typical shift of 5 hours. Initial conditions are full buffer levels of blocks from sawing that the trolley has delivered after the end of the last turn. At the end of each simulation, average buffer level before cutting machines and splitters' pick up trolley utilization are measured. For each level 230 replications are made. To check differences on buffer level distribution, Wilkoxon difference's test has been applied due to lack of normality, unlike the trolley utilization, on which the F-test could be applied.

The results do not show any significant difference between the independence assumption and the developed correlation model proposed as it is shown in the following box plots.

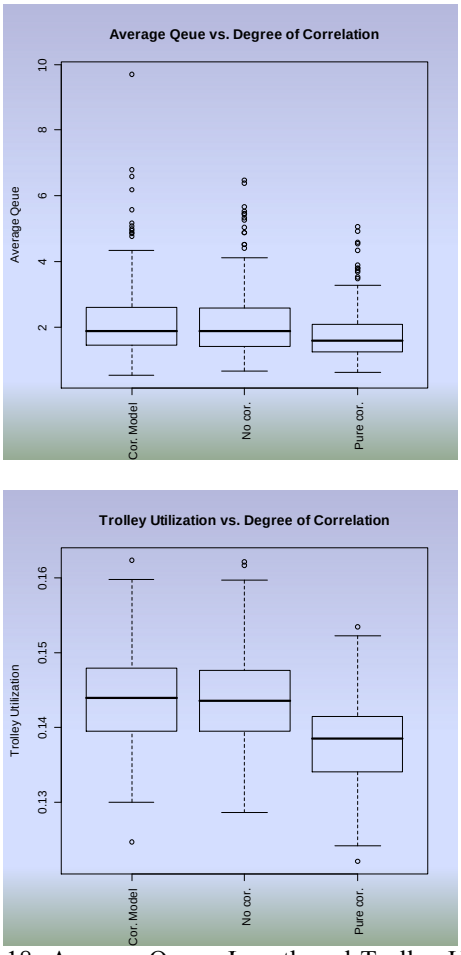


Figure 18: Average Queue Length and Trolley Utilization versus Degree of Correlation

Only in the case of perfect correlation differences are met. An arguably explanation might be that the first two models imply a higher variability on model's behaviour, and thus queues generated are bigger. The conclusion of this result is that it is unnecessary to include the correlation model when attending to study the effect of splitter's behaviour on the downstream steps of the process. It would be enough to accept the independence hypothesis and so employ the marginal distributions. This result is then considered in our statistical analysis so further effort in data analysis was discarded.

CONCLUSIONS

An M&S approach has been proposed as a means to characterize a poor structured and highly variable system. The model's building process allows a better understanding of real processes as well as the definition of following stages in the own simulation project. In the

case of the natural roofing slates sector, this is a total innovative initiative that is being accomplished.

A splitting statistical model has been developed and by a combined use of early simulation experimentation the extent of statistical analysis has been limited so the model verification and validation can reliably go on. Besides, as a result of simulation, a simple and cheap recommendation regarding classification policies has been proposed leading to increased productivity in the final process stage.

Finally, further research is necessary in order to complete the whole systems' characterization and once a database is built, to define experimentation for a validating simulation environment for process optimization, both operational and morphological.

ACKNOWLEDGMENTS

We wish to express our gratitude to the Xunta de Galicia, which has funded this work through the research project "Análise mediante simulación para a optimización do proceso de fabricación de lousas" (IN841C-2008/108).

REFERENCES

Alfaro, M.D., Sepulveda, J.M., 2006. Chaotic behavior in manufacturing systems. *International Journal of Production Economics*, 101, 150–158.

Clymer, J.R., 2009. *Simulation-Based Engineering of Complex Systems*. 2nd ed. Hoboken (USA): Wiley.

Dassault Systemmes, 2009. *PLM Glossary*. Available from: http://www.3ds.com/plm-glossary/ [accessed 24 June 2009]

Deming, W.E., 1986. *Out of the crisis*. MIT Press.

Leemis, L., 2000. Input Modeling. *Proceedings of the 2000 Winter Simulation Conference*. December 10-13, Orlando (Florida, USA).

R Development Core Team, "R: A Language and Environment for Statistical Computing", R Foundation for Statistical Computing. Vienna, Austria, 2005.

Skoogh, A., Johansson, B., 2008. A Methodology for input data management in discrete event simulation. *Proceedings of the 2008 Winter Simulation Conference*. December 7-10, Miami (Florida, USA).

Skoogh, A., Johansson, B., 2007. Time-consumption analysis of input data activities in discrete event simulation projects. *Proceedings of the 2007 Swedish Production Symposium*. Gothenburg (Sweden).

AUTHORS BIOGRAPHY

David del Rio Vilas holds an MSc in Industrial Engineering (Ingeniero Industrial) and has been studying for a PhD since 2007. He works full-time in the research group GII of the University of A Coruna as a research engineer since 2007. He is mainly involved in R&D projects development related to industrial and logistical processes optimization. His area of major interest is the use of simulation in industrial and logistical environments, especially in maritime container terminals.

Diego Crespo Pereira holds an MSc in Industrial Engineering (Ingeniero Industrial) and he is currently studying for a PhD. He works full-time in the GII of the University of A Coruna as a research engineer since 2008. He is mainly involved in the development of R&D projects related to industrial and logistical processes optimization. He is especially interested in the field of dynamic vehicle routing problems.

Juan Luis Crespo Mariño graduated from the University of Santiago de Compostela in 1996 and has been working in the GII since 1998. He led the Artificial Vision Unit of the CIS-Galicia between 2003 and 2005. He is at present Adjunct Professor in the University of A Coruna and had previous experience in the University of Costa Rica between 2002 and 2003. He holds an MSc in Physics and a PhD in Industrial Engineering.

Alejandro Garcia del Valle is an Industrial Engineer by the Polytechnic University of Madrid and holds a PhD in Industrial Engineering. He is Professor and one of the founders of the GII at the University of A Coruna. He has led many research projects about simulation and optimization of industrial, logistical and service systems.

PRODUCTION PROCESS MODELING AND PLANNING WIH SIMULATION METHOD, MOUNTING PROCESS OPTIMISATION

János, Jósvai[(a)]

[(a)] Széchenyi István University, Professorship for Material Sciences and Automobile Production

[(a)]josvai@sze.hu

ABSTRACT
The paper focuses on the establishment of the production program using simulation technology in a structure, where several products and high amount of variants per product are produced. The topic of the paper addresses the discrete event simulation technology which is used to model the material flow and the manufacturing processes in the production area. This paper would like to show and describe the modelling steps of a complex production system with a lot of products and three different line parts, which are connected with buffers.

Keywords: discrete event simulation, production planning, scheduling, genetic algorithm.

1. INTRODUCTION

Today the production tasks have got a very complex planning process. This is caused by the high amount of variants of one product. We can speak here about a vehicle or engine production. Most of the production structures are established as lines and have the task to produce several product types and several variants of the products. This means a very difficult planning and execution of production. The establishment of the production program is complicated, the times of work tasks are different, and the material delivery on the line and the inventory has to be taken into consideration, too.

The production planning has several goals, some of them are:

- the scheduling of the tasks to ensure delivery accuracy,
- to determine the lot size of product batches,
- to ensure smoothed workloads at the workplaces,
- to determine the buffer sizes in the production line,
- to handle the lead times – depending on the complexities of the products,
- to determine and handle the bottlenecks – can change with the system dynamic behaviour, etc.

Mostly the production system is not configured as a whole integrated line. To plan a system, which is separated by buffers between two or perhaps three main lines, has a lot of influential parameters. The main question is either to plan these part lines together, or to plan the production on the lines separately because of some reasons. For example if the mean cycle time is different on the lines then this could be a reason to make the planning separately.

These properties show the complexity of this field. The influence parameters are not only a large number, but the combination of these parameters causes a lot of option and problems to solve. In practice there is not enough time to fulfil the mathematical analysis manually, even if the right behaviour functions are ready to use.

There is another possible method which is useful to plan such complex systems. The modelling and dynamic simulation are able to answer most of the questions, and show the time dependent behaviour of the concerned production system. This modelling technique is the time discrete event controlled simulation.

This paper would like to show and describe the modelling steps of a complex production system with a lot of products and three different line parts, which are connected with buffers.

2. PROBLEM DEFINITION

The considered production system was an engine production line with three separated line parts. These were connected by buffers. The simulation model and study had to investigate, how the line output, usage statistics changes with the different production sequences.

The product mix changes time to time, this had many influences and plus tasks while the planning of the model. We will see how it works when a product has to be changed in the model. This could mean for instance the end of production of one product type, or new type has to be launched on the line. This data handling procedure and the amount of handled data causes a great model size.

The modelling had to consider, that a lot of flexible parameters were needed to ensure enough planning roam. Lot size determination had to be fixed, that the actual pre-planned production program could be changed and set on new levels by the simulation.

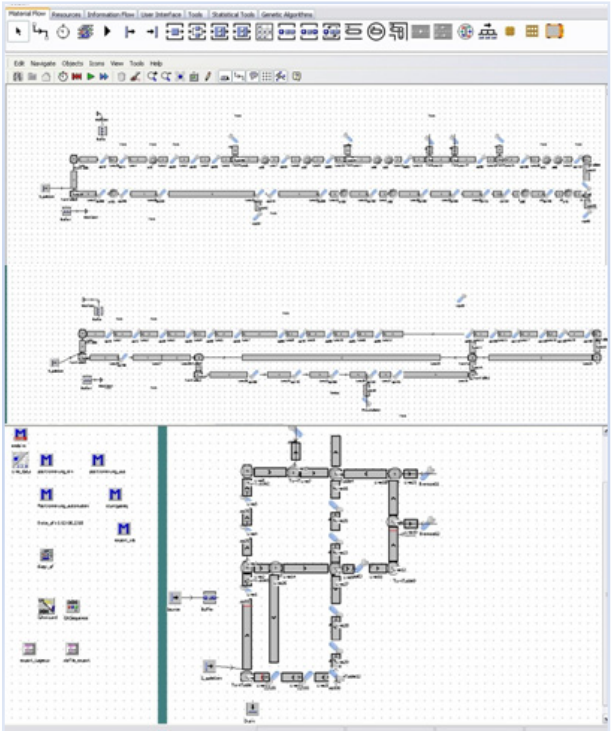


Figure 1: Simulation models of the line parts

Another main goal was to determine the computational achievable "right" production sequence. The hand-made production program should be optimized by the simulation. A genetic evolution algorithm was used to solve this difficult problem with a large search area.

For planning the line balancing there was needed an option, to ensure handling functionality, when workload change has to be planned. The mounting tasks can be assigned to various places in the line. This means that the variation of workloads at the stations in the line has a large number. The line balancing has the goal to put the tasks in the right order after each other and approximately hold the average cycle time at one station. In case of production changes - product type, produced volume, technological, and production base time – there was a need to pre-calculate the changed line behaviour. There are different changes in the task load of the stations, we make such influences which determine the throughput, working portion of the stations and gives different optimal sequence combination of products.

3. SIMULATION AND SCHEDULING

There are similarities and differences as well between general research- and simulation case studies. Simulation case studies are typically focused on finding answers to questions through simulation-based experiments. In the social science arena, experimentation is considered to be a distinct research method separate from the case study. Social science case study researchers use observation, data collection, and analysis to try to develop theories that explain social phenomena and behaviours. Simulation analysts use observation and data collection to develop "as-is" models of manufacturing systems, facilities, and organizations. The analysts test their theories and modifications to those models through simulation experiments using collected data as inputs. Data sets may be used to exercise both "as-is" and "to-be" simulation models. Data sets may also be fabricated to represent possible future "to-be" conditions, e.g., forecast workloads for a factory. (McLean 2003.).

In (Standridge 2000.), teaching simulation through the use of manufacturing case studies is discussed. He organizes case studies into four modules:

- Basic manufacturing systems organizations, such as work stations, production lines, and job shops.
- System operating strategies including pull (just-in-time) versus push operations, flexible manufacturing, cellular manufacturing, and complete automation.
- Material handling mechanisms such as conveyors, automated guided vehicle systems, and automated storage/retrieval systems.
- Supply chain management including automated inventory management, logistics, and multiple locations for inventory.

Simulation case study problem formulations and objectives define the reasons for performing the simulation. Some examples of study objectives might be to evaluate the best site for a new plant, create a better layout for an existing facility, determine the impact of a proposed new machine on shop production capacity, or evaluate alternative scheduling algorithms. (McLean 2003.)

Simulation textbooks typically recommend that a ten to twelve step process be followed in the development of simulation case studies. The recommended approach usually involves the following steps: (1) problem formulation, (2) setting of objectives and overall project plan, (3) model conceptualization, (4) data collection, (5) model translation into computerized format, (6) code verification, (7) model validation, (8) design of experiments to be run, (9) production runs and analysis, (10) documentation and reporting, and (11) implementation (Banks et al. 1998).

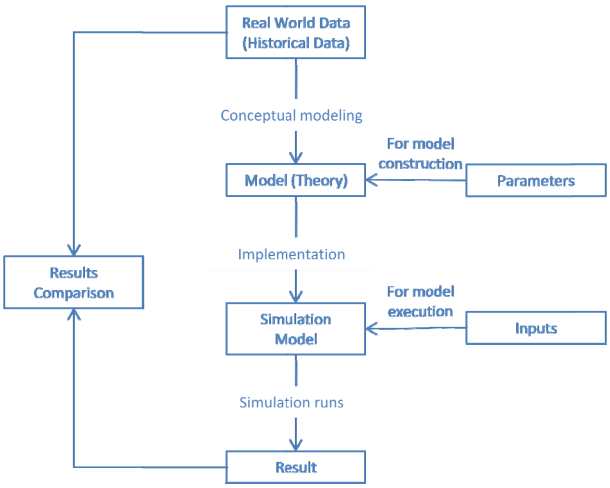


Figure 2: Simulation modelling and executing steps (Shao 2008.)

What is manufacturing simulation? In The Handbook of Simulation, Jerry Banks defines simulation as:

"…the imitation of the operation of a real-world process or system over time. Simulation involves the generation of an artificial history of the system and the observation of that artificial history to draw inferences concerning the operational characteristics of the real system that is represented. Simulation is an indispensable problem-solving methodology for the solution of many real-world problems. Simulation is used to describe and analyze the behaviour of a system, ask what-if questions about the real system, and aid in the design of real systems. Both existing and conceptual systems can be modelled with simulation." (Banks 1998.)

Manufacturing simulation focuses on modelling the behaviour of manufacturing organizations, processes, and systems. Organizations, processes and systems include supply chains, as well as people, machines, tools, and information systems. For example, manufacturing simulation can be used to:

- Model "as-is" and "to-be" manufacturing and support operations from the supply chain level down to the shop floor
- Evaluate the manufacturability of new product designs
- Support the development and validation of process data for new products
- Assist in the engineering of new production systems and processes
- Evaluate their impact on overall business performance
- Evaluate resource allocation and scheduling alternatives
- Analyze layouts and flow of materials within production areas, lines, and workstations
- Perform capacity planning analyses
- Determine production and material handling resource requirements
- Train production and support staff on systems and processes
- Develop metrics to allow the comparison of predicted performance against "best in class" benchmarks to support continuous improvement of manufacturing operations (McLean 2002.)

3.1. Genetic Algorithms

An implementation of a genetic algorithm begins with a population of (typically random) chromosomes. One then evaluates these structures and allocates reproductive opportunities in such a way that those chromosomes which represent a better solution to the target problem are given more chances to reproduce than those chromosomes which are poorer solutions.

The goodness of a solution is typically defined with respect to the current population. This particular description of a genetic algorithm is intentionally abstract because in some sense, the term genetic algorithm has two meanings. In a strict interpretation, the genetic algorithm refers to a model introduced and investigated by John Holland (1975) and by students of Holland (e.g., DeJong, 1975). It is still the case that most of the existing theory for genetic algorithms applies either solely or primarily to the model introduced by Holland, as well as variations on what will be referred to in this paper as the canonical genetic algorithm. Recent theoretical advances in modelling genetic algorithms also apply primarily to the canonical genetic algorithm (Vose, 1993).

In a broader usage of the term, a genetic algorithm is any population-based model that uses selection and recombination operators to generate new sample points in a search space. Many genetic algorithm models have been introduced by researchers largely working from an experimental perspective. Many of these researchers are application oriented and are typically interested in genetic algorithms as optimization tools. (Whitley 1995)

The use of genetic algorithms requires five components:

1. A way of encoding solutions to the problem - fixed length string of symbols.
2. An evaluation function that returns a rating for each solution.
3. A way of initializing the population of solutions.
4. Operators that may be applied to parents when they reproduce to alter their genetic composition such as crossover (i.e. exchanging a randomly selected segment between parents), mutation (i.e. gene modification), and other domain specific operators.
5. Parameter setting for the algorithm, the operators, and so forth. (Jones 1996)

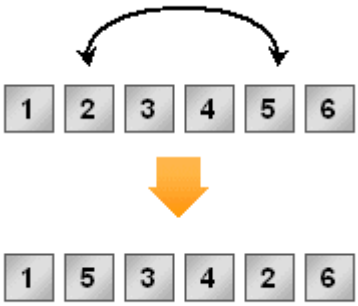


Figure 3: Mutation for a sequential task (Tecnomatix 2008)

The simulation model uses the genetic algorithm for a sequential task. The logic to produce a new population is shown on Figure 3. Several test runs were made in order to identify the right settings of the algorithm. The statistical operators were configured after real life data test runs, to make the algorithm converge faster. The runs showed at last, that the population size has to be set to 10 and the simulated generations' numbers were 20. This was a main question among others, because the simulation running time was limited up to one and half an hour.

3.2. Scheduling

Scheduling has been defined as the art of assigning resources to tasks in order to insure the termination of these tasks in a reasonable amount of time. The general problem is to find a sequence, in which the jobs (e.g., a basic task) pass between the resources (e.g., machines), which is a feasible schedule, and optimal with respect to some performance criterion. A functional classification scheme categorizes problems using the following dimensions:

1. Requirement generation,
2. Processing complexity,
3. Scheduling criteria,
4. Parameter variability,
5. Scheduling environment.

Based on requirements generation, a manufacturing shop can be classified as an open shop or a closed shop. An open shop is "build to order", and no inventory is stocked. In a closed shop the orders are filled from existing inventory.

Processing complexity refers to the number of processing steps and workstations associated with the production process. This dimension can be decomposed further as follows:

1. One stage, one processor
2. One stage, multiple processors,
3. Multistage, flow shop,
4. Multistage, job shop.

The one stage, one processor and one stage, multiple processors problems require one processing step that must be performed on a single resource or multiple resources respectively.

In the multistage, flow shop problem each job consists of several tasks, which require processing by distinct resources; but there is a common route for all jobs.

Finally, in the multistage, job shop situation, alternative resource sets and routes can be chosen, possibly for the same job, allowing the production of different part types.

The third dimension, scheduling criteria, states the desired objectives to be met. They are numerous, complex, and often conflicting. Some commonly used scheduling criteria include the following:

1. Minimize total tardiness,
2. Minimize the number of late jobs,
3. Maximize system/resource utilization,
4. Minimize in-process inventory,
5. Balance resource usage,
6. Maximize production rate.

The fourth dimension, parameters variability, indicates the degree of uncertainty of the various parameters of the scheduling problem. If the degree of uncertainty is insignificant, the scheduling problem could be called deterministic. For example, the expected processing time is six hours, and the variance is one minute. Otherwise, the scheduling problem could be called stochastic.

The last dimension, scheduling environment, defined the scheduling problem as static or dynamic. Scheduling problems in which the number of jobs to be considered and their ready times are available are called static. On the other hand, scheduling problems in which the number of jobs and related characteristics change over time are called dynamic. (Jones 1998)

According to the previous classification the modelled system can be classified as:

- Open shop
- Multistage, flow shop
- The processing times are treated as deterministic
- Job characteristic is dynamic

4. MODELING AND SIMULATION RUNS

This model is a planning tool which is able to answer several questions of the complex production planning. The creation of the model followed the physical parameters of the real system. The iteration process of the modelling was difficult because it had to handle the product mounting time. The mounting times were gained from the real production system, but the collection and filtering was made inside the simulation model, to prepare the data ready for production inside the simulation.

4.1. Model building

Plant Simulation provides a number of predefined objects for simulating the material flow and logic in a manufacturing environment. There are five types of main object groups from Plant Simulation:

- **Material flow objects**: Objects used to represent stationary processes and resources that process moving objects.
- **Moving objects**: Objects used to represent mobile material, people and vehicles in the simulation model and that are processed by material flow objects. Moving objects are more commonly referred to as MUs.
- **Information flow objects**: Objects used to record information and distribute information among objects in the model.
- **Control objects**: Objects inherently necessary for controlling the logic and functionality of the simulation model.
- **Display and User interface objects**: Objects used to display and communicate information to the user and to prompt the user to provide inputs at any time during a simulation run.

SimTalk is the programming language of Plant Simulation; it was specifically developed for application in Plant Simulation models. The Method objects are used to dynamically control and manipulate models. SimTalk programs are written inside method objects and executed every time the method is called during a simulation run.

The logical structure of the model was created on basis of Plant Simulation provided level structure. So it was a "simple" planning step to divide the model into specified functional levels. Different folders and frames are used in order to implement the line structure, the data handling for manufacturing programs and the basic data for the manufactured products. However, the scheduling of the production program has its own separate level.

The data input and output of the model work with the Excel Interface of Plant Simulation. Users can manipulate the parameter settings and see the results of the simulation runs on this easy way independently from Plant Simulation – no special simulation knowledge is asked.

User interface has been implemented for the model in order to handle the simulation model and the several built-in functions, which are to test the simulated line behaviour. This handling tool, which is shown on Figure 4, helps the manufacturing engineer to plan tasks and solve rescheduling problems on the line.

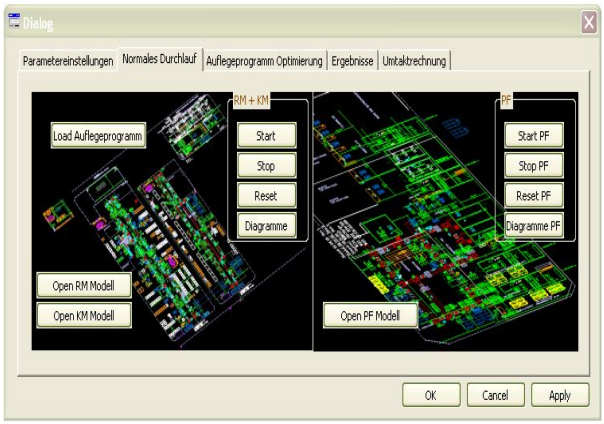


Figure 4: Simulation models of the line parts

4.2. Model validation and verification

Validation and verification of the model is formulated as follows:

Model validation: process of demonstrating that a model and its behaviour are suitable representations of the real system and its behaviour w.r.t. intended purpose of model application.

Model verification: process of demonstrating that a model is correctly represented and was transformed correctly from one representation form into another, w.r.t. transformation and representation rules, requirements, and constraints. (Rabe 2008)

There are many techniques to validate and verify the model. The physical environment has high influence on the method which is adaptable to verify and validate the model. In this particular case together with experts from the enterprise a structured walkthrough was possible to use for this system model. For special throughput data of the line it was possible to make historical data validation.

4.3. Simulation runs and results

The regular use of the simulation was secured with the several setting function, among them the line speed, the different value setting of the palettes on the separated lines, lot size limitations, and daily production program definition function.

With the simulation model it is possible to gain information about the system elements, for example how they are working in time, their occupation and empty time – waiting, etc. (Figure 5). Not only the elements can be obtained, but also the different working scenarios of the planned system load are about to be tested.

The simulation test runs with manufacturing data brought the following most important results:

- The simulation model is capable for everyday usage.
- To bring more efficiency 2-3 days are to be handled with the rescheduling algorithm.
- It is able to reduce lead time with 1-10%, this depends on product mixtures.

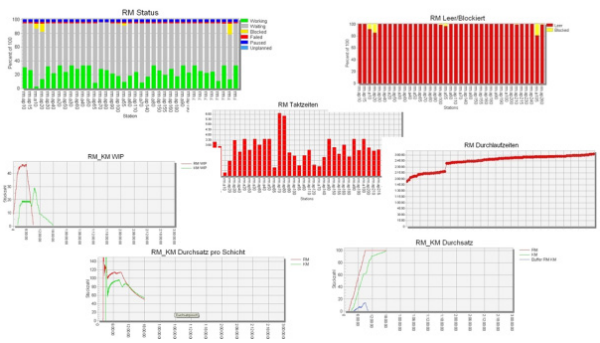


Figure 5: Evaluation diagrams of the modelled production system

The simulation model building and the test runs at the enterprise show that the simulation technique is suitable for the manufacturing planning. The model and the line connection mean in this case that the real data application could be made much better. This depends on both sides; the model structure has to be modified if the physical system is able to give over real time data. In this matter the rescheduling and the simulation tool could be not only the planning tool, but also it would be the production control tool.

5. CONCLUSIONS

The paper focuses on the applicability of simulation technology in production schedule of a production oriented firm and on the possibilities of planning and controlling the manufacturing process with simulation method in the automotive industry. A simulation model for manufacturing line planning and its establishment process is presented.

The paper discusses the questions of simulation and scheduling problems, these questions help to classify the physical system and the simulated problem.

Model validation and verification are taken into consideration after the presentation of the implemented genetic algorithm for production sequence optimization.

The most important benefits were highlighted based on the results of simulation runs.

REFERENCES

Edited by: Jerry Banks, 1998. *Handbook of simulation, Principles, Methodology, Advances, Application and Practice*. Atlanta:JohnWiley & Sons Inc.

Gudehus T., 1999. *Logistik Grundlagen Strategien Anwendungen*. Berlin : Springer.

Hopp, W. J., Spearman, M.L., 2008. *Factory Physics*. Boston : McGraw-Hill.

Jones, A., Riddick, F., Rabelo, L., 1996. *Development of a Predictive-Reactive Scheduler Using Genetic Algorithms and Simulation-based Scheduling Software*. National Institute of Standards and Technology, Ohio University. Available from: http://www.nist.gov [accessed 18 May 2009]

Jones, A., Rabelo, L., 1998. *Survey of Job Shop Scheduling Techniques*, National Institute of Standards and Technology, California Polytechnic State University. Available from: http://www.nist.gov [accessed 18 May 2009]

Kuhn A., Tempelmeier H., Arnold D., Isermann H., 2002. *Handbuch Logistik*. Berlin : Springer.

Rabe, M., Spieckermann, S., Wenzel, S., 2008. *Verifikation und Validierung für die Simulation in Produktion und Logistik*, Berlin:Springer.

McLean, C., Leong, S., 2002. *The Role of Simulation in Strategic Manufacturing*. Manufacturing Simulation and Modeling Group National Institute of Standards and Technology. Available from: http://www.nist.gov [accessed 18 May 2009]

McLean, C., Shao, G., 2003. *Generic Case Studies for Manufacturing Simulation Applications*. National Institute of Standards and Technology. Available from: http://www.nist.gov [accessed 18 May 2009]

Pfeiffer, A., 2007. *Novel Methods for Decision Support in Production Planning and Control*. Thesis (PhD). Budapest University of Technology and Economics.

Schönsleben, P., 2007. *Integrales Logistikmanagement, Operations and Supply Chain Management in umfassenden Wertschöpfungsnetzwerken*. Berlin : Springer.

Shao, G., McLean, C., Brodsky, A., Amman, P., 2008. *Parameter Validation Using Constraint Optimization for Modeling and Simulation*. Manufacturing Simulation and Modeling Group, National Institute of Standards and Technology. Available from: http://www.nist.gov [accessed 18 May 2009]

Standridge, C., 2000. Teaching Simulation Using Case Studies. *Proceedings of the 32nd on Winter Simulation Conference*, pp. 1630-1634. December 10-13, Orlando, (Florida, USA).

Tecnomatix Technologies Ltd., 2006. *Tecnomatix Plant Simulation Help*.

Vollmann, T.E., Berry, W. L., Whybark, D.C., Jacobs, F.R., 2005. *Manufacturing Planning and Control Systems for Supply Chain Management*. New York : McGraw-Hill.

Whitley, D., 1995. A Genetic Algorithm Tutorial. *Statistics and Computing* Vol. 4:65-85.

AUTHORS BIOGRAPHY

János Jósvai is working at the Széchenyi István University, Győr, Hungary. He is a PhD candidate, his field is manufacturing planning and simulation methods. He has several year experiences in material flow simulation of manufacturing systems and in production process planning. In the field of research and development he spent significant time aboard with international cooperation in matter of digital factory.

SELF-BUILDING SIMULATION TOOL FOR DAILY DECISION SUPPORT IN PRODUCTION CONTROL

András Pfeiffer[(a)], Botond Kádár[(a)], Marcell Szathmári[(a)], Gergely Popovics[(a,b)], Zoltán Vén[(a,b)], László Monostori[(a,b)]

[(a)] Computer and Automation Research Institute, Hungarian Academy of Sciences
Kende u. 13-17, Budapest, H-1111, Hungary
[(b)] Faculty of Mechanical Engineering, Budapest University of Technology and Economics, Budapest, Hungary

[(a)]pfeiffer@sztaki.hu, [(a)]kadar@sztaki.hu, [(a)]szmarcell@sztaki.hu, [(a,b)]popovics@sztaki.hu, [(a,b)]ven.zoltan@sztaki.hu, [(a,b)]laszlo.monostori@sztaki.hu

ABSTRACT
The paper introduces a decision support architecture with an integrated, self-building simulation module for the validation of the calculated manufacturing capacities, a-priori recognition of due date deviations and analysis of the effect of possible actions. In the underlying research special attention is devoted to the prediction and evaluation of the production on a daily, rolling time horizon basis (e.g., work in process (WIP) trajectories, machine utilizations).

The paper addresses the simulation module of a higher level integrated system in which the simulation model is automatically built o the basis of a real-time connection to Manufacturing Execution System. The main functionalities and advantages are highlighted through a real industrial case study.

Keywords: self-building simulation, production control, decision support, data mining

1. INTRODUCTION

In manufacturing systems, at the operational level, difficulties arise from unexpected tasks and events, non-linearities, and a multitude of interactions while attempting to control various activities in dynamic shop floors. The selection of the most appropriate production control decision for a given assignment, as well as the prediction of waiting times, workloads or utilisations of the resources are not trivial tasks, although they can be supported by simulation-based evaluations (Pfeiffer 2007, Bagchi et al. 2008, Watt 1998).

Therefore, based on previous results (Monostori et al. 2007), we propose a decision support architecture, in which the integrated, self-building simulation module can be applied for validation of the calculated manufacturing capacities, a-priori recognition of due date deviations and analysis of the effect of possible actions taken. In the research presented in this paper special emphasis is given to the prediction and evaluation of the production on a daily, rolling time horizon (e.g., Work in process (WIP) trajectories, machine utilizations).

The paper addresses the simulation module of the proposed architecture highlighting its main functionalities and advantages through a real industrial case study. An important issue regarding short-term (operational level) simulation is the automatic collection and definition of simulation input data. Therefore a new operation time definition method is presented in the paper, as well as a self-building simulation procedure is described in details.

2. PRODUCTION CONTROL AND SIMULATION – BACKGROUND

2.1. Simulation-supported production control

The discrete-event simulation (hereafter referred to as simulation) approach has been applied to decisions in scheduling and control, related to production applications (see e.g., Banks 1998, Law and Kelton 2000, O'Rielly and Lilegdon 1999). The simulation models that are used for making or evaluating these decisions (e.g., by projecting different key performance indicators, KPI-s) generally represent the flow of materials to and from processing machines and the operations of machines themselves (Rabelo et al. 2003). Potential problems can be identified and can be corrected using a simulation model. By far the most common use of simulation models is for operational decisions such as scheduling or dispatching (Law and Kelton 2000).

Simulation captures those relevant aspects of the Production Planning and Control (PPC) problem which cannot be represented in a deterministic, constraint-based optimization model. The most important issues in this respect are uncertain availability of resources, uncertain processing times, uncertain quality of raw materials, and insertion of conditional operations into the technological routings.

In simulation supported KPI evaluation, simulation is often used for evaluating the different scheduling or dispatching logics and methods. The usefulness of simulations lies in detecting and preventing the problems concerning KPI-s before they might occur at the shop floor. Thus, the key benefit of a simulation-based evaluation system is the feedback about system performance and system state (e.g., WIP, tool

utilisation) which, in turn, can be used for improving subsequent solutions.

2.2. Simulation-based schedule and control evaluation

As it is stated above, one stream of research in simulation-supported production control focuses on the simulation-based evaluation of schedules and scheduling rules. In these cases the main goal is to find solutions with the help of the simulation to daily scheduling issues including on-time order completion, priority changes, and unexpected changes in resource availability. Discrete Event Simulation (DES) helps a system engineers in detecting potential scheduling problems through checking the resource and schedule performances during the scheduling interval (shift, day, or week). The new alternative policies are then executed and performances of alternatives are compared. This process is repeated until a feasible and desired schedule is achieved. Indicated in another way, a schedule is created by simply simulating the execution of the factory and taking the recorded execution history as the schedule (Smith 1992).

Previous solutions introducing simulation-supported evaluation of schedules and scheduling policies are given in Cowling and Johansson (2002), Honkomp et al. (1999), Sabuncuoglu and Kizilisik (2003) and Vieira et al. (2000).

Watt (1998) presents a case-study where several information sources and applications are integrated. Simulation is applied in both off-line and scheduling modes. The current plant status and static data from the ManuMES are collected periodically and schedules are generated by a commercial scheduling package. Off-line simulations are performed in order to test what-if scenarios and reuse the same information for scheduling. New rules can be created and tested against history data. The improved rules are then applied in the scheduling system.

In several manufacturing areas the amount of products, product variants, resources, variation and fluctuation in the production processes does not facilitate predefined schedules to be adhered to. In these cases production is mainly controlled by production control rules (e.g., dispatching rules, for order management, and resource allocation, etc.) Thus, another direction of research does not focus on the schedule formulation but on the trajectories of the most relevant KPI-s during the simulation execution of the production. Solutions in the field of semiconductor manufacturing are presented in the followings.

Bagchi et al (2008) describe a discrete event simulator developed for daily prediction of WIP position in an operational wafer fabrication factory to support tactical decision-making. The model parameters are automatically updated using statistical analyses performed on the historical event logs generated by the factory.

Bureau et al. (2007) use simulation to compare alternative WIP management policies, while Sivakumar and Chong (2001) quantify the effects of lot size, lot release controls and machine dispatching rules, on selected KPI-s (throughput, process time and process time spread) for manufacturing steps. Klein and Kalir (2006) introduce a simulation system in order to analyse transient behaviour of a factory in case of new products ramp up and old products ramp down. Lin et al. (2001) and Kim et al (1998) present a simulation study on lot release control and dispatching rules for batch scheduling in semiconductor industry.

As it was shown in the previous literature review, related works describing some of the application areas, as well as the recent solutions of simulation in production control, simulation has been typically used for off-line decision making. Consequently, effective integration into the control process of production was restrained. One of the limitations of its use in on-line decision making is the considerable amount of time spent in gathering and analysing data. In quasi real-time control (hours, minutes), however, the three key issues are data acquisition, quick response and instantaneous feed-back. As a result, decision makers mostly apply simulation primarily for off-line decision support and not for the critical on-line decision making that may arise.

3. SELF-BUILDING SIMULATION SYSTEM

To support factory wide short-, midterm capacity as well as production control (utilization, WIP), scheduling and planning decisions, a simulation system has been developed – in form of a cooperative research project. This simulation system builds the discrete-event, object-oriented simulation model of the example shop-floor system automatically on the base of the data retrieved from the Manufacturing Execution System (MES) database.

This section presents the results of the systems' analysis and the developed self-building simulation tool.

3.1. Novel simulation architecture – self-building simulation tool

At the beginning of the research activity related to this paper, in several preliminary examinations it turned out that the drawbacks of the existing simulation-based dispatching system are as follows:

- The input of the simulation model is oversimplified, collected and/or generated manually, which requires a huge amount of human effort;
- The data provision from shop-floor and engineering generates frequent errors;
- The data maintenance is poor.
- The responses in dynamic manufacturing environments are generally slow;

The main goal of development is the enhancement of the simulation-based analysis and dispatching system by eliminating the manual data collection through

automatic interfaces, creating a more realistic model of the real factory and improving the dispatch logic inside the simulation model. Furthermore the self-building production simulation should provide both, prospective (e.g. locate anticipated disturbances, identify trends of designated performance measures), and retrospective (e.g. gathering statistics on resources) simulation functionalities. *Self-building simulation* means that the simulation model is build up by means of the combination of the MES data as well as the knowledge extracted from the MES data (e.g. resource and execution model). In addition to the automatic model building feature, main requirement of the solution is to minimize the response time of the experiments and to enable the quasi "real-time" applicability of the simulation.

3.1.1. Automated, component-based simulation model generation

In order to meet all the requirements and achieve the desired functionality for a flexible, self-building simulation system, a so-called *component-based simulation* method has been developed (Pfeiffer 2007). Resources, products, routings, production information, i.e., directly and indirectly usable data are gathered from MES database, and transformed as well as processed to the same form for all system components (e.g., for the simulation system, or for the Decision-maker to analyse KPI-s). Note that simulation relevant data (e.g. resource model, execution policies, process flow model) are redundantly stored locally in the simulation model.

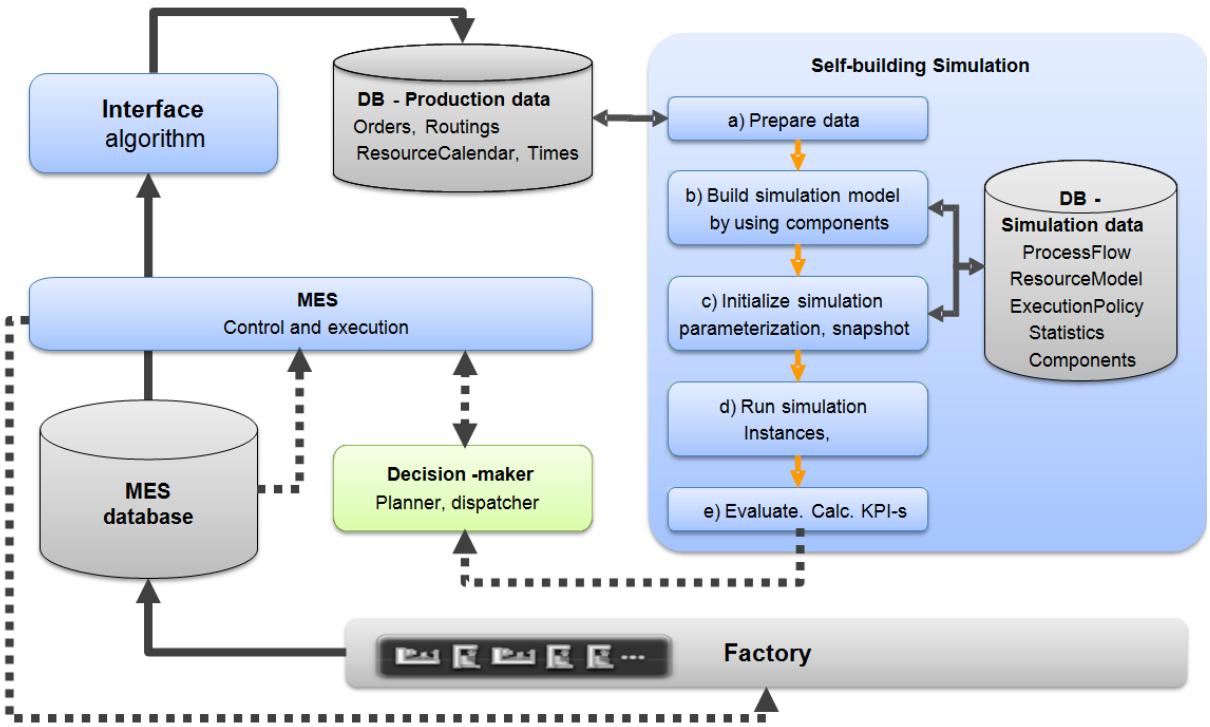


Figure 1: The structure and the data-flow of the new self-building simulation architecture (dotted lines represent information flow, while solid lines represent data flow)

Figure 1 represents the dataflow of the new self-building simulation architecture. All the data necessary for simulation is retrieved from the MES database (represented on left side in Figure 1). The first interface splits the aggregated data of the MES into separated inputs and performs statistical analysis by applying previously developed algorithms. Production data are then further transferred and processed by the second interface into the simulation model (Prepare data in Figure 1). The simulation model keeps its own, internal, simulation-specific database to support fast response time. The exchange of the data is necessary only in the early state of the simulation when the automatic model building and data initialization is performed.

Data preparation is carried out before the overall simulation (production related data is refreshed on a weekly, while MES data for factory snap-shot is refreshed on a daily basis, see Figure 1). The redundant data storage in the simulation model is compensated by the advantage of the shorter response time. Modelling real production systems frequently brings up the problem of handling hundreds of resources in a simulation model. Having the modelling objects in hand, which were created on the base of the conceptual model, in our architecture the simulation model is created automatically based on the pre-processed data (*phase b*).

The automatic generation of the model is followed by initialization (*phase c*). There, besides classical parameter settings, the procedure involves the generation of input parameter specific model components (entities such as products, tools, machines

and the snapshot of the factory, detailed in Section 3.3). Contrary to the previous phase, this one is carried out for each replication. The simulation model incorporates a number of new dispatching rules with which the simulation runs can be manually initialized by simulation experts. In a following version of the simulation the automatic selection of proper dispatching rules and rule combinations will be implemented in order to achieve better performance in the shop-floor. The simulation is started on the base of these statistics by generating random production orders which cover the product type distribution calculated from the MES database. Naturally, instead of randomly generated orders, the users of the simulation can also provide the input for the simulation model on the base of real customer order data.

The simulation runs are repeated until the required number of replications is obtained (*phase d*). Each replication is a terminating, non-transient simulation run.

In the last phase, the results are evaluated (critical values for defined KPI-s) and the results of the evaluation process are interpreted by the Decision-maker (e.g., planner or dispatcher) who is responsible for taking the necessary actions. Several simulation results and statistics are calculated inside the simulation model and a graphical user interface (GUI) is provided for the visualization via a web browser of both, input settings, and statistical results. The first version of this self-building simulation prototype is developed in Plant Simulation v8.1 software, while the GUI is implemented in a web application.

3.1.2. Production information – data in the MES database

The production environment to be modelled by the simulation covers a whole production section of a real industrial factory. The factory produces several different products which are identified by different product types. Each product type is assigned to a routing. The sequence number is the position of a defined operation in the routing. In order to have a better traceability of products moving through the factory a number of operations are logically grouped into so called groups. Regarding resources in the factory we can differentiate between operators and machines. Products are processed on machines with predefined tools and machines are grouped into cells.

Operators are responsible for loading, starting, finishing, as well as unloading the machines. The data in the MES database are collected manually by the operators of the machines during the manufacturing process in the shop-floor. The products are transferred from station to station in lots. When a product arrives to a station the operator of the station registers the arrival and the product is waiting for the manufacturing in the queue of the station. Depending on the type of the tool different products might be processed together.

When an operation is completed the operator of the station registers the completion of the work. As a consequence the MES database contains a large list of data records about the life-cycle of the products cataloguing the entrance and completion events.

The data about the resources, their types and availability, routings, process times are also calculated from the MES database. MES-based statistics are collected about the most important product types and their distribution in the production order.

3.2. Defining operation times on MES data

In simulation systems representing a complex, wide scale production system, exact processing times are crucial for successful and credible simulation results.

Defining process times for simulation models based on logged production (e.g., MES) data is a well known and widely used method. Bagchi et al (2008) present a linear regression method for calculating process times based on raw process times (RPT) collected for single, batch and sampling tool types. Sivakumar and Chong (2001) present a case study, where the theoretical process times are defined by the mean values, however the authors state that based on the wide distribution in theoretical process times, theoretical ratio based on mean is indefinite.

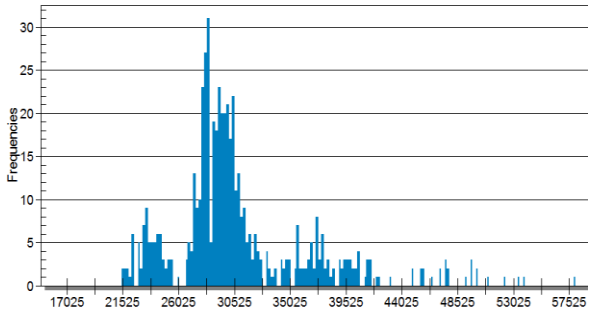


Figure 2: Frequencies of raw process times for a defined operation subtype on batch tool machines in seconds

In Figure 2 an example is given, representing the characteristic of the raw process time distribution. The main goal is to define the so called basic process time for the given operation and tool relations. Since several factor influences the raw process time as for instance different waiting times in the input buffer of tools, different operators, etc., the most relevant way is to scrape the raw process times from the effect of disturbing factors, and thus, to identify the relevant lower bound (excepting unnecessary, problematic data), i.e., the shortest possible raw process time of the data presented in Figure 2. It is clear that by calculating the mean (31193 s) or the mode (29520 s) values for the data set does not provide the necessary lower bound.

Thus, as one of the most important issues regarding the self-building simulation system, a significantly more effective method (compared to e.g., mean or modus statistical methods) had to be developed in order to have reliable and up-to date process times in the simulation. It is also important that the calculated process times will be applied in the near future for static capacity calculations in the factory.

Intensive research is still ongoing in our laboratory to calculate the raw process times. The system presented in this paper applies the preliminary results of the first algorithms. The detailed algorithm will be published soon in an upcoming paper of the authors.

3.3. Simulation model - Creating snapshot of the production

In order to have as much precise simulation results of a daily WIP prediction as possible, the initial state of the simulations (from which the system starts evolving) is critical. Therefore, a snapshot of the production (WIP, status of resources, etc.) is necessary (Figure 3).

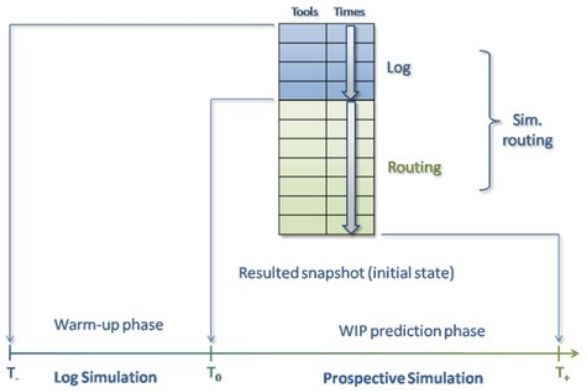


Figure 3: Proposed architecture of the simulated execution of a MES section, the formulation of the WIP snapshot and the prospective simulation

In Figure 3 the blue box represents the operations executed during the Log Simulation, while green boxes represent the steps to be processed in the routing. The main challenge is the transition phase at time point T_0 , where the execution of the log-based operations should be changed to routing-based operations for each unit to be processed.

We propose the following simulation procedure to ensure a short warm-up phase and a reasonable initial state of the production status.

- The user should define the simulation time horizon
  - Starting time (T_0) at which the prediction starts
  - Length of the WIP prediction phase (e.g. 1 week), i.e., $T_+ - T_0$
- Estimation is given on the length of the play-back period, required for the snapshot formulation (T_0 - T_-), in order to have the actual WIP state of the factory (*warm-up phase*)
  - Define the level of significance and the number of simulation runs required
- Play-back of the log (*Log Simulation*)
- From the snapshot point in time (T0) execute the WIP prediction phase (several independent simulation runs, *Prospective Simulation*)
- Evaluation of the results, notify user

In the simulation model we do not simulate the movement of individual products, but the movement and processes done on the lots. Therefore, the selected section (defined by T_0 - T_-) of the log (for creating the log file in the simulation) is decomposed and forwarded to lot objects in an aggregated way. This procedure improves the speed of the simulation, while this lower level of model granularity does not significantly influences the output quality.

4. PRELIMINARY SIMULATION RESULTS

The main goal of the simulation studies are the prospective analysis of the WIP movements in the factory, as well as the monitoring and reporting on bottle-neck machines and machine groups (for capacity analysis).

Our solution is distinct in the sense that the extent of modelling is relatively large (whole factory) while the data available for model formulation is limited only to the – continuously updated and statistically analysed – MES data (no conventional process and resource data is available).

The production system to be modelled has the following main features:

- several million rows in the MES database, 20-50.000 new rows every day;
- 3000-4000 products for snapshot (Factory WIP);
- approx. 500 resources;
- 500-1200 operations in a routing for a defined product type;
- avg. lead time 20-40 days;
- ~ 10-15% rework process.

The section presents an example of the validation process of the simulation model (Section 4.1), and preliminary results of the simulation experiments. Section 4.2 introduces the steady state analysis of system performance, while in Section 4.3, results on testing improved execution and sorting logics are described.

4.1. Validation – comparison of throughput

In order to have credible results computed by the simulation model, a comprehensive validation procedure is mandatory. In the followings a short example is given about the validation of throughput values.

One stream of the validation procedure of the proposed system is based on the comparison of the simulated WIP prediction results (simulated log) with the real original data (Figure 4). This validation procedure serves as a feed-back for the iterative fine-tuning (trace technique, Law and Kelton (2000)) of the execution policies, tool models, process times and dispatching rules applied in the simulation.

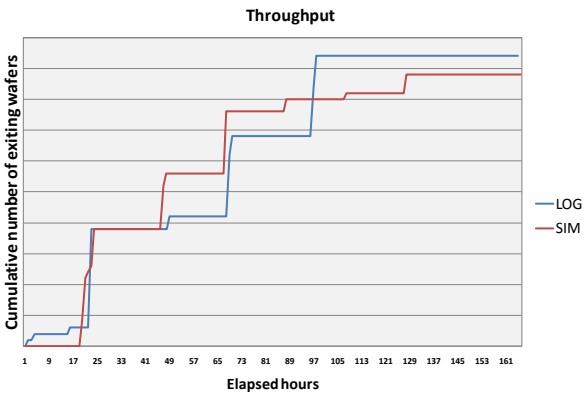


Figure 4: Simulated and real throughput diagram

In Figure 4 the results of a comparison for one week period is highlighted. If the total number of products in the system (WIP) is considered as 100%, then the ratio of exiting products all together in the simulation for the one week time horizon in consideration is 1.222%, while the real, log based ratio is 1.306%. The difference between the predicted and the real total throughput is relatively low (6.38%), however the distinct steps in the log based curve can not be represented exactly with the simulation.

4.2. Testing production control rules - steady state performance

Regarding the shop-floor level control of the production, several control rules are implemented and used in the factory. One of the most relevant groups of these control logics is the work-load balancing dispatching rules of the machines. These rules define the logic of choosing the successor machine in case of the next operation can be processed on more than one (alternative) machine. At the current phase of implementation the following rules are available for the simulation:

- *Random*: the simplest logic to send products randomly to the next tool.
- *Min queue*: this logic prefers the machine where the queue is the minimum.
- *Relative occupation (Tool)*: this logic gives priority for the machine that has the Tool with the lowest relative occupation.
- *Relative occupation (Buffer)*: similar the one above, but considers the Buffer.
- *Next free (Tool)*: this logic prefers the machine where the Tool is free. It chooses randomly if there is no free Tool among the alternative machines.
- *Next free (Buffer)*: similar the one above, but considers the Buffer.
- *Min wait time*: in this logic a method calculates the time remaining to the start time of the process for the product on every possible machine. The product is sent to the Tool where this time is the minimum.

In order to be able to analyse the effect of dispatching rules on bottle-neck machines, first the long-term, steady-state analysis of the simulation model was performed (Table 1). It can be stated that the different dispatching rules significantly affect the systems' throughput performance, e.g. the rule *Min wait time*).

Table 1: The average number of finished products (simulation time horizon is 50 days, number of replications is 100)

Dispatching rule	**No. of finished products**
Next free (Tool)	2619
Relative occupation (Tool)	2188
Relative occupation (Buffer)	2434
Min wait time	2865

As an important output of the self-building simulation system, situation-dependent selection of the most relevant control rules are intended to be tested and analysed. Therefore, contrary to the above described analysis, not only the steady-state performance of the defined control rules is of interest, but the short-term effect regarding the WIP in the input buffers of the machines. Moreover, the dynamic behaviour, e.g., the movement of these WIP "waves", from machine to machine has to be considered.

In Figure 5 and Figure 6 the predicted trajectories of the WIP are presented for the selected bottle-neck machines, simulated for the dispatching rules *Min wait time* and *Random*, respectively. Based on the simulation results it can be stated that, the control rules with a reasonable high performance for steady-state studies, not always perform similarly good regarding short-term decisions (e.g., for eliminating WIP in buffers).

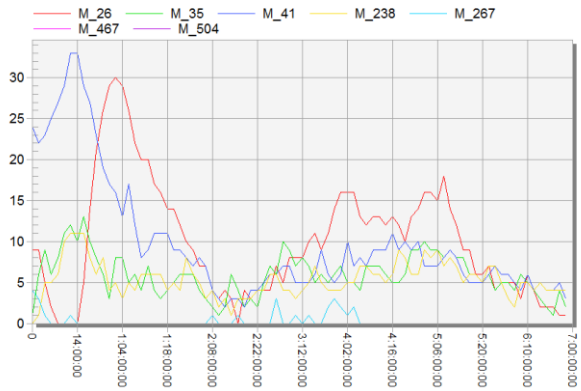


Figure 5: Simulated WIP movement diagram, for selected bottle neck machines (number of products in the input Buffers/ time, for a one week horizon, applied dispatching logic: *Min wait time*)

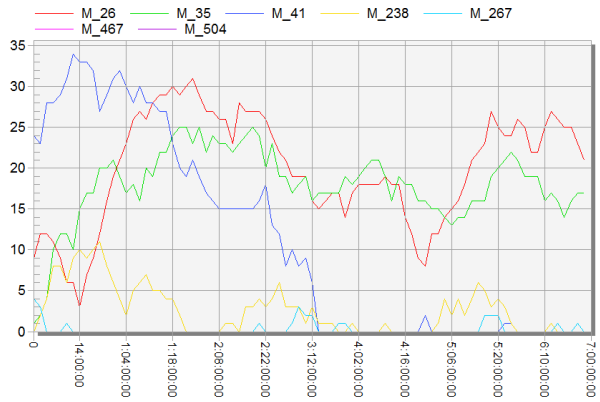


Figure 6: Simulated WIP movement diagram, for selected bottle neck machines (number of products in the input Buffers/ time, for a one week horizon, applied dispatching logic: Random)

4.3. Suggestions for improving utilisation of parallel machines

During the analysis of the real systems' data in the MES database, it turned out that the actual load of some machines, and thus the utilization is constantly low.

These machines are of principally one type, namely, special parallel machines, characterised by a certain maximal load capacity and with high operation costs. In case of the processing of the products starts, no more products can be loaded into the machine unless the processing is finished. Thus, the main goal is to minimize the idle time of these machines and to maximize the utilization (load). Therefore, three different loading logics have been developed and compared through simulation experiments. The three logics are as follows:

- *FIFO* (or *no group*) means the simplest logic: product lots are processed based on the FIFO rule;
- *Group* logic regroups the product lots, based on the operation type (OpType) required, and then it sorts according to the number of products in the new group. The order of processing is FIFO;
- *Group and sort* logic first regroups the product lots as described in *Group*. Each group (based on the OpType) has a predefined "starting level". It defines a loading level at which processing is allowed to be started. A *waiting time* is also assigned to each OpType The group which first exceeds the *starting level* or exceeds the *waiting time* will be first processed on the machine.

The upper chart in Figure 7 presents the case, where *No group (FIFO)* rule is applied to load the parallel machines, while the lower chart presents the case, where the special sorting logic (*Group and sort*) is applied in order to have a higher machine utilization ratio. In the first case, the machine does not wait for *OpType1* operations to load the machine up to full capacity, but starts processing immediately after receiving products for *OpType1,* or *OpType2*. Contrary, in the second case, the processing of products requiring *OpType1* are postponed, because the minimum level assigned to this type of operation is not exceeded.

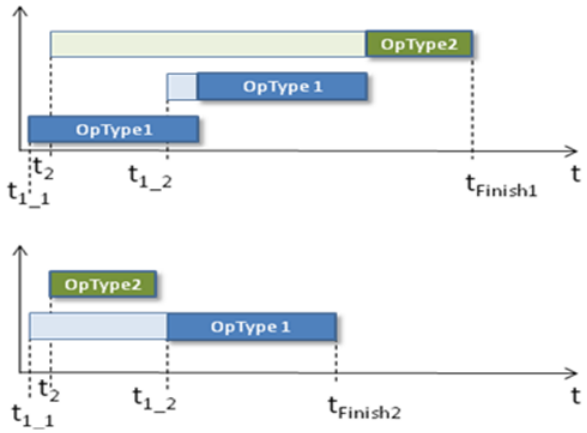


Figure 7: The effect of *no group* logic (upper chart) and *group and sort* logic (lower chart) on lead time, for loading a parallel machine

The results of simulation experiments regarding the effectiveness to the new sorting logics are presented in Figure 8 and Figure 9 for two selected different machines. It is clear that the new *Sort and group* algorithm has a positive effect on the WIP level of machine input buffers.

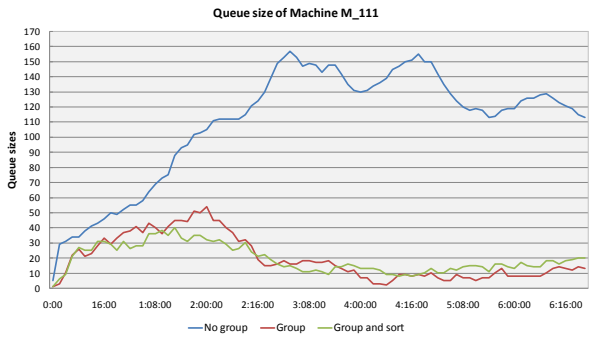


Figure 8: The WIP of machine M_111 for one week, with the three different sorting logics

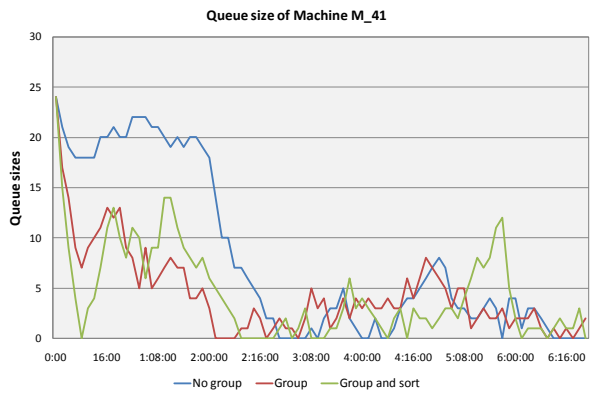


Figure 9: The WIP of machine M_41 for one week, with the three different sorting logics

5. SUMMARY AND FURTHER STEPS

The features provided by the new generation of simulation software facilitate the integration of the simulation models with the production scheduling and control systems. Additionally, if the simulation system is combined with the production database of the factory, it is possible to instantly update the parameters in the model and use the simulation parallel to the real manufacturing system supporting and/or reinforcing the decisions on the shop-floor.

However, some tasks are still under development stage (e.g. calculation of reliable process times), in this paper the authors presented a novel solution to build simulation models from the MES database. Moreover, the initial validations, which were carried out with real industrial data, showed that the behaviour and the results of the simulation model are close to the ones of the real system.

As a further step, regarding the usage of our self-building simulation approach, we propose the re-initialization and memorization of sequential simulation runs in a rolling time intervals. In case of high difference arises between the simulated and the real production KPI-s (WIP, waiting times, etc.), an examination is initiated in the log of the last interval to identify the root cause of the deviation.

The main challenges related to this rolling time horizon analysis are the identification of relevant rolling interval (i.e., the frequency of simulation experiments), defining significant ΔKPI-s and the definition of look-ahead horizon.

With predefined simulation protocols, this method will result in an easy-to-use decision support tool for shop floor engineers, which gives prediction of future events at a time by which necessary actions can be taken in advance.

ACKNOWLEDGMENTS

This research has been supported by the Hungarian Scientific Research Fund (OTKA) grants "New mathematical models and methods for integrated production planning and scheduling" K-76810, and "Production Structures as Complex Adaptive Systems" T-73376.

REFERENCES

Bagchi, S., Chen-Ritzo, C., Shikalgar, S.T., Toner, M., 2008. A full-factory simulator as a daily decision-support tool for 300mm wafer fabrication productivity, *in Proceedings of the 2008 Winter Simulation Conference, S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler eds.*, pp. 2021-2029.

Banks, J., 1998. Handbook of Simulation, Principles, Methodology, Advances, Application and Practice. *John Wiley & Sons Inc*.

Cowling, P., Johansson, M., 2002. Using real time information for effective dynamic scheduling, *European Journal of Operational Research*, 139, 230-244.

Honkomp, S.J., Mockus, L., Reklaitis, G.V., 1999. A framework for schedule evaluation with processing uncertainty, *Computers and Chemical Engineering*, 23, 595–609.

Hyndman, R.J., 1995. The problem with Sturges' rule for constructing histograms, *Department of Econometrics and Business Statistics, Monash University, Clayton, Victoria, Australia*, pp. 1-2.

Kim, Y., Lee, D., Kim, J., Roh, H., 1998. A simulation study on lot release control, mask scheduling, batch scheduling in semiconductor wafer fab facilities, *Journal of Manufacturing Systems*, 17 (2), 107–117.

Klein, M., Kalir, A., 2006. Improved simple simulation models for semiconductor wafer factories. *In Proceedings of the 2006 Winter Simulation Conference, ed. L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto*, pp. 1836-1839.

Law, A., Kelton, D., 2000. Simulation modeling and analysis, *McGraw-Hill, New York*.

Lin, J.T., Wang, F., Yen, P., 2001. Simulation analysis of dispatching rules for an automated interbay material handling system in wafer fab, *International Journal of Production Research*, 39 (6), 1221–1238.

Monostori, L., Kádár, B., Pfeiffer, A., Karnok, D., 2007. Solution approaches. to real-time control of customized mass production, *CIRP Annals - Manufacturing Technology*, 56(1), 431-434.

O'Reilly, J.J., Lilegdon, W.R., 1999. Introduction to FACTOR/AIM. *In: Proc. of the 1999 Winter Simulation Conference,* pp. 201-207.

Pfeiffer, A., 2007. *Novel Methods for Decision Support in Production Planning and Control*, Thesis (PhD). Budapest University of Technology and Economics.

Rabelo, L., Helal, M., Jones, A., Min, J., Son, Y.J., Deshmukh, A., 2003. A hybrid approach to manufacturing enterprise simulation. *In: Proc. of the 2003 Winter Simulation Conference*, pp. 1125-1133.

Sabuncuoglu, I., Kizilisik, O.M., 2003. Reactive scheduling in a dynamic and stochastic FMS environment, *International Journal of Production Research*, 41(17), 4211-4231.

Sivakumar, A.I., Chong, C.S., 2001. A simulation based analysis of cycle time distribution, and throughput in semiconductor backend manufacturing, *Computers in Industry*, 45 (1), 59-78.

Smith, S.F., 1992. Knowledge-based production management: Approaches, results and prospects. *Production Planning & Control*, 3(4), 350-380.

Vieira, G.E., Herrmann, J.W., Lin, E., 2000. Predicting the performance of rescheduling strategies for parallel machines systems, *Journal of Manufacturing Systems*, 19, 256-266.

Watt, D.G., 1998. Integrating simulation-based scheduling with MES in a semi conductor fab, *in*

Proc. 30th WSC 1998, ed. Benjamin et al., pp. 1713-1715.

Wilhelm, R., et al., 2000. Finite Schedule Monitoring and Filtering in a Computer Integrated Manufacturing Environment, *Annals of the CIRP*, 49(1), 335-338.

AUTHORS BIOGRAPHY

András Pfeiffer, earned his PhD in 2008 at the Budapest University of Technology and Economics. Currently he is a senior research fellow at the Engineering and Management Intelligence Laboratory of 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 modelling of complex production systems, self-building simulation systems.

Botond Kádár is a senior researcher at the Engineering and Management Intelligence Laboratory of the Computer and Automation Institute of the Hungarian Academy of Sciences (SZTAKI). He obtained his MSc and Ph.D. degrees at the Budapest University of Technology and Economics, Hungary, in 1993 and 2002, respectively. His current interest includes production control, simulation and multiagent approaches for production engineering and manufacturing systems and he is involved in several research and development projects from these fields. Dr. Botond Kádár is author or co-author of 70 publications with over 120 citations.

Marcell Szathmári, graduated in 2006 at the Budapest University of Technology and Economics. Currently he is a research associate at the Engineering and Management Intelligence Laboratory of the Computer and Automation Research Institute, Hungarian Academy of Sciences (SZTAKI). His current interest includes tracking information systems and automatic identification technologies.

Gergely Popovics, graduated in 2008 at the Budapest University of Technology and Economics. Currently he is a research associate at the Engineering and Management Intelligence Laboratory of the Computer and Automation Research Institute, Hungarian Academy of Sciences (SZTAKI). His current interest includes the simulation modelling of complex production systems and automatic identification technologies.

Zoltán Vén, graduated in 2006 at the Budapest University of Technology and Economics. Currently he is a research associate at the Engineering and Management Intelligence Laboratory of the Computer and Automation Research Institute, Hungarian Academy of Sciences (SZTAKI). His current interest includes reconfigurable manufacturing systems, as well as the simulation and emulation modelling of complex production systems, self-building simulation systems.

Prof. László Monostori has been with the Computer and Automation Institute of the Hungarian Academy of Sciences (SZTAKI) since 1977, now he serves as Deputy Director Research and Head of the Engineering and Management Intelligence Laboratory. He is also full time professor at the Department of Manufacturing Science and Technology, Budapest University of Technology and Economics. He is a Fellow and Council Member of the International Academy for Production Engineering (CIRP); Full Member of the European Academy of Industrial Management (AIM); Chairman of the Coordinating (CC) on Manufacturing and Logistics Systems, International Federation of Automatic Control (IFAC), and Chairman of the TC on Technical Diagnostics, International Measurement Confederation (IMEKO). He is Chairman of the Editorial Committee of CIRP Annals; Editor-in-Chief of the CIRP Journal of Manufacturing Science and Technology; Associate Editor of Computers in Industry, as well as the Measurement, and member of the editorial boards of other international scientific periodicals. For his research achievements published in more than 340 publications resulted in about 1500 independent citations and for his development activities – among others – the Dennis Gabor Prize was given to him in 2004. Prof. Monostori is a member of the Hungarian Academy of Engineering.

HOW TO SUPPORT FIRM ENVIRONMENTAL INNOVATION: THE CASE OF THE VOCLESS MECHANICAL PULPING PROJECT

Stefano SAETTA, Lorenzo TIACCI

Dipartimento di Ingegneria Industriale, Università degli Studi di Perugia

stefano.saetta@unipg.it

ABSTRACT

One of the most discussed issue is the environmental sustainability. The increase depletion of natural sources, the high level of pollution, also in emergent countries, will push the economic world towards more environmental friendly decisions and actions. Eventually, the environment dimension will play an important role in the decision making in future. But the environment, as the sustainability in general, must face, today, other dimensions such as economic and short term competitiveness. Not always it is possible to show the advantage, especially in the long term of an environment oriented investment. It is necessary to develop an environmental solution in order to show its feasibility. It is important so, to study how to effectively integrate environmental issues with company strategy making the right trade-offs between strategic and environmental goals. It is important to focus on procedure by which it is possible to show the economic viability of environmental investment. In the paper an interesting case is discussed: the VOCless project. This project thanks to the support of the EU within Life program which has allowed to demonstrate that solutions like regenerative catalytic incinerator and moreover biofiltering are effective in VOC emission control of Mechanical Pulping processes. Results are shown and the feasibility and transferability to other sector are discussed.

Keywords: VOC abatement, Biofiltering, catalytic incineration, Environamental Management

1. INTRODUCTION

Nowadays there are still strong barriers to implement a real environmental friendly policy. This for many reasons: there is not a global support to environment problems, markets do not always gives value to green policies, there is a tight worldwide competition.

Nevertheless some improvement in the green dimension is starting. EU and US policy in favor of renewable energies is one example. Market promotion of environmental friendly products is another. But one of the barrier further environmental industry improvement is due to the risk of environmental investment. Sometimes environmental investments are just a cost, which do not improve any impacts. This is also not good in the case companies are acting proactively towards the environment and they are looking towards improvements below legislation limits. So, even if there are theoretically possible improvements with traditional and new technologies, very often companies does not really implement anyone of them. So limits to the investments are not only financial but also because there are not certain environmental returns. In this sense it is important to be able to activate some industrial research and/or pre-competitive development in order to show the feasibility of one investment. In this manner it is possible to give evidence about achievable results and support better environmental decision making. In the following of the paper this issue is discussed referring to a particular case, the "VOCless Pulping", VOCless, project, funded by European Union under LIFE-Environment Demonstration projects 2006 (website http://voclesspulping.com/front-page). Thanks to VOCless it was possible to demonstrate the feasibility of VOC abatement in the Mechanical Pulping industry.

Mechanical Pulping companies were interested in the VOC abatement, even if it is not compulsory by the EU law. The problem has been about the effectiveness of VOC abatement solutions. There was not any study concerning the most viable techniques or hybrid combination of them. In VOCless project it was possible to investigate through measurements and literature survey several techniques and give the right suggestions to the companies. In particular one traditional technique, such as regenerative catalytic incinerator and one more innovative technique, a new biofiltering were tested and compared. VOCless project is an example of how to proceed in order to improve the environmental performances of the companies.

Sometimes the barriers to the innovation in the companies rely on the fact that there are not possibilities to show the feasibility of a solution. The support of research and demonstrative actions could give good results in many fields of industrial engineering.

In the following, after a brief related research, a description of the VOCless project is made, a description of the Mechanical Pulping industry is made, with VOC emissions. Afterwards a description of regenerative catalytic incineration and biofiltering is made and an evaluation, based on the pilot plant test

results was made. In the conclusions final considerations and future development are discussed.

2. RELATED RESEARCH

There has been a growing interest about the environmental management strategies since at least the last 2 decades, as in (Graedel and Allenby 1995; Hart 1997; Porter 1990; Russo and Fouts 1997). The Strategic Environmental Management links to the neoclassical Economy theory and Capabilities Based theory was discussed (Goldstein 2002). Governmental decisions, starting from EU and Japan, support the Sustainable Development by promoting Environmental Industries and Management (Elder 2007). This is in order to stimulate innovation and increase efficiencies. But the needs of economic growth of the developing countries can limit the environmental management in multinational corporations (Chaiporn 2009). But the problem on how to quantify the benefits of a good environmental management is still under debate. (Bhat 1999) thinks that it seems that green management means to make more profit, also because there is a low risk in a green company. But (King 2001) it has been shown that even if statistically seems that green firms make more profits, it can not be stated that this is due to environmental management. In synthesis (Lankovsky 2008), Environmental Management belongs to Corporate Responsibility activities. For those activities there are three possible outputs: learning, reputation and Corporate Responsibility Outputs. The three outputs must be taken into account not only for sustainability but also from the economic performance.

It is not easy for companies to evaluate environmental performances, even if it seems that there are benefits. It became very important to be able to effectively invest in environmental actions. In this sense innovative solutions are particularly risky. Firms that have in any case difficulties in evaluating the economic benefits of environmental management, can not stand under performing investments. It is important to find a way to support such industrial innovations. In this manner it will possible to take into account the risk of the innovation. This is relevant in any environmental investment. In Modeling and Simulation, M&S, environment is becoming a very important component. We are moving from classical problems, where the aim of M&S were in essence to improve production perfomances (Guash and Piera 2005), towards new models (Longo and Mirabelli 2008), based also on other aspects, such as security (Longo and Massei 2008) and, of course the green dimension (Bruzzone and Merkuriev 2009), (Bruzzone, Tremori, Massei and Tarone 2009).

The case considered in the paper concern a VOC abatement study for Mechanical Pulping in Europe. VOC emissions, a very wide spectrum of organic compounds and are naturally generated or are emitted especially from industries using traditional solvents (Seinfeld and Spyros 1998;EU 1999; Mannschereck Mannschereck, Bächmann, Becker, Kurtenbach, Memmesheimer, Mohnen, Obermeier, Poppe, Steinbrecher, Schmitz, Volz-Thomas and Zabel 2002; Schnitzler, Bauknecht, Brüggemann, Einig, Forkel, Hampp, Heiden, Heizmann, Hoffmann, Holzke, Jaeger, Klauer, Komenda, Koppmann, Kreuzwieser, Mayer, Rennenberg, Smiatek, Steinbrecher, Wildt and Zimmer 2002).

Their impact depend on the particular organic compound but generally all VOC emissions contribute to the photochemical pollution. Governmental legislation concern essentially limits on the anthropogenic VOC emissions like solvent based paintings, but not other emissions concerning, for instance natural wood or some food and beverage industries. This because those VOC emissions are biogenic, generated from nature. Biogenic emissions are considered mainly for other issue and mainly odour abatement (Revah and Morgan-Sagastume 2005) (Kleinheinz and Wright 2009). Mechanical Pulping belongs to biogenic emissions. Mechanical Pulping is a studied subject, but, since very often pulping production is integrated with paper production, there are studies that considers the environmental impact of pulp and paper all together. Mechanical Pulping environmental studies concerns mainly the wastewater effluent (Zuo, Qun 2009), (O'Connor, Kovacs and Voss 1992), (Johnsen, Grotell and Tana 2000). Another aspect of Mechanical Pulping considered is the energy consumption (Hart, Waite, Thibault, Tomashek, Rousseau, Hill, Sabourin 2009; Rausdepp 2009). There are not relevant studies concerning VOC abatement and Mechanical Pulping, excepted (Sueiro and Gill 1995). In this paper VOC emissions are analyzed in the case of Thermomechanical Pulp. In particular the inventory of VOC emissions is considered. The refining stage and chip washing contribuite to VOC emission rate. The are not any measurement about any control techniques for the abatement.

There several abatement techniques that could be used for the VOC abatement (Moretti and Mukhopadhyay 1993) (Moretti 2002), but there are not any study concerning VOC emissions in Mechanical Plant, event if this emissions are relevant. In VOCless project the aim has been to study which are the best VOC abatement techniques for Mechanical Pulping processes.

In particular two technologies were tested. First one is a regenerative catalytic incineration, that is common in pulp and paper industries. The other was a biofiltering, that is new for Mechanical Pulping. Biofiltering is one of the more promising techniques for its sustainability in term of environmental protection and also safety. For a success in biofiltering is important to keep the good operating conditions that allows bacteria eliminate VOC from the inflow air. More on biofiltering can be found in (Adler 2001; Boswell 2002; Deshusses, Shareefdeen 2005; Kim 2004).

The two abatement techniques, one considering catalytic incineration and the other considering

biofiltering, were tested and compared on technical, environmental and economic point of view.

3. INDUSTRIAL INNOVATION SUPPORT: THE CASE OF LIFE-VOCLESS PROJECT

3.1. Followed Procedure

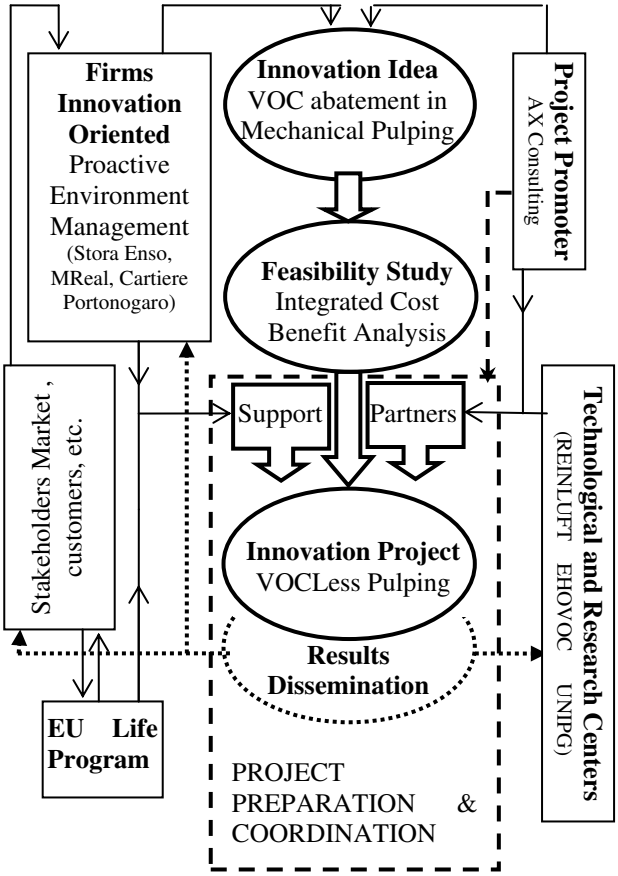


Figure 1: VOCless procedure

In Figure 1 there is a Flux diagram about the VOC less project. This represents the procedure followed to support an innovation initiative. The input and innovation idea in the case considered came out from two actors: firms and the project promoter.

Firms considered were very oriented to innovate and in environment protection. The proactive behaviour towards environment impact is due also to the action of the various stakeholders. Stakeholders, from firm owners, to government, to citizens are driving companies to anticipate environmental issues in the normal management. Main goals of firms must be economic, of course, but environment has became one of the most important driver in the global competition. It is necessary to understand as soon as possible any potential environmental impact in order to be ready before competitors and using the environment as a competition driver. In the case of VOCless project there were leading companies in pulping and in paper industry such as Stora Enso and M-Real. Also a small but dynamic company from Italy joined the project for the piloting test, the Cartiere di Portonogaro. In order to start the innovation process is then necessary to find an activator a promoter. AX Consulting, from Tampere, Finland. AX is a consulting company with experience in environmental impact and also in VOC emissions and European Research Projects. The Project Promoter in AX Consulting understood that there was the possibility to treat VOC abatement in Mechanical Pulping, even if this was not necessary, because Mechanical Pulping does not belong to any VOC abatement directive.

VOC abatement could be a future issue and could create benefits in terms for instance of odour abatement and safety. The problem was mainly to investigate what could be the better technology for the abatement. One well known technology in the paper industry is incineration, but no data about performance in Mechanical pulping VOC emissions were available. This is a risk for companies because there are not any pratical usable references. Moreover, there are other technologies and in particular biofiltering will be interesting both for CO2 emission reduction respect to catalytic incineration and also for safety issue, since there is not any risk of fire. But for firms was not possible to be confident on any solution, due to the lack of any application.

The basic idea of the Project Promoter was to activate a research project in order to create a reference for the application of VOC abatement in mechanical pulping. Project Promoter started preparing the project, by finding supporter and partners. In the case of partner there were both technological partner, as Ehovoc and Reinluft and also research center such as UNIPG, the Department of industrial Engineering of the University of Perugia. Support to the project came from the European Union, via the Life Environmental Program. In this manner has been possible, as shown in the following to activate the research and to have interesting results concerning VOC abatement in Mechanical Pulping. Results and outcome diffusion and dissemination will contribute to give the right information to the interested parties. In this manner it will be possible, for firms, for technology developers and for any stakeholder to take consciences Environmental decisions.

3.2. VOCless Project

The aim of the project is to minimise emissions of volatile organic compounds (VOC) from mechanical and semi-chemical pulping processes due to their contribution to smog formation and odour problems. The emissions minimization are obtained via a VOC and odour abatement system that takes into account the particular conditions of the pulping process. The abatement system is optimized by considering several abatement technologies and combination of them. In the working team of the VOCless project there are partners with experience in pulp industry and VOC emissions that play different roles. In particular there are: 1 university research center, 2 pulp and paper companies, and 3 companies operating in the environment protection.

Partners

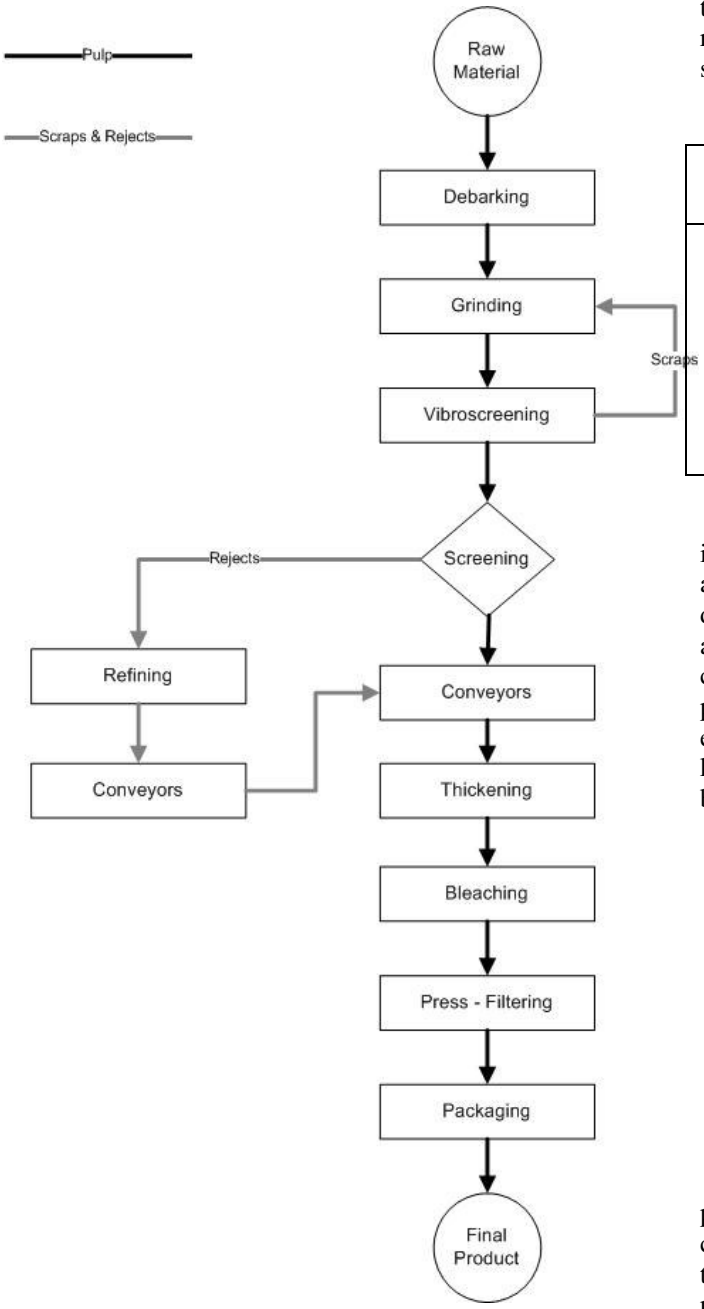


Figure 2: Process Flow of one firm involved in VOCless project

VOCless project includes several activities. First of all a Taxonomy of mechanical, Thermomechanical and chemi-thermomechanical pulping processes, with its VOC emissions.

In Table 1 is shown a classification of Mechanical Pulping. It is a well known processes but with variations due to the kind of wood used.

Starting from the taxonomy of Mechanical pulping, VOC emissions were considered and literature information was searched and also measurements were made in order to have whole information about the typology and the quantities of emissions. In Table 2 the results from measurements in 2 plants are synthetically shown.

Table 1: Mechanical Pulping processes classification

Grinding processes	**Refiner processes**
Stone groundwood pulping (SGW) Pressure groundwood pulping (PGW)	Thermomechanical pulping (TMP) Chemi-thermomechanical Pulping (CTMP) BCTMP if bleaching is integrated in CTMP

The next step in the VOCless project is the investigation about performances and limitis of VOC abatement techniques. In Figure 3 a classification of different abatement technique is shown. VOC abatement techniques may differ in many aspects, considering costs, environment and economic performance and also characteristics for VOC air emissions. Not all techniques are then usable for one kind of emission. Also it is possible to optimize results by a combination of 2 or more techniques.

Table 2: VOC from one PGW plant and TMP plant.

		PGW plant		TMP plant	
Variable	Unit	Min	Max	Min	Max
VOC	mg/ Nm^3	26	226	11	1100
Airflows	Nm^3/s	0,2	20	0,26	1,7
Temper.	°C	38	74	39	85
Relative humidity	%	30	99	83	100

Results of the measurement are used to design a pilot plant, considering 2 abatement techniques: catalytic incineration and biofiltering. Enviromental and technical results are considered in order to figure out potential benefit of VOC abatement techniques.

3.3. VOCless Project Results

VOC measurements were carried out in 3 different plant. In pulp production was integrated with paper production also the emission from drying pulp were considered in order to asses its magnitude. Sample from wastewater were considered to evaluate the VOC transfer to water. There are several sources of VOC emissions. Debarking was considered, while emissions from sawing processes before Pulping were not included.

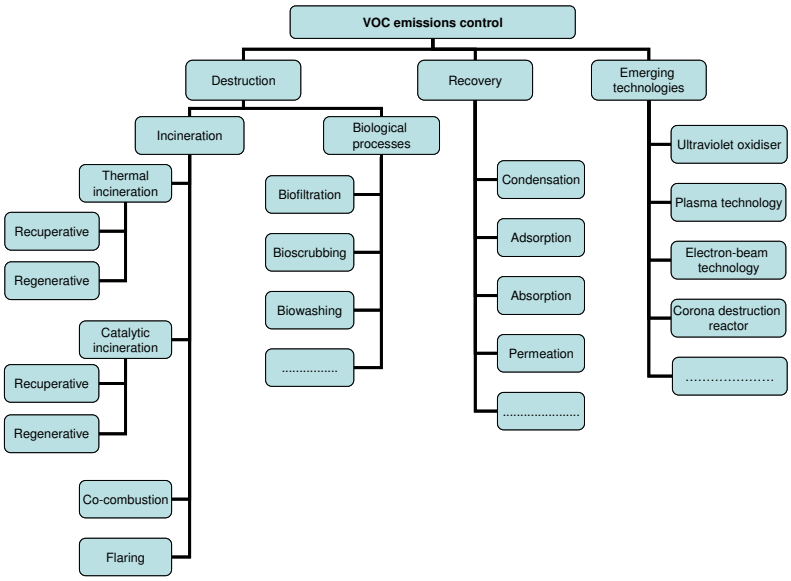


Figure 3 Different VOC abatement techniques

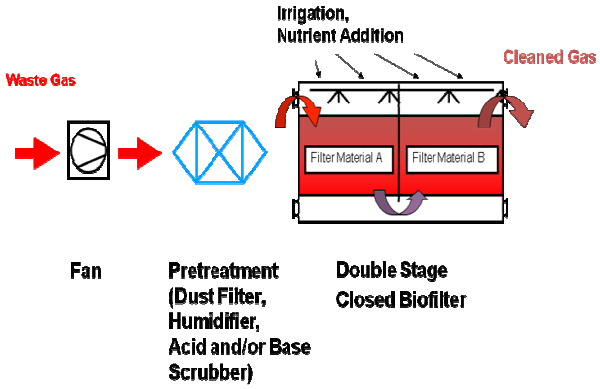


Figure 4 A schema of Biofilter for the VOC abatement

Figure 5 A real installation of a modern biofilter [REINLUFT]

From emissions it was possible to note that:

- Numerous exhausts in the pulping processes that complicated measurements.
- Total VOC emissions per production site might be significant depending on the size of the facility.
- The VOC emissions to air from the pulping process normally consist of a group of terpenes and in some cases also methanol and ethanol. There are always also other unidentified components present.
- Fugitive sources of emissions, in which the concentrations are low but from which the total emissions might be remarkable (e. g. debarking and waste water treatment)

Other relevant results concern the piloting testing of the plant. Total results are not yet available, because testing is still going on. But some consideration can be made from the first measurements. In Particular there were 2 pilot testing measurements, one catalytic incinerator and one biofilter.

The cleaning and thermal efficiencies of the catalytic incinerator were measured before and after the incinerator.

In the PGW plant the pilot plant cleaning efficiency was an average 90% and in the TMP plant the pilot catalytic incinerator reached the cleaning efficiency of 82-94%. The "total" VOC concentration of the combined exhausts (before catalytic incinerator) in the TPM plant was measured to be 260 mg/m^3 and the total airflow was 4,0 m^3/s. The "total" VOC concentration of the combined exhausts (before catalytic incinerator) in the PGW plant was measured to be 460 mg/m^3 and the total airflow was 46,5 m^3/s.

The VOC concentrations after catalytic incinerator were up to 20 mg/m^3 both in TMP and PGW pilot plants. This cleaning performance guarantees the emission value of 50 mg/m^3 which is the normal requirement of VOC emission level according to the Solvent Directive (1999/13/EC) and which can be set as the target level for the pulping process exhausts in VOCless pulping project.

The thermal efficiency of the pilot catalytic incinerator was 90% at TMP plant. The thermal efficiency depends usually on the optimization. The higher efficiency is needed, the higher investment cost is required. The higher the thermal efficiency is, the less heat energy is needed in the oxidation process. The catalytic incinerator is working at reasonably low temperature level (300 °C) and thus does not consume as much energy as thermal incinerators (operating temperature usually 850 °C).

Results of the measurements demonstrate efficient operation and good cleaning efficiency of the catalytic incinerator pilot plant at TMP and PGW plants (Table 3).

Table 2: Cleaning efficiency of regenerative catalytic incinerator pilot plant

Pilot plant	Cleaning efficiency (%)	Thermal efficiency (%)
TMP	82-94	88
PGW	94	94

From a first estimation, not considering, in particular differences in energy costs that can affect the global performance, the total cost of abatement range from 0,4 to 1 € per ton pulp.

A pilot biofiltering plant was also considered. A schema where there are the sampling points is in Figure 5:

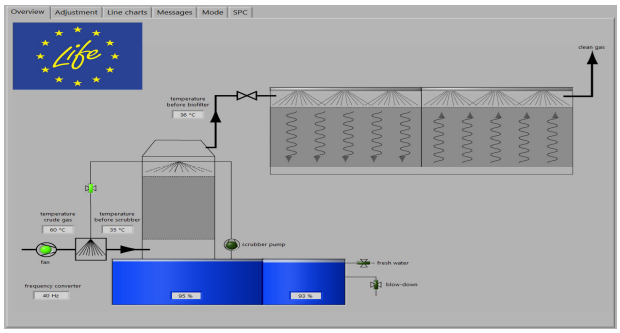


Figure 6: Sampling point in the biofilter pilot plant

Generally speaking biofiltering can give very good results, but it requires a right working conditions. This is not always easy to do. But in the case tested, probably because of concentration of VOC, results are very good. In Figure 7 is shown the FID results in terms VOC abatement in the sampling points. The efficiency is very high. Also in Figure 8 it is shown the good efficiency in VOC abatement of pilot biofiltering plant. Even if the evaluation is not complete, biofiltering should have low costs. In addition biofiltering is a good solution concerning CO_2 consumptions. It does not need fuel to work and energy consumption are quite small.

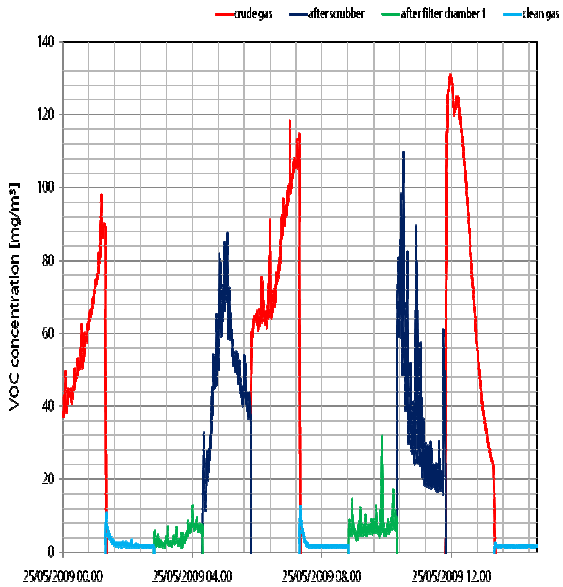


Figure 7: FID output with VOC concentration showing a good abatement

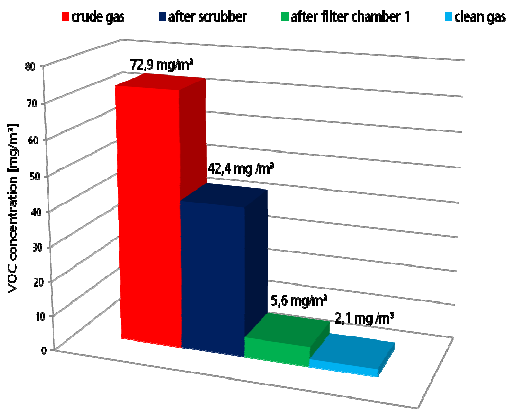


Figure 8: VOC concentration abatement with Biofilter Pilot Plant

4. CONCLUSIONS

In the paper a case of how to support innovative solution in the case of industrial engineering is shown. In particular the possibility to reduce VOC emissions in Mechanical Pulping is considered. This idea needs support because is a proactive action and there were not enough results from past experience. The evaluation of several techniques show the technical feasibility and the costs of this environmental investment can be estimated. Obtained results show the importance of such cooperation projects. Thanks to project like VOCless is possible to study new impact and to improve the knowledge about environmental issues.

ACKNOWLEDGMENTS
The researchers shown in this paper funded by the European Union, Life Programme, Project: "VOC less pulping: Control of VOC emissions from mechanical pulping beyond BAT" contract LIFE06 ENV/FIN/000201. The authors really thanks all the partner in the VOCless project: AX-LVI Consulting Ltd, coordinator, Reinluft Umwelttechnik BmbH, Ehovoc Oy, M-real Tako Board Massatehdas, Stora Enso Publications Papers Oy and Cartiere di Portonogaro sas, for their invaluable support.

REFERENCES

Adler, Stephen F., 2001. Biofiltration--a primer. *Chemica*l *Engineering Progress* 97: 4.

Boswell, Jim, 2002. Understand the capabilities of bio-oxidation. *Chemical Engineering Progress* 98: 12.

Bruzzone, A.G., Merkuriev, Y. 2009. Advances in Supply Chain Simulation. Proceedings of the *Third Asia International Conference on Modelling & Simulation.* May 25-26 Bandung, May 29 Bali.

Bruzzone, A.G., Tremori, A., Massei, M., Tarone, F. 2009. Modeling Green Logistics. Proceedings of the *Third Asia International Conference on Modelling & Simulation,* 543-548. May 25-26 Bandung, May 29 Bali.

Bhat, Vasanthakumar N.,1999. Does it pay to be green? *International Journal of Environmental Studies*. 56, 497-507.

Chaiporn, Vithessonthi, 2009. Corporate ecological sustainability strategy decisions: the role of attitude towards sustainable development. *Organisational Transformation and Social Change* 6:49-64.

Deshusses, M. A., Shareefdeen, Z. 2005. Modeling of Biofilters and Biotrickling Filters for Odor and VOC Control Applications. in: Shareefdeen, Z., Singh, A., eds. *Biotechnology for Odor and Air Pollution Control*, ISBN (PDF) 9783540270072

Elder, M., 2007. Making Business Green and the Environment Profitable: The Promotion of Environmental Industries and Environmental Management in the EU, Japan, and the US. *Proceedings of 48th Annual Convention of the International Studies Association.* Chicago.

EU, 1999. VOC Directive number 13/99.

Goldstein, D., 2002. Theoretical perspectives on strategic environmental management. *Journal of Evolutionary Economics* 12: 495-524.

Graedel, T. E., Allenby, B. R., 1995. *Industrial ecology*. Englewood Cliffs, NJ: Prentice Hall..

Guash, T., Piera, M. A., 2005. Preface Introduction to Special Issue. *SIMULATION* 81:611.

Hart, P. W., Waite, D. M., Thibault, L., Tomashek, J., Rousseau, M. E, Hill, C., Sabourin M. J., 2009. Selective enzyme impregnation of chips to reduce specific refining energy in alkaline peroxide mechanical pulping. *Holzforschung: International Journal of the Biology, Chemistry, Physics, & Technology of Wood* (0018-3830) 63: 418-423

Hart, S. L., 1997. Beyond Greening: Strategies for a Sustainable World. *Harvard Business Review* 75: 66-76.

Johnsen, K., Grotell, C., Tana, J, 2000. Impact of mechanical pulp mill effluent on egg hatchability on brown trout, *Bulletin of Environmental Contamination and Toxicology* (0007-4861) 64:873-9

Kim, I., 2004. The Promising World of Bio-oxidation, *Chemical Engineering Progress* 100, 1.

King, A. A., Lenox, M. J., 2001. Does It Really Pay to Be Green?: An Empirical Study of Firm Environmental and Financial Performance. *Journal of Industrial Ecology* 5: 105-116.

Kleinheinz, G. T., Wright, P. C., 2009. Biological Odor and VOC Control Process, in: Wang L. K., Pereira N. C., Hung, Y.T., eds. *Handbook of Environmental Engineering, Volume 8: Biological Treatment Processes*, Totowa, NJ: The Humana Press,.

Lankoski, L., 2008. Corporate responsibility activities and economic performance: a theory of why and how they are connected. *Business Strategy and the Environment Business* 17:536–547.

Longo, F., Massei M., 2008. Advanced Supply Chain Protection & Integrated Decision Support System,. Proceedings of The *Second Asia International Conference on Modelling & Simulation,* 716-721. Kuala Lumpur, Malaysia, 13-15 May.

Longo, F., Mirabelli, G., 2008. An Advanced Supply Chain Management Tool Based on Modeling & Simulation. *Computer and Industrial Engineering* 54: 570-588

Mannschereck, K., Bächmann, K., Becker, K.H., Heil, T., Kurtenbach, R., Memmesheimer, M., Mohnen, V., Obermeier, A., Poppe, D., Steinbrecher, R., Schmitz, T., Volz-Thomas, A., Zabel, F., 2002. A Database for Volatile Organic Compounds. *Journal of atmospheric Chemistry* 42:281-286.

Moretti, E. C., Mukhopadhyay, N, 1993. VOC Control – current practices and future trends. *CHEMICAL ENGINEERING PROGRESS* 89:20-26.

Moretti, E. C, 2002. Reduce VOC and HAP emissions. *Chemical Engineering Progress* 98.

O'Connor, B.I., Kovacs, T.G., Voss, R.H.. 1992. The Effect of Wood Species Composition on the Toxicity of Simulated Mechanical Pulping Effluents. *Environmental Toxicology and Chemistry* 11, 1259-1270.

Pietrangeli, B., Bragatto, P., Pittiglio, P., 2008. Potential of biofiltration for VOCs emission control and safety aspects. *International Journal of Environment and Pollution* 32:57-67.

Porter, M. E., 1990. *The Competitive Advantage of Nations*. New York: Free Press.

Raudsepp, W., 2009. Heat recovery in mechanical pulping. *APPITA JOURNAL* 62:179-181

Revah, S., Morgan-Sagastume, J. M., 2005. Methods of Odour and VOC Control. Shareefdeen, Z., Singh, A., eds. *Biotechnology for Odour and Air*

Pollution Control. Berlin Heidelberg, Springer-Verlag.

Russo, M. V., Fouts, P., 1997. A Resource-Based Perspective on Corporate Environmental Performance and Profitability. *Academy of Management Journal* 40:534-559.

Schnitzler, J. P., Bauknecht, N., Brüggemann, N., Einig, W., Forkel, R., Hampp, R., Heiden, A. C., Heizmann, U., Hoffmann, T., Holzke, C., Jaeger, L., Klauer, M., Komenda, M., Koppmann, R., Kreuzwieser, J., Mayer,H., Rennenberg, H., Smiatek, G., Steinbrecher, R., Wildt, J., Zimmer, W., 2002. Emission of biogenic volatile organic compounds: an overview of field, laboratory and modelling studies performed during "Tropospheric research program" 1997-2002. *Journal of Atmospheric Chemistry* 52:159-177.

Seinfeld, J. H., Spyros, N. P., 1998. *Atmospheric Chemistry and Physics – From air pollution to climate change*, Wiley Interscience.

Sueiro, L., Gill, J, 1995. Control of VOC emissions from a TMP plant. Proceedings of *TAPPI International Environmental Conference*, 925-930. May 7, Atlanta, GA, USA.

Zuo, X., Li, Q., 2009. Contamination Characteristics of Effluent from APMP Pulping of Poplar. *Tianjin Keji Daxue Xuebao / Journal of Tianjin University of Science & Technology (1672-6510)* 24:,44-46, 78

AUTHORS BIOGRAPHY

Stefano SAETTA is associate professor at the Industrial Engineering Department of the University of Perugia since 22/12/2004. He obtained the Electronic Degree in 1991 summa cum laude at the University of Perugia where he became assistant professor in 1995. He teaches courses: Industrial Plants for mechanical engineers, Industrial Plants, Management of Logistic systems, Logistics. He was visiting professor at Rutgers University, NJ, USA in 2005 and at the Arizona University, AZ, USA in 2007 and 2008. His research field covers: modelling and simulation of logistic and productive processes, methods for the management of life cycle assessment, discrete event simulation, supporting decision methods, lean production. He was involved in several national and international research projects such as: ISACOAT (2001-2004)- Network on Integrated Scenario Analysis of Metal Coating, as co-responsible, supported by The European Union within V research Framework; GIPIETRE (2005-2006), as responsible, supported by ICE, an Italian agency for international trading and in cooperation with companies from Brazil and China; VOC-LESS (2006-2009) responsible for the Perugia Unit of the project VOC less pulping, supported by European Union within the Life framework. He is responsible of many research projects of his department, funded by Italian ministry of university or by other entities. He is active in ERASMUS exchange programs with the University of Marseille and Salford University. His Internationally Affiliated Faculty of the Industrial and Systems Engineering Department of the Rutgers University. His the director of MISS- McLeod Institute of Simulation Science - Perugia Center and International Co-Vice Director. He is author of more than 70 publications.

Lorenzo TIACCI is assistant professor at the Department of Industrial Engineering of the University of Perugia. He obtained the degree summa cum laude in Mechanical Engineering in 1998 and the Phd in Industrial Engineering at the University of Perugia in 2004. He teaches Inventory Management and Production Planning & Scheduling at the Faculty of Engineering – University of Perugia. He worked as consultant in many research projects on production planning and control, logistics and operations management, in cooperation with research centers and companies. His research fields of interest are: production system design, planning and control; assembly line balancing (argument of my Phd final dissertation: "A simulation based methodology as supporting decision tool in the assembly line balancing problem") through genetic algorithms; discrete event simulation; Interactions between Demand Forecasting and Inventory control policies; Lateral shipments policies in multi-echelon networks. He is author of more than 40 publications.

INDUSTRIAL WORKSTATIONS DESIGN: A REAL CASE STUDY

Antonio Cimino[(a)], Francesco Longo[(b)], Giovanni Mirabelli[(c)], [(d)]Enrico Papoff

[(a) (b) (c) (d)] Modeling & Simulation Center - Laboratory of Enterprise Solutions (MSC-LES)
Mechanical Department, University of Calabria, 87036 Rende (CS), Italy

[(a)]acimino@unical.it, [(b)]f.longo@unical.it, [(c)]g.mirabelli@unical.it, [(d)]e.papoff@unical.it

ABSTRACT
The research work advances the effective design of the most critical workstations belonging to a real industrial plant; to this end, the authors propose an approach based on the integration of Modeling & Simulation tools, several ergonomic standards and the most known work measurement tools. The Modeling and simulation tools allow to implement a three-dimensional environment capable of recreating, with satisfactory accuracy, the evolution over the time of the real workstations. The ergonomic standards consent to evaluate the ergonomic risks level within the system being considered. The work measurement tools permit to calculate the time required for performing all the workstations operations. The effective design of the workstations is achieved by using the simulation model for comparing workstations' alternative configurations. The comparison is based on ergonomic and time indexes related to the ergonomic standards and the work measurement tools. Such comparison allows to choose the workstations final configurations.

Keywords: industrial plants, industrial workstations design, ergonomic standards, work measurement tools

1. INTRODUCTION

Today, industrial workstations design continuously provides challenging problems in terms of interaction between operators and their working environment (ergonomics) and productivity enhancement as well. Let us consider the ergonomics within the industrial environment. The research works developed in the late '90 consider a single ergonomic performance measure, based upon a specific ergonomic standard for the ergonomic redesign of the workstation belonging to industrial plants. Among the ergonomic standards, the following have to be regarded as the most widely used: (i) the NISOH 81 and the NIOSH 91 equations for lifting tasks (NIOSH stands for National Institute for Occupational Safety and Health); (ii) the OWAS for analyzing working postures (OWAS stands for Ovako Working Analysis System); (iii) the Burandt-Schultetus analysis for lifting tasks involving a large number of muscles; (iv) the Garg analysis for assessing the energy expenditure for performing an operation. Further information about the cited ergonomic standards can be found in Garg (1976), Schultetus (1980), Kharu et al. (1981), the Niosh Technical Report 81-122 (1981), the Scientific Support Documentation for the Revised 1991 NIOSH Lifting Equation (1991) and Waters et al. (1994). Examples of research works that propose an ergonomic redesign of industrial workstations based on a single or multiple ergonomic performance measure are Scott and Lambe (1996), Wright and Haslam (1999), Temple and Adams (2000), Waters et al. (2007), Russell et al. (2007), Cimino et al. (2009a). Let us consider now the workstations productivity enhancement within industrial plant workstations. There are many alternatives available for improving productivity. Clearly the effective design of work methods is one of the most important aspects of increased productive output (Cimino et al. 2008a). In this regards, Methods Engineering is a systematic technique for the design and the improvement of work methods, for the introduction of those methods into the workplace, and for ensuring their solid adoption (Zandin, 2001). Motion and time study is at the heart of methods engineering (Ben-Gal and Bukchin, 2002). As reported in Lawrence (2000) the motion study is to determine the best way to perform a job and the time study is to measure the time required for a job to be completed using the best method. As time study tools (also known as work measurement tools), the following have to be regarded as the most important: MTM (Methods Time Measurement) and MOST (Maynard Operation Sequence Techniques). Further information about the cited work measurement tools can be found in Karger and Bayha (1987), Zandin (2001), Cimino et al. (2008b).

Another important issue to take into consideration in the industrial workstations design is whether to analyze directly the real industrial system or by using computerized models. Usually analyzing directly the real workstations requires huge amount of money and time for testing all the workstations configurations, work assignments, work methods as well as "disturbs" processes and activities of the industrial plant. For this

reason, researchers and practitioners very often use simulation as problem solving methodology for creating an artificial history of the system, analyzing its behaviour, choosing correctly, understanding why, diagnosing problems and exploring possibilities (Banks, 1998). Example of research works that propose the effective design of indutrial workstations using Modeling & Simulation tools are Bruzzone (1996), Wilson (1997), Feyen et al. (2000), Bruzzone et al. (2004), Bruzzone and Williams (2004), Santos et al. (2007) and Cimino et al. (2009b), Cimino and Mirabelli (2009), Longo and Mirabelli (2009).

The main contribution of the paper to the state of the art is to propose the effective design of the most critical workstations (Skinning and Assembly workstations) of an industrial plant devoted to produce high pressure hydraulic hoses. The authors propose an approach based on the integration of Modeling & Simulation tools (eM-Workplace and Pro-Engineer), several ergonomic standards (NIOSH 81, NIOSH 91, Burandt Schultetus, OWAS, Garg) and the most known work measurement tools (MTM, MOST). In particular, the simulation is jointly used with virtual three-dimensional environments in which observe the system evolution over the time. The ergonomic standards allows to evaluate the ergonomic risks level affecting the workstations and finally the work measurement tools consent to calculate the time related to each operation performed in the Skinning and Assembling workstations. After the workstations simulation model development, the authors propose several workstations modifications evaluating their effectivness in terms of both ergonomic risks level and time performace.

Before going into details of the study, let us give a brief overview of each section of the paper. Section 2 describes the industrial plant under investigation. Section 3 gives specific details on the workstations simulation model implementation and validation. Section 4 reports a brief description of the ergonomic standards and the work measurement tools. Sections 5 and 6 present the simulation results and the workstations effective redesign, respectively. The last section reports the conclusions that summarize the scientific contribution of the work and the research activities still going on.

2. THE INDUSTRIAL PLANT

The industrial plant considered in this research work manufactures ring nuts, fittings and high pressure hydraulic hoses for the following sectors: industrial, naval, aeronautical, construction, iron, mechanic and railway. The industrial plant was established in 1994 and it is located in the North of Italy covering a surface of about 3700 m^2. The plant-layout is subdivided into 4 different areas:

1. Raw materials warehouse;
2. Mechanical area;
3. Assembly area;
4. Final products warehouse.

A brief description of each area is reported as follows.

The Raw materials warehouse

Here the raw materials for manufacturing ring nuts, fittings and high pressure hydraulic hoses are stored in shelves and pallets located along the whole area. Note that the pallets are placed on the bottom level of each shelf in order to full use the warehouse area. The raw materials are manually moved by means of a multi order picking cart as well as several forklifts are used for the pallets placement. The storage area is 10 m high and covers a surface of 930 m^2.

The Mechanical area

It produces fittings and ring nuts; some of them are used for manufacturing the high pressure hydraulic hoses, the others are final products of the industrial plant. The workstation employs five operators and it is made by 5 numerically controlled machine. The workstation layout covers a surface of 350 m^2.

The Assembly area

It assembles rubber hoses with ring nuts and fittings in order to obtain the final high pressure hydraulic hoses. Note that each hydraulic hose is made by a rubber hose, two fittings and two ring nuts. The assembly area consists of 6 different workstations each one performing specific operations of the hydraulic hoses assembly process. The assembly area employs 12 operators and cover a surface of about 1112 m^2.

The Final Products Warehouse

Here the final products (ring nuts, fittings, rubber hoses and high pressure hydraulic hoses) are stored in shelves and pallets located along the whole area. As the *Raw Materials Warehouse*, note that the pallets are placed on the bottom level of each shelf in order to full use the warehouse area and the final products are moved by means of a multi-order picking cart as well as by using several forklifts. The storage area is 10 m high and covers a surface of about 1395 m^2. Figure 6 shows the plant layout of the warehouse.

2.1. The Skinning workstation and the Assembly workstation

A preliminary analysis carried out by the company top management shows that the Skinning and the Assembly workstations are characterized by several ergonomic issues as well as low productivity levels. In this context, the company top management asked us to redesign such workstations, from the one side, for preventing the workers' health, from the other side, for increasing their productivity and in turn, the industrial plant one. Before getting into details of the research study, let us present the work methods and the layout of both the workstations.

The Skinning workstation employs 2 workers performing the following operations:

1. They pick up the Shop Order sheet, read the information they need and put it back;
2. They set the skinning machine up;

3. They pick manually up a rubber hose located on a pallet 15 cm high;
4. They insert the rubber hose into the skinning machine, perform the security procedure and start the skinning phase;
5. They remove the rubber hose from the skinning machine and put it on a bin placed on a manual hand chart 30 cm high;
6. They set the status “end of the operation” on the company informative system;
7. They moves the rubber hoses to the successive workstation by means of a manual hand chart.

Note that the skinned rubber hoses are used for manufacturing the high pressure hydraulic hoses as well as are directly sold to the customers; in this context, some of them is moved to the assembly workstation, the others are moved to the final product warehouse.

Table 1 consists of description, dimensions (length, width and height) and quantity of all the objects of the Skinning workstation.

Table 1. Skinning workstation objects

Objects description	Dimensions (cm) (L x W x H)	Quantity
Empty bin	60 x 40 x 30	4
Rubber hose	Depending of Shop Order	Depending of Shop Order
Pallet	80 x 120 x 15	2
Skinning Machine	300 x 150 x 250	2
PC Worktable	100 x 65 x 95	2
Support Table	180 x 60 x 95	2
Manual hand chard	100 x 140 x 15	2

Figure 1 shows the objects position as it takes place in the Skinning workstation.

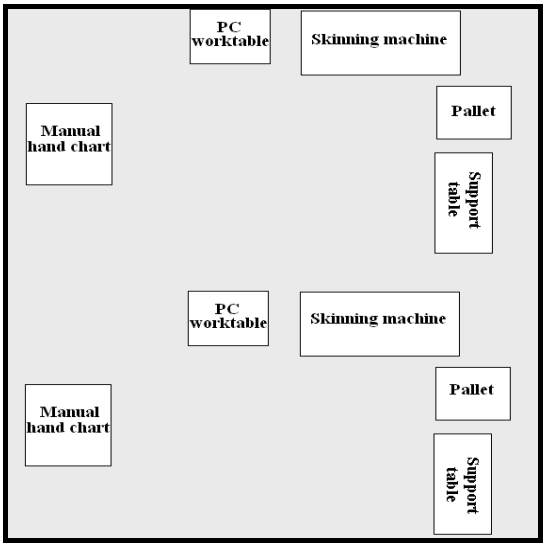


Figure 1. Skinning workstation layout

The Assembly workstation employs 2 workers and it presents the following work method:

1. The workers pick the Shop Order sheet up, read the information they need and put it back;
2. The workers picks manually up a rubber hose located on a pallet 15 cm high and bring it to the work table;
3. The workers pick manually two ring nuts and two fittings up from bins located on their backs and bring them to the work table;
4. The workers manually perform the assembly operation;
5. The workers move the assembled hydraulic hoses to a pallet located on a manual hand chart 30 cm high;
6. The workers set the status “end of the operation” on the company informative system;
7. The workers moves the hydraulic hoses to the successive workstation by means of a manual hand chart.

Table 2 consists of description, dimensions (length, width and height) and quantity of all the objects of the Assembly workstation.

Table 2. Assemvly workstation objects

Objects description	Dimensions (cm) (L x W x H)	Quantity
Ring nuts and fittings bins	20 x 15 x 15	98
Rubber hose	Depending of Shop Order	Depending of Shop Order
Pallet	80 x 120 x 15	2
Worktable	400 x 150 x 95	2
PC Worktable	180 x 60 x 95	2
Manual hand chard	100 x 140 x 15	2

Figure 2 shows the objects position as it takes place in the Assembly workstation.

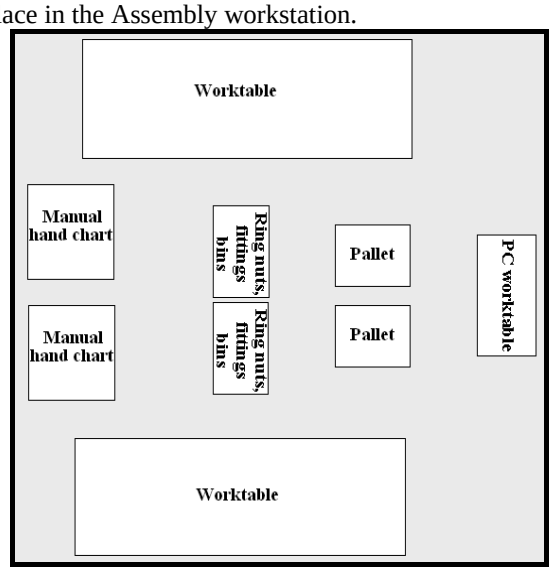


Figure 2. Assembly workstation plant layout

It is important to underline that the Skinning workstation productivity is higher than the Assembly workstation one. It generates a skinned rubber hoses inventory to be managed by the operators. The rubber

hoses are stored by means of several shelves located between the two workstations and their management causes a notable reduction of the workstations productivity; for this reason, the company top management asked us to face and solve the problem in order to avoid such noteworthy loss of time.

3. SIMULATION MODEL DEVELOPMENT

Industrial plants workstations are very complex systems characterized by several design parameters such as objects dimensions, tools positions and operators work method. As a consequence, the workstations redesign process should be supported by an approach capable to correctly recreate the complexity of the real system. To this end, the authors decide to adopt an approach based on Modeling & Simulation supported by a three dimensional virtual environment. In this context, the first step of the research work was the development of a simulation model capable of recreate with satisfactory accuracy the evolution over the time of the Skinning and Assembly workstations. The steps to guide the model builder in the simulation study are reported as follows:

1. Data collection phase: collect data concerning the system under consideration;
2. Modeling phase: reproduce the real system from a geometric point of view;
3. Simulation phase: reproduce the real system from a work method point of view;
4. Validation phase: verify if the simulation model is an accurate representation of the real system.

Sections 3.1–3.4 get into details of the four phases of the simulation model development.

3.1. Data collection phase

As first step, the authors submitted a schedule of data requirement to the company top management. However the company top management indicated that the data required were unavailable. Therefore the authors spent a three months period at the Skinning and Assembly workstations collecting data and information about operators' characteristics (age, gender, height, weight and physical condition), dimensions (length, width and height) and weights of all the objects being modeled and analyzing the work methods used by workers for performing the manufacturing operations. Operators' characteristics were used for selecting human models capable of representing as much as possible the real workers. Objects' dimensions and weights were used for designing the geometric models of each workstation. The observation of the work methods was used for reproducing correctly in the virtual environment the manufacturing operations of each workstation.

3.2. Modeling phase

After the data collection phase, the second step was the reproduction of the Skinning and Assembly workstations from a geometric point of view. The geometric models of all the workstations objects were developed by means of the CAD software Pro-Engineer (Pro-E). The geometric models generated by using Pro-E contain all the information regarding dimensions, weights and type of materials. For each workstation the geometric models recreate the following elements: machines, equipment and tools, worktables, manual hand charts, raw materials, containers and cases.

Figure 3 and figure 4 show respectively the Assembly workstation worktable and a pallet with several rubber hoses.

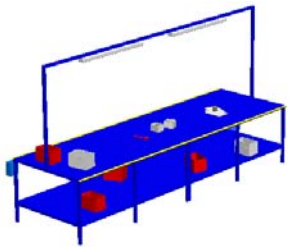


Figure 3. Worktable of the Assembly workstation

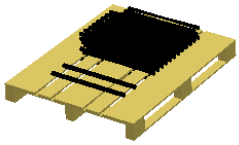


Figure 4. Pallet and rubber hoses

3.3. Simulation phase

After the modeling phase, the successive step was the reproduction of the Skinning and Assembly workstations from a work method point of view. The first step of the Simulation phase requires to import the geometric models into the virtual environment provided by the simulation software eM-Workplace.

Figure 5 and figure 6 provide a panoramic view of the Skinning and the Assembly workstation.

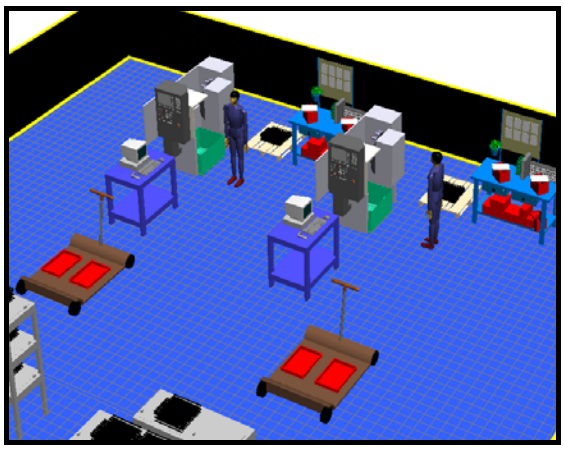


Figure 5. Panoramic view of the Skinning workstation.

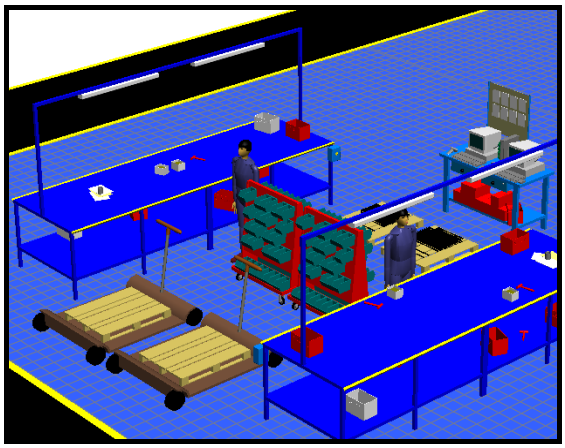


Figure 6. Panoramic view of the Assembly workstation.

The insertion and training of the human model are the successive steps required for completing the simulation phase. The eM-Workplace provides the user with different human models libraries. The selection of the human models takes into account the characteristics of the real operators (age, gender, height, weight and physical condition) with the aim of importing in the virtual environment human models as much as possible similar to the real workers. Table 3 consists of operators' characteristics in terms of age, gender, height, weight.

Table 3. Operators' physical characteristics

Operator ID	Age	Gender	Height (cm)	Weight (kg)	Workstation
Op-1	32	Male	172	73	Skinning
Op-2	29	Male	175	71	Skinning
Op-3	44	Male	169	74	Assembly
Op-4	39	Male	178	78	Assembly

Obviously each human model needs to be trained in order to perform the manufacturing operations. To this end, eM-Workplace provides the user with a programming language for teaching the basic motions of each operation.

3.4. Validation phase

The last step of the simulation model development is the validation that aims at determining if the simulation model is an accurate representation of the real industrial plant workstations.

The validation phase have been carried out by analyzing and discussing the simulation model with workers and employees of the industrial plants. With the help of the workers the authors checked all the basic motions of the human models and deleted some errors concerning the work methods (wrong working postures, wrong motions or redundant motions). At the end of this phase the simulation model was ''reasonable'' both to workers, company's engineers and technicians for its capability to recreate correctly the workstations layout and all the manufacturing operations as well.

4. MEASUREMENT AND STUDY OF WORK

The eM-Workplace simulation software provides the user with several ergonomic and time indexes based on the most known ergonomic standards and work measurement tools. The ergonomic standards allow the authors to establish the ergonomic risks level affecting the workstations; the work measurement tools consent to calculate the time required for performing the operation within the workstations being considered. As concerns the ergonomic standards, the Burandt Schultetus, NIOSH 81 and NIOSH 91 were used for evaluating the stress related to the lift operations, OWAS analysis for evaluating the stress associated to the working postures and finally the Garg analysis for calculating the energy expenditure associated to each activity. As regards the work measurement tools, the simulation model uses the MTM and the MOST for evaluating the process time of each operation.

Sections 4.1 – 4.6 present a brief description of the ergonomic standards and work measurement tools being used in this research work. Note that further information concerning the ergonomic standards can be found in Cimino et al., 2008c.

4.1. Burandt Schultetus

The analysis detects the maximum weight that a working person can lift (maximum permissible force). The analysis requires several input parameters regarding the physical conditions, age and gender of the worker, the load weight, the lifting frequency (measured in lifts per minute) and the total task duration.

4.2. NIOSH 81 and NIOSH 91

The NIOSH 81 method calculates the Action Limit (AL) and the Maximum Permissible Limit (MPL). AL is the weight value, which is permissible for 75% of all female and 99% of all male workers. MPL is the weight value, which is permissible for only 1% of all female and 25% of all male workers. Concerning the NIOSH 91 analysis, additionally to the NIOSH 81, it includes the Recommended Weight Limit (RWL) and the Lifting Index (LI). The RWL is the load that nearly all healthy workers can perform over a substantial period of time for a specific set of task conditions. The LI is calculated as a ratio between the real object weight and the RWL.

4.3. OWAS analysis

The analysis calculates the stress associated to each body posture and classifies them in one of the following four stress categories:

- *Category 1*: If the stress level is optimum, no corrective interventions are required
- *Category 2*: If the stress level is almost acceptable, corrective interventions are necessary in the near future
- *Category 3*: If the stress level is high, corrective interventions are required as soon as possible

- *Category 4*: If the stress level is very high, corrective
- interventions must be carried out immediately.

4.4. Garg analysis

The total amount of energy spent during the manual operations is calculated. The analysis splits up a specified operation into smaller steps calculating for each of them the EE; the sum of these separate steps represents the total EE for the activity.

4.5. MTM

The Method Time Measurement is the most widely used system of predetermined times (Rice, 1977). The MTM is a procedure for analyzing any manual operation or method by breaking out the basic motions required to perform it and assigning to each a predetermined standard time based on its nature and the conditions under which it is made (Karger and Bayh 1987). The total time for the manual operation is then calculated as sum of the time of each basic motion it is made by.

4.6. MOST

MOST concentrates on the movement of objects (Zandin and Kjell 1990). The primary work units are no longer basic motions, but fundamental activities (collection of basic motions) dealing with moving objects. These activities are described in terms of sub activities fixed in sequence. In other words, to move an object, a standard sequence of events occurs. Only three activity sequences are needed for describing manual work. Summarizing the MOST technique is made up of the following basic sequence models:

- The *general move sequence* for the spatial movement of an object freely through the air.
- The *controlled move sequence* for the movement of an object when it remains in contact with a surface or is following a controlled path during the movement.
- The *tool use sequence* for the use of common hand tools.

5. SIMULATION RESULTS

In this section the authors present the simulation results of both the Skinning and the Assembly workstations. In particular the authors use the simulation model, the ergonomic standards and the work measurement tools for evaluating the ergonomic risks level affecting the workstations and for calculating the total time required for performing the workstations operation as well. Note that the simulation runs consider the production of a typical industrial plant Shop Order made by 20 hydraulic hoses. Section 5.1 and 5.2 present the simulation results for the Skinning and the Assembly workstations, respectively.

5.1. Skinning workstation

The activities performed by the operators do not require heavy lifting tasks; in effect the Burandt Schultetus, the NIOSH 81 and the NIOSH 91 analysis do not reveal any particular lifting problem. Significant results for the effective design have been obtained in terms of uncomfortable working postures and Energy Expenditure (EE) respectively for the OWAS and the Garg analysis. When the OWAS analysis is applied to the Skinning workstation, the program assigns a category 3 (body posture characterized by high stress level) to the following operations:

1. picking manually up a rubber hose located on a pallet 15 cm high before the skinning operation (see section 2.1, operation 3 of the Skinning workstation);
2. putting a rubber hose on a bin located on a manual hand chart 30 cm high after the skinning operation (see section 2.1, operation 5 of the Skinning workstation).

Considering both the operations, the most affected body part is the workers' back.

The Garg analysis completes the ergonomic evaluation process calculating about 2340 Kcal as the total amount of energy spent during the whole shift.

Consider now the work measurement analysis. The operations performed in this workstation have been subdivided in 4 different groups (each group has to be regarded as a macro-activity), described as follows.

- *Macro-activity 1* – the operators set the workstation for starting the skinning operations;
- *Macro-activity 2* – the operators move the component (rubber hoses) into the skinning machine and start the skinning phase;
- *Macro-activity 3* – after the skinning phase the operators remove the components from the skinning machine and put them into a bin;
- *Macro-activity 4* – the operators complete the Shop Order and move the rubber hoses to the successive workstation.

The authors suppose to subdivide the macro-activities in two different categories: preparation operations (performed just once for the entire Shop Order) and cyclic operations. The macro-activities 1 and 4 (workstation set-up and Shop Order completion) belong to the first category. The macro-activities 2 and 3 belong to the second category. Table 4 and table 5 consist of process times for each macro-activity

Table 4. Process time evaluated by the MTM analysis

MTM analysis	
Preparation operation	**Time (sec.)**
Macro-activity 1	6.19
Macro-activity 4	184.97
Total Preparation Time	191.16
Cyclic operation	**Time (sec.)**
Macro-activity 2	222.04
Macro-activity 3	415.12

Total Cyclic Time	637.16
Total time for completing the Shop Order	**828.74**

Table 5. Process time evaluated by the MOST analysis

MOST analysis	
Preparation operation	**Time (sec.)**
Macro-activity 1	7.01
Macro-activity 4	185.52
Total Preparation Time	192.53
Cyclic operation	**Time (sec.)**
Macro-activity 2	222.94
Macro-activity 3	416.24
Total Cyclic Time	639.18
Total time for completing the Shop Order	**831.71**

As concerns the MTM, the total process time is 828.74 sec. (about 12 min and 8 sec.). As concerns the MOST, the total process time is 831.71 sec (about 12 min and 11 sec). Let us focus on the Skinning workstation productivity. It has been evaluated by taking into account the total time required for completing a Shop Order (process time), the 8 hours shift of the workstation and the operators' allowance for physiological needs, fatigue and delay (calculated as 20% of the process time). Regardless of the work measurement tools (MTM or MOST) used for the evaluation of the process time the workstation productivity is about 29 Shop Orders per day.

5.2. Assembly workstation

As the Skinning workstation, the activities performed by the operators do not require heavy lifting tasks and significant results for the workstation effective design have been obtained only in terms of uncomfortable working postures and Energy Expenditure. The OWAS analysis identifies the following operations as the most critical ones:

1. picking manually up a rubber hose located on a pallet 15 cm high before the assembly operation (see section 2.1, operation 2 of the Assembly workstation). The analysis classifies such operation within the OWAS category 3 and identifies the workers' back as the most affected body part;
2. moving the assembled hydraulic hoses to a pallet located on a manual hand chart 30 cm high after the assembly operation (see section 2.1, operation 5 of the Assembly workstation). As the previous operation, a high stress level affects the workers' back.

As concerns the Garg analysis, the total amount of energy spent for the whole shift is about 2870 Kcal.

Let us consider the work measurement analysis.

The operations performed in this workstation have been subdivided in 5 different macro-activities, reported as follows.

- *Macro-activity 1* – the operators set the workstation for starting the assembly operations;
- *Macro-activity 2* – the operators move the rubber hoses to the work table;
- *Macro-activity 3* – the operators move the ring nuts and fittings to the work table and start the assemby phase;
- *Macro-activity 4* – after the assembly phase the operators move the high pressure hydraulic hose to the pallet locate on the manual hand chart;
- *Macro-activity 5* – the operator completes the Shop Order.

As the skinning workstation, the authors subdivide the macro-activities in preparation operations and cyclic operations. The macro-activities 1 and 5 (workstation set-up and Shop Order completion) belong to the first category. The macro-activities 2, 3 and 4 belong to the second category. Table 6 and table 7 consist of process times for each macro-activity.

Table 6. Process time evaluated by the MTM analysis

MTM analysis	
Preparation operation	**Time (sec.)**
Macro-activity 1	8.14
Macro-activity 5	19.94
Total Preparation Time	28.08
Cyclic operation	**Time (sec.)**
Macro-activity 2	180.94
Macro-activity 3	584.54
Macro-activity 4	380.13
Total Cyclic Time	1145.61
Total time for completing the Shop Order	**1173.69**

Table 7. Process time evaluated by the MOST analysis

MOST analysis	
Preparation operation	**Time (sec.)**
Macro-activity 1	8.67
Macro-activity 5	20.74
Total Preparation Time	29.41
Cyclic operation	**Time (sec.)**
Macro-activity 2	181.78
Macro-activity 3	586.23
Macro-activity 4	380.96
Total Cyclic Time	1148.97
Total time for completing the Shop Order	**1178.38**

According to the MTM analysis the total process time is 1173, 69 (about 19 min and 33 sec). According to the MOST, the total process time is 1178,38 (about 19 min and 38 sec.). The Assembly workstation productivity (evaluated by taking into account the total time required for completing a Shop Order, the 8 hours shift, the operators' allowance for physiological needs, fatigue and delay and regardless of the work measurement tools) is about 20 Shop orders per day.

6. THE WORKSTATIONS EFFECTIVE DESIGN

In this section the authors achieve the effective design of both the Skinning and the Assembly workstations. Section 6.1 and section 6.2 present respectively the Skinning and the Assembly workstations redesign.

6.1. The Skinning workstation redesign

Let us present the changes the authors propose for reducing the ergonomic risks and for increasing the productivity level of the workstation.

- A manual dolly replaces the pallet being used for locating the rubber hoses before the skinning operations. This change allows the operators to avoid the continuous bending needed for picking the rubber hoses up.
  Figure 7 and figure 8 shows the actual configuration and the final solution, respectively.

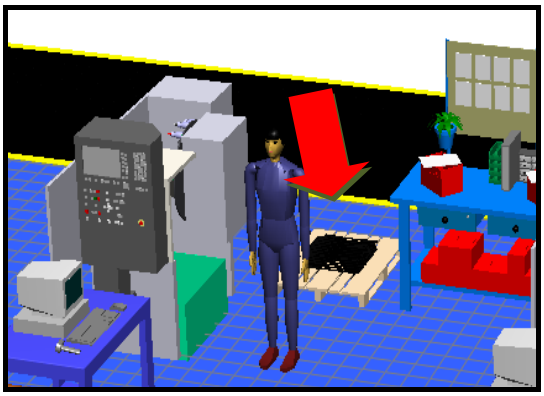


Figure 7. Actual workstation configuration for picking the rubber hose up before the skinning operation.

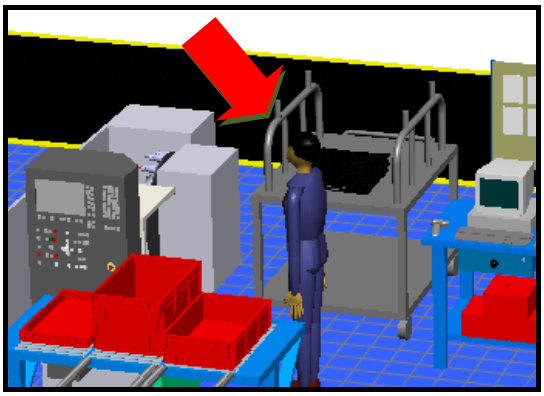


Figure 8. Final workstation configuration for picking the rubber hose up before the skinning operation.

- The PC being used by the operator for setting the status "end of operation" on the company informative system, has been moved to the support table. Note that such change allows to reduce the number of steps required by the operator for reaching the PC worktable; figure 9 and figure 10 show respectively the actual and the final workstation configurations.

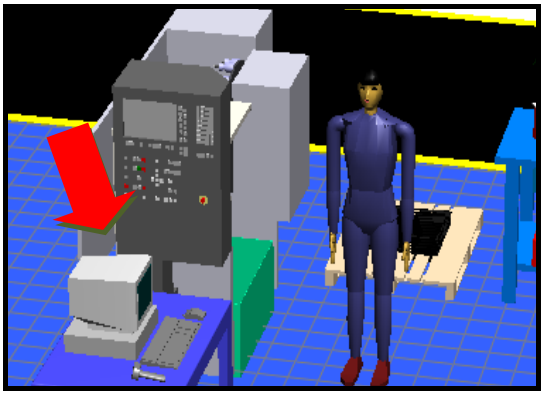


Figure 9. Actual workstation configuration of the PC location.

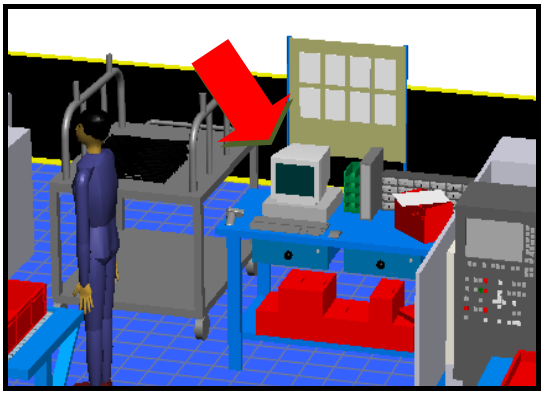


Figure 10. Final workstation configuration of the PC location.

- A manual conveyor replaces the manual hand chart for moving the skinned rubber hoses to the successive workstation. Such change allows to notably reduce the time required for performing this operation; in fact, after the skinning phase, the workers put the skinned rubber hoses on a bin located on the manual conveyor and then, providing a slight push to the bin, move the rubber hoses to the Assembly workstation. Moreover the new configuration consents to manage effectively and efficiently the rubber hoses inventory owing to the different productivity levels of the Skinning and the Assembly workstations; in effect, the rubber hoses can be directly stored into the bins placed on the manual conveyor, instead of being stocked on the shelved located between the workstations. Figure 19 and figure 20 show the actual configuration and the final solution, respectively.

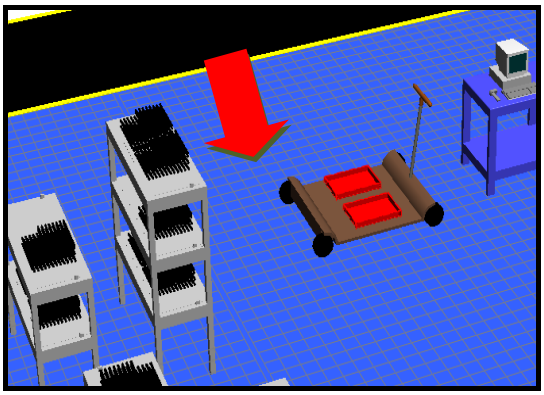


Figure 11. Actual workstation configuration for moving the skinned rubber hoses to the successive workstation.

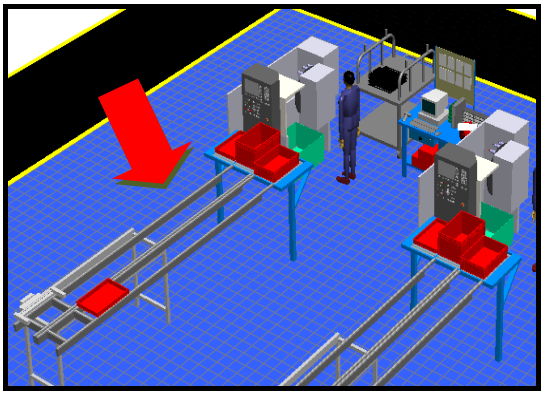


Figure 12. Final workstation configuration for moving the skinned rubber hoses to the successive workstation.

The workstation changes have been tested by means of the simulation model evaluating the ergonomic risk levels and the time required for performing the skinning operations within the new configuration. The ergonomic evaluation process did not detect any ergonomic problem. Moreover the new configuration requires to the workers about 1780 Kcal, as total amount of energy spent during the whole shift. Note that the EE reduction respect to the initial configuration is about 24%.

Concerning the work measurement analysis, Table 8 and table 9 reports the process times for each macro-activity performed within the new workstation configuration.

Table 8. Process time evaluated by the MTM analysis within the new Skinning workstation configuration

MTM analysis	
Preparation operation	**Time (sec.)**
Macro-activity 1	6.19
Macro-activity 4	70.16
Total Preparation Time	76.35
Cyclic operation	**Time (sec.)**
Macro-activity 2	192.04
Macro-activity 3	255.12
Total Cyclic Time	397.16
Total time for completing the Shop Order	**523.51**

Table 9. Process time evaluated by the MOST analysis within the new Assembly workstation configuration

MOST analysis	
Preparation operation	**Time (sec.)**
Macro-activity 1	7.01
Macro-activity 4	72.56
Total Preparation Time	79.57
Cyclic operation	**Time (sec.)**
Macro-activity 2	193.01
Macro-activity 3	256.32
Total Cyclic Time	449.33
Total time for completing the Shop Order	**528.9**

According the MTM analysis the total time required for completing the Shop Order is 523,51 sec. (about 8 min and 43 sec.). According the MOST analysis the total time required for completing the Shop Order is about 528,9 sec. (about 8 min and 48 sec.). Note that the process time improvement respect to the initial configuration is about 58%. The workstation producitity is 45 Shop Orders per day and the productivity improvement is about 56% respect to the initial workstation configuration.

6.2. The Assembly workstation redesign

Let us list the workstation changes the authors propose for reducing the ergonomic risks levels and for increasing the productivity levels within the workstation.

- A work table replaces the pallet being used for locating the rubber hoses before the assembly operations. This change allows the operators to avoid the continuous bending needed for picking the rubber hoses up. Figure 13 and figure 14 shows the actual configuration and the final solution, respectively.

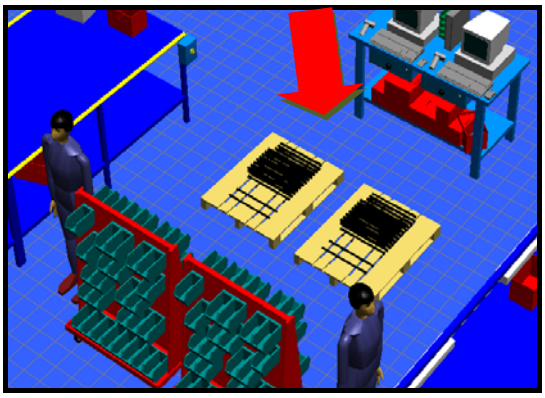


Figure 13. Actual workstation configuration for picking the rubber hose up before the assembly operation.

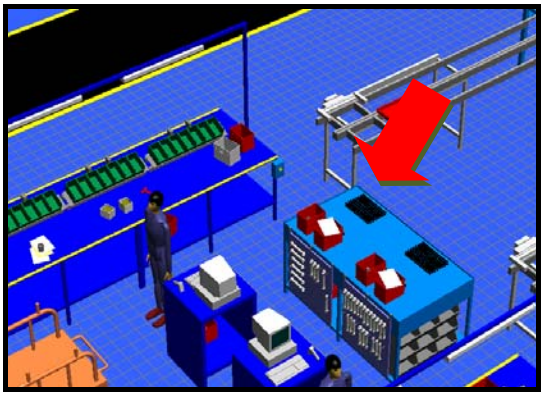


Figure 14. Final workstation configuration for picking the rubber hose up before the assembly operation.

- A manual dolly replaces the pallet being used for locating the high pressure hydraulic hoses after the assembly operations. Such change allows the operators to avoid the continuous bending needed for performing this operation. Figure 15 and figure 16 shows the actual configuration and the final solution, respectively.

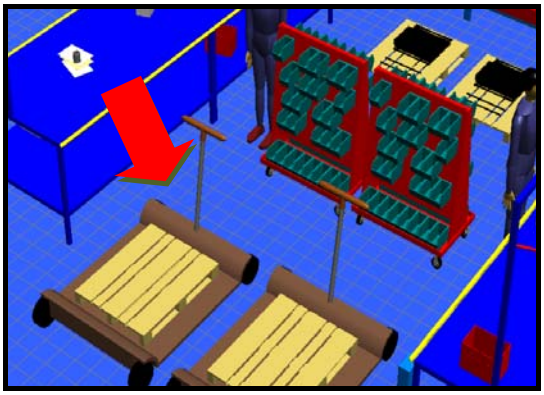


Figure 15. Actual workstation configuration for moving the high pressure hydraulic hoses to the successive workstation.

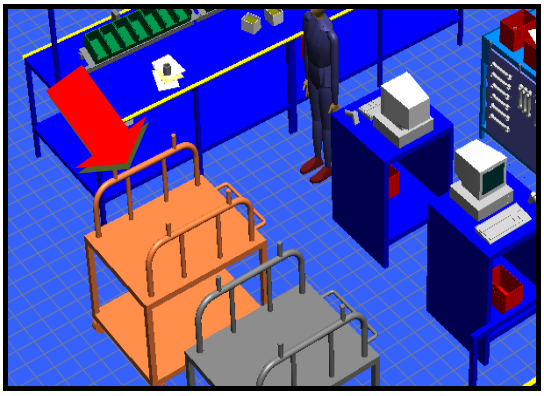


Figure 16. Final workstation configuration for moving the high pressure hydraulic hoses to the successive workstation.

- The PC being used by the operator for setting the status "end of operation" on the company informative system, has been moved closer to the assembly work table. Note that such change allows to reduce the number of steps required by the operator for reaching the PC worktable; figure 17 and figure 18 show respectively the actual and the final workstation configurations.

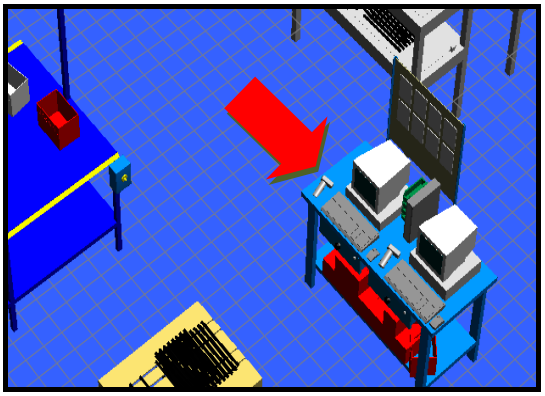


Figure 17. Actual workstation configuration of the PC location.

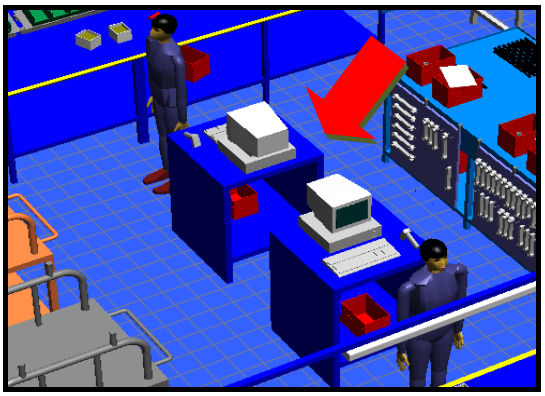


Figure 18. Final workstation configuration of the PC location.

- The ring nuts and fittings bins have been placed to the work table in order to reduce the time required to the operators for reaching and managing such components. Figure 19 and figure 20 show respectively the actual and the final workstation configurations.

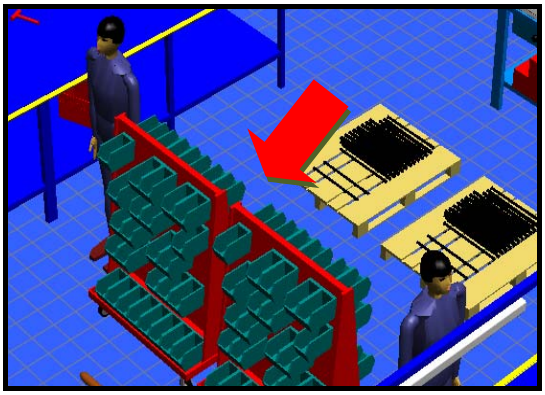


Figure 19. Actual workstation configuration of the ring nuts and fittings bins location.

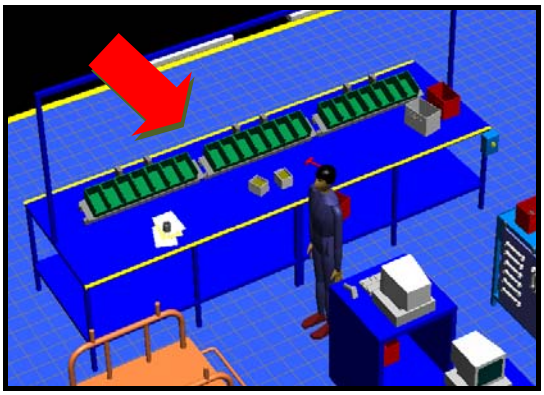


Figure 20. Final workstation configuration of the ring nuts and fittings bins location.

After the workstation changes implementation within the simulation model, the next step was the evaluation of the ergonomic risk levels and the calculation of time required for performing the assembly operations.

As the skinning workstation, the ergonomic evaluation process did not detect any problem related to the lifting tasks as well as no harmful body postures were identified. The total amount of energy spent during the whole shift is about 2340 Kcal. Note that the EE rduction respect to the initial configuration is about 18%.

Concerning the work measurement analysis, Table 10 and table 11 reports the process times for each macro-activity performed within the new workstation configuration.

Table 10. Process time evaluated by the MTM analysis within the new Skinning workstation configuration

MTM analysis	
Preparation operation	**Time (sec.)**
Macro-activity 1	8.14
Macro-activity 5	13.94
Total Preparation Time	22.08
Cyclic operation	**Time (sec.)**
Macro-activity 2	124.85
Macro-activity 3	368.57
Macro-activity 4	255.14
Total Cyclic Time	748.56
Total time for completing the Shop Order	**770.64**

Table 11. Process time evaluated by the MOST analysis within the new Assembly workstation configuration

MOST analysis	
Preparation operation	**Time (sec.)**
Macro-activity 1	8.67
Macro-activity 5	14.45
Total Preparation Time	23.12
Cyclic operation	**Time (sec.)**
Macro-activity 2	125.01
Macro-activity 3	369.05
Macro-activity 4	255.68
Total Cyclic Time	749.74
Total time for completing the Shop Order	**772.86**

According the MTM analysis the total time required for completing the Shop Order is 770,64 sec. (about 12 min and 50 sec.). According the MOST analysis the total time required for completing the Shop Order is about 772,86 sec. (about 12 min and 52 sec.). Note that the process time improvement respect to the initial configuration is about 53%. The workstation producitity is 31 Shop Orders per day and the productivity improvement is about 55% respect to the initial workstation configuration.

7. CONCLUSIONS

The paper advances the effective design of the most critical workstations (the Skinning workstation and the Assembly workstation) belonging to a real industrial plant. The authors propose an approach based on the integration of Modeling & Simulation tools, several ergonomic standards and the most known work measurement tools. The first step of the research work was the development of the workstations simulation model. The simulation model has been developed by using the CAD software Pro-Engineer and the simulation software eM-Workplace. After the simulation model validation, the ergonomic standards were accomplished for evaluating the ergonomic risks level affecting the workstations as well as the work measurement tools were used for calculating the time related to each operation performed within the Skinning and the Assembly workstations. The next step was the achievemnt of the workstations effective deisgn; in particular, the authors use the simulation model for comparing several workstations alternative configurations. Each workstations configuration was recreated within the simulation model and then all the configurations were compared by means of ergonomic and time indexes related to the ergonomic standards and work measurement tools used in this research work. Such comparison allows the authors to choose the workstations final configurations. The Skinning and the Assembly workstations final configurations do not present any ergonomic issue and are characterized by productivity levels higher than the initial ones. Further research activities are still going on (in cooperation with the same industrial plant) for analyzing the remaining workstations of the Assembly area.

REFERENCE

Banks. J., 1998. Principles of Simulation, in: Banks J. (Eds.), Handbook of Simulation. Wiley Interscience, New York, pp. 3-30.

Ben-Gal, I., Bukchin, J., 2002. The ergonomic design of workstations using virtual manufacturing and response surface methodology. IIE Transaction, 34, 375-391.

Bruzzone A.G., 1996. *Object Oriented Modelling to Study Individual Human Behaviour in the Work Environment: a Medical Testing Laboratory*, Proc. of WMC'96, San Diego, January.

Bruzzone A.G., Williams E., 2004. *Modeling and Simulation Methodologies for Logistics and Manufacturing Optimization*. Journal of Simulation, Vol 80, No.3 pg 119-174.

Bruzzone A.G., Viazzo S., Briano C., Massei M., 2004. *Modelling Human Behaviour in Industrial Facilities & Business Processes*. Proc. of ASTC2004, Arlington, VA, April.

Cimino A., Curcio D., Longo F. Papoff E., 2008a. *Workstation productivity enhancement within hydraulic hoses manufacturing process*. Proceedings of the 7th International Workshop on Modeling & Applied Simulation, 17-19 September, Campora S. Giovanni, Amantea (CS), Italy, pp. 268-274.

Cimino A., Longo F., Mirabelli G., Papoff E., 2008b. *MOST and MTM for work methods optimization: a real case study based on Modeling & Simulation.* Proceedings of the 11th International Workshop on Harbor, Maritime & Multimodal Logistics Modeling & Simulation, 17-19 September, Campora S. Giovanni, Amantea (CS), Italy, pp. 35-41.

Cimino A., Curcio D., Longo F., Mirabelli G., 2008c. *Workplaces effective ergonomic design: a literature review*. Proceedings of the 20th European Modeling & Simulation Symposium, *17-19 September, Campora S. Giovanni, Amantea (CS), Italy, pp. 853-862.*

Cimino A., Mirabelli G., 2009. *Modeling, simulation and ergonomic standards as support tools for a workstation design in manufacturing system*. International Journal of Simulation and Process Modeling, Vol. 5, No. 2, pp. 138-148

Cimino A., Curcio D., Longo F., Mirabelli G., 2009a. *Improving workers conditions within Industrial Workstations.* Proceedings of the Summer Computer Simulation Conference, Istanbul, Turkey.

Cimino A., Longo F., Mirabelli G., 2009b. *Multi-Measures Based Approach for the Ergonomic Effective Design of Manufacturing System Workstations.* International Journal of Industrial Ergonomics, *39, pp 447-455.*

Feyen, R., Liu, Y., Cha, D., Jimmerson, G., Joseph, B., 2000. Computer-aided ergonomics: a case study of incorporating ergonomics analyses into workplace design. Appl. Ergon. 31, 291-300.

Garg A., 1976. A metabolic rate prediction for manual materials handling jobs. Dissertation, University of Michigan.

Karger, O., Bayh, F., 1987. Engineered Work Measurement, New York: Industrial Press.

Kharu, O., Harkonen, R., Sorvali, P., Vepsalainen, P., 1981. Observing working postures in industry: Examples of OWAS application. Appl Ergon 12, 13-17.

Lawrence, S.A., 2000. Work Measurement & Methods Improvement, New York: John Wiley & Sons.

Longo F., Mirabelli G., 2009. Effective Design of an Assembly Line using Modeling & Simulation, International Journal of Simulation, 2009(3), pp. 50-60.

Niosh Technical Report 81-122. National Institute for Occupational Safety and Health (Hrsg.). Work practices guide for manual lifting. Center for Disease Control, U.S. Department of health and human services, Cincinnati, OH, USA: NTIS 1981.

Russell S. J., Winnemuller L., Camp J. E., Johnson P. W., 2007. Comparing the results of five lifting analysis tools. Appl. Ergon. 38, 91-97.

Santos, J., Sarriegi, J. M., Serrano, N., Torres, J. M., 2007. Using ergonomic software in non-repetitive manufacturing processes: A case study. Int. J. Ind. Ergonom. 37, 267-275.

Schultetus, W., 1980. Daten, hinweise und beispiele zur ergonomischen arbeitsgestaltung. Montagegestaltung, Verlag TÜV Rheinland Gmbh, Köln.

Scientific Support Documentation for the Revised 1991 NIOSH Lifting Equation: Technical Contract Reports, May 8, 1991, NTIS No. PB-91-226-274.

Scott, G. B., Lambe, N. R., 1996. Working practices in a pherchery system, using the OVAKO Working posture analysing System (OWAS). Appl. Ergon. 27, 281-284.

Temple, R., Adams, T., 2000. Ergonomic Analysis of a Multi-Task Industrial Lifting Station Using the NIOSH Method. The Journal of Industrial Technology, 16, 1-6.

Waters, T., R., Vern, P., A., Garg, A., 1994. Application Manuals for the Revised NIOSH Lifting Equation. U.S. Department of health and human services, National Institute for Occupational Safety and Health, Cincinnati, OH, USA.

Waters, T.R., Lu, M.L., Occhipinti, E., 2007. New procedure for assessing sequential manual lifting jobs using the revised NIOSH lifting equation. Ergonomics 50, 1761–1770.

Wilson, J. R., 1997. Virtual environments and ergonomics: needs and opportunities. Ergonomics, 40, 1057-1077.

Wright, E.J., Haslam, R.A., 1999. Manual handling risks and controls in a soft drinks distribution centre. Appl. Ergon. 30, 311–318.

Zandin K.B., 2001. Maynard's Industrial Engineering Handbook, 5th ed. New York: McGraw-Hill.

AUTHORS BIOGRAPHY

ANTONIO CIMINO was born in Catanzaro (Italy) in October the 1th, 1983. He took his degree in Management Engineering, summa cum Laude, in September 2007 from the University of Calabria. He is currently PhD student at the Mechanical Department of University of Calabria. His research activities concern the integration of ergonomic standards, work measurement techniques, artificial intelligence techniques and Modeling & Simulation tools for the effective workplace design. He collaborates with the Industrial Engineering Section of the University of

Calabria to research projects for supporting innovation technology in SMEs.

FRANCESCO LONGO took the degree in Mechanical Engineering from University of Calabria (2002) and the PhD in Industrial Engineering (2005). He is currently researcher at the Mechanical Department (Industrial Engineering Section) of University of Calabria. His research interests regard modeling & simulation of manufacturing systems and supply chain management, vulnerability and resilience, DOE, ANOVA. He is Responsible of the Modeling & Simulation Center – Laboratory of Enterprise Solutions (MSC-LES), member organization of the MS&Net (McLeod Modeling & Simulation Network) He is also member of the Society for Computer Simulation International and Liophant Simulation.

GIOVANNI MIRABELLI was born in Rende in 1963 and he took the degree in Industrial Engineering at the University of Calabria. He is currently researcher at the Mechanical Department of University of Calabria. His research interests include ergonomics, methods and time measurement in manufacturing systems, production systems maintenance and reliability, quality. He has published several scientific papers participating as speaker to international and national conferences. He is actively involved in different research projects with Italian and foreign universities as well as with Italian small and medium enterprises.

ENRICO PAPOFF was born in Naples (Italy) on February the 03rd, 1948. He took the degree in Mechanical Engineering from University of Napoli Federico II, in 1973. He is currently Associate Professor at the Mechanical Department (Industrial Engineering Section) of the University of Calabria. His research interests regard project management and business plans.

PROBLEMATIC RESEARCH DESIGN IN QUALITY MANAGEMENT- IS AGENT-BASED SIMULATION A PANACEA?

Bizhan Jamshidnezhad [(a)], Alan Smith [(b)]

[(a), (b)] Department of Mechanical Engineering
University of Melbourne
Victoria 3010 Australia

[(a)]b.jam@pgrad.unimelb.edu.au, [(b)]ajrs@unimelb.edu.au

ABSTRACT
Quality Management (QM) problems are mostly researched by means of inductive reasoning based on statistics. Statistical models, albeit having several limitations, have been applied for a ling time in QM-related researches. Although some people present different definitions for QM and Total Quality Management (TQM), they are assumed to be the same in this research. Quality Management Systems (QMSs) are defined as a different notion later. This paper is an attempt to shed light on the alternative research methods to explore interactions between organizational performance and Quality Management (QM). To reach this objective, the problems of inductive reasoning, as the dominant research method, are categorized in three categories, namely research design, definitions and measurements, and analysis. The problems in each category are scrutinized and then the potentials of Agent-based Simulation (ABS), as new research method, are presented. It is argued that by applying ABS a virtual environment is provided to improve the quality research efforts in QM and deduce more reliable results.

Keywords: Quality, Performance, research method, Agent-based simulation

1. INTRODUCTION

Improving organizational performance has long been among significant activity for engineering managers. Performance is universally related to the efficient conversion of resources into products. It is important to managers because resources are scarce and local and international competition is severe and ever-increasing. Therefore, to overcome competitors and survive in the market, astute resource allocation and investment are of paramount importance to organizations in order to maximize their gains.

Improving quality has become a priority for some companies only in the last few decades of the twentieth century. One reason for the slow adoption of quality programs could be the widely-held belief in industry that improving quality would reduce productivity. However, global competition has forced companies to seek continuously competitive advantages and in this regard, quality initiatives may be helpful.

The relationships between quality management and performance have long been of interest of both management academics and practitioners. No matter what the exact definitions of quality and organizational performance are, it is common knowledge that quality and organizational performance are interrelated concepts; however there does not appear to be an explanatory model for clarifying the hidden causal relationships between these two important concepts. In modeling the relationship between performance and QM, in addition to productivity, other business performance measures like innovation, flexibility, financial performance, market share and competitiveness are discussed (Neely 1999, Neely et al 2000, 2004, Rathore et al 2005, Bititici and Turner 2000, White 1996). Within this framework, there are research works by which the effects of Total Quality Management (TQM) on financial performance (Wayhan and Balderson 2007, Sila 2005), or the impacts of competition on quality (Das et al 2000), or the effects of contextual factors (Sila 2007) or Just in Time (JIT) and ISO 9000 (Dreyfus, Ahire, and Ebrahimpour 2004) on TQM, are studied. A review of literature reveals that QM-related studies suffer from some fundamental defects in the research method itself and its implementation.

The aim of this paper is to outline major problems of research on QM-performance relationship. To alleviate the existing problems, a new methodology of research based on complexity theory is proposed. This methodology is an attempt to explore the potentials of complexity science in QM-related research. It begins with the description of TQM and Quality Management Systems (QMSs) followed by the problems associated with QM-performance. They are grouped in three categories: research design and approach, definitions and measurements, and analysis. Next, the outlooks of ABS as a new tool of conducting research are discussed.

2. PRINCIPLES OF TOTAL QUALITY MANAGEMENT

Quality draws the attention of economists, operation managers and marketing strategists, each with their own vantage point. Competitive pressures and technological developments have affected the core knowledge of operation and production management and introduced many new concepts in the literature. In fact, these developments are a response to managing the ever-increasing complexities associated with production and operation management. "The increased competition led to a proliferation of new inventions in how to manage for quality" (Juran 2004, page: 343). Feigenbaum (1991) categorizes the development of the quality field into *operator quality control, foreman quality control, inspection quality control, statistical quality control, Total Quality Control,* and *Total Quality Management (TQM).* It is seen that the scope of quality has been extending from a single workstation to the upper levels of management. Within TQM framework, it is suggested that quality be considered as a major business strategy. Gradually "concepts, methods, and tools accelerated the evolution of a new science- the science of managing for quality" (Juran 2004, page: 343).

Nowadays, Total Quality Management (TQM) and QMS are two terms that are used repeatedly and are presumed to be perceived by those who apply them, but especially this is not the case for TQM. Unlike QMS which has a systematic definition developed by International Organization for Standardization (ISO), there is no formal definition of TQM. However, there are numerous definitions, presented by different quality consulting organizations, academic scholars and quality practitioners. A review of different definitions shows that there is no consensus on the basic definitions of these buzzwords.

According to Harrington (1995), the term TQM was popularized in mid-1980s to label the extension of the quality discipline into all areas of an organization. In a TQM model presented by Oakland (2003), constituent elements of TQM are the so-called 4Ps including Planning, Performance, People, and Process along with three Cs- Culture, Communication and Commitment. Harrington (1995) defines TQM as a methodology dedicated to quality improvement efforts, whilst Cole and Mogab (1995, page: XI) consider TQM and continuous improvement as "different perspectives of the same phenomenon: a lean firm". The former's focal point is the beliefs and practices required whilst the latter acts as a change tool.

Feigenbaum (1991, page: 14) argues that TQM is an extension of Total Quality Control to cover the "full scope of the product and service life cycle from product conception through production and customer service" while a quality system "is a company-wide and plant-wide operating work structure, documented in effective, integrated technical and managerial procedures,…, to assure customer quality satisfaction and economical costs of quality" which is the foundation of Total Quality Control.

In another school of thought, TQM is defined, in a broader perspective, as a system approach to management rather than quality management. For instance, Stahl (1999) adopts this approach and states that TQM is targeted at continuously increasing value to customers by designing improving organizational processes and systems. TQM is a combination of a philosophy (e.g. compete on superior quality (Stahl 1999) and tools or a set of guiding principles (as Besterfield et al (2003) cite). ISO 8402:1994 contains the same definition of TQM as a management method: "a management approach of an organization, centered on quality, based on the participation of all of its members and aiming at long-term success through customer satisfaction and benefits to the members of the organization and to society". However, this viewpoint of TQM has vanished from the recent versions of ISO 9000 series (refer, for instance, to ISO 9001: 2008).

The key elements of TQM are: commitment and leadership of the top management, planning and organization, customer focus, using tools and techniques, education and training, workforce involvement, teamwork, measurement and feedback, continuous improvement culture, treating suppliers as partners (Oakland 2003, Besterfield et al 2003, Harrington and Harrington 1995). A part of researches about TQM include identification of different constructs and their interrelationships in order for conceptualizing TQM principles. The logical outcome of this conceptualization process seems to be developing casual models to relate TQM constructs to performance measures because the ultimate objective of management theory in general may be summarized as acquiring profit and competitiveness through organizational performance.

The same argument holds for QMS. ISO provides a rigorous definition of QMS and its elements. A QMS is a "management system to direct and control an organization with regard to quality" (AS/NZS ISO 9000:2006). A QMS along with other management sub-systems like the financial management system form the organization management system (ibid). In this view, QMS acts as a framework for continuous improvement, encouraging customer requirement analysis, process definition and process control with the aim of enhancing stakeholders' satisfaction (EN ISO 9000:2005). Acting as a stabilizer of quality efforts, a quality management system is a prerequisite to ensure consistent achievement of customer satisfaction and determines how organizations operate to institutionalize a continuous improvement culture throughout the company. In other words, QMS is the means for sustaining the improvement gains, a solid foundation of continuous improvement initiatives (Harrington and Harrington, 1995). A quality management system, in this viewpoint, is contemplated as the elementary step towards TQM (Stahl 1999, Besterfield et al 2003). Based on the several field studies, Dale et al (2007) reports the following benefits from a successful quality program:

- Increased productivity
- Better marketplace performance
- Improved business performance

Here again, the question is whether QMS is effective to produce desirable outcomes for organizations.

To develop a sound, operational model between quality management and organizational performance, an efficient methodology is needed, but inductive reasoning by means of statistical models falls short of expectations. Instead, Agent-based Modeling raises promising prospects that are discussed later in this paper.

The relationship between quality and performance could be analyzed in two layers: first quality as an outcome and second quality management system. Output quality is an outcome of quality management systems implemented all around the organization. Quality as an outcome cannot be achieved at random but by a well-established quality management system. If we consider output quality, it is directly related to performance because it could be construed as having lower defect rates and more usable products but the interaction of performance and quality management is not straightforward because discovering the effects of QM on organizations is a rigorous research due to the involvement of several interacting elements such as people, organization, technology and contextual factors. In the next section, the ubiquitous research method in QM, inductive reasoning, is critically analyzed.

3. RESEARCH METHODS IN QM AND THEIR DRAWBACKS

Based on the literature review, it seems that the majority of QM-related research is done by applying inductive reasoning. Different statistical methods like regression analysis, factor analysis, and structural equation modeling are used to analyze the data, mainly collected by questionnaires as the research instrument. Although inductive reasoning seems to be the only method for conducting research on QM-performance problems, it is not error-proof. In fact, some flaws are fundamental, yet hidden. The problems of current research on QM and performance relationship may be categorized into three groups: research design, definitions and measurements, and analysis. The details of each category are explained herein.

3.1. Research Design

Research design means planning the overall research. In QM literature, the dominant research method is induction based on statistics but it has several problems, discussed here.

The Drawbacks associated with induction reasoning in comparison to deduction is the first problem. Induction based on statistical methods is a dominant approach in QM-performance research, applied when deduction is very arduous (sometimes even impossible). In Bowell and Kemp (2002) words, "to say that an inference is an inductive inference is to say: (a) it is not deductively valid; (b) its premise is a generation about a sample of a given population, and (c) its conclusion is a generalization about the total population from which the sample is drawn". The main problem is that causal inductive arguments can provide us only probable conclusions (Vaughn 2005). Therefore, inductive conclusions must be viewed with caution (Powell 1995). Despite this weakness, it is still a useful tool in doing research in many fields because there is no other equally efficient alternative. In addition, there are major difficulties in modeling synergies and interactions of QM elements through statistical methods (Stahl 1999). In other words, interactions, for example between organizational culture changes due to TQM and performance, are nonlinear that is difficult to take into account with statistical models.

The second problem is lack of universally accepted theoretical foundation for QM. In the quality management literature, a considerable number of papers deal with basic quality concepts, which is argued to be because of a dearth of well-founded theory of quality management (Sousa 2003, Ghobadian and Speller 1994, Singh 2006, Foley 2000). In fact, QM look more like a collection of different methods and tools without any structure) and there is no theory to test it or expand it. Ghobadian and Speller (1994) report the lack of conceptual framework and sound instructional methodology as the main drawbacks of previous theories of quality. This idea is reiterated by Singh and Smith (2006) where other barriers like difficulties in successfully imitating quality management implementation and underestimation of the magnitude of change that is needed are noted.

The third problem is lack of originality in research on QM-performance relationship based on statistical models. There are many works in this area already done albeit with various methods and depth of research (see for example: Singh 2008, Dow et al 1999, Sila 2005, and Kaynak 2003). On the one hand, these similar researched may test the falsifiability of conclusions, which is a criterion for demarcation in Popperian philosophy (see Popper 2005 page: 17). Therefore, similar researches could be interpreted as a tool for strengthening evidence for relatively similar hypotheses. For instance, the effects of communication, information flow, and measurement are emphasized under different concepts in QM research, each addressing relatively a similar notion (see for example: Sila 2007 or Singh 2008). On the other hand, there has yet to be a theoretical system for QM, as it has been explained in this paper, without which falsifiability or verification makes no sense. It means that each author, as is a standard method in social science, defines his/her own conceptual model first and then tries to test some relevant hypotheses to prove the validity of the conceptual model. Therefore, similar research projects

with different conceptual models provide little evidence for a hoped-for general theory of QM.

Existence of external disturbance obscuring the links between QM and performance measures is the forth drawback. The hoped-for theory for QM-performance includes supposedly a set of cause-effect propositions but they cannot be deduced without deliberately controlling different variables that is very difficult in inductive reasoning. Existence of some important contextual parameters like organization culture or competition pressures exacerbates the situation (Das et al 2000, Sila 2007, Davies and Kochhar 2002, Stahl 1999).

Finally, the last problem is existence of tangible and intangible (sometimes behaviorally-oriented) factors of performance and TQM. Like QM, organizational performance is a multi-faceted concept that may be interpreted as a combination of output quantity and quality, flexibility and service-related indicators. Output could be intangible which makes it difficult to be quantified such as short order lead times, on-time deliveries, and high quality (Tangen 2005, Ray and Sahu 1990, Bernolak 1997, Yousif et al 1990, Jackson and Petersson 1999, Kaydos 1991). The performance indicators in era of mass production (e.g. early 20th century) sometimes criticized as being just an efficiency indicator, when the emphasis was put on more production, not contemplating intangible outputs and ignoring important issues like profitability, quality, on-time delivery and after-sale services. With the introduction of lean manufacturing concepts, customers have become more important than production itself. In other words, owing to market pressures and the shift from mass production to mass customization, traditional definitions of performance could not evaluate the organization success, which is related directly to the enterprise objectives and the quality of output as well. It means that traditional indicators like output to input ratio does not measure "*effectiveness*", another index used by scholars like Son (1994) in an attempt to alleviate the inherent exposed weakness of the traditional definitions. Nevertheless, contemplating intangible outputs in performance measurement models is not an straightforward job. This is also the case for QM with its behavioral factors.

To be successful, QM requires changing individual mindsets and the organization (Stahl 1999). In other words, some major changes in behaviors are inevitable but they cannot be captured and modeled within mathematics/statistics framework. This provides fresh impetus for more research on the feasibility of alternative research methods, discussed later in this paper.

3.2. Definitions and Measurements

The first requirement of an efficient research is basic definitions. Without definitions, nobody could define appropriate, measures. Both performance and QM suffer from fussiness of definitions and measures. This is explained in the following.

Lack of exact definitions of basic concepts like quality, QM, performance, etc is the first problem. There is no universally agreed definition for quality and it remains subjective, and easily misunderstood (Crosby 1996, Juran 1992, Deming 1986). This is also the case for performance (Tangen 2002, Bernolak 1997). The definitions of quality as "degree of goodness" and "the minimum number of conceptions required" were first suggested by Shewhart (1931). Despite the wide acceptance of simple definitions suggested by some quality gurus like "fitness for use" by Juran (1989) or "conformance to requirements" by Feigenbaum (1991), Garvin (1988) presented a thorough, multi-faceted definition of quality, based on what he called quality dimensions: performance, features, reliability, conformance, durability, serviceability, aesthetics and perceived quality. Although this definition is multi-faceted but it is difficult to define appropriate measures based on it.

Second, the differences between QM, TQM, quality awards, and business excellence models are unclear. There was a change in scope of quality initiatives from mere inspection in a workstation to company-wide quality management systems (QMSs) and further to Total Quality Management (TQM). Unlike a QMS, which is a management system to direct and control an organization with regard to quality (AS/NZS ISO 9000:2006), TQM has no formal definition. However, there are numerous suggestions, presented by different quality consulting organizations, academic scholars and quality practitioners (Harrington 1995, Oakland 2003, Cole and Mogab 1995, Stahl 1999, Besterfield 2003). This vagueness of differentiation hinders developing a consistent mental model among managers, consultants as well as academics and is reflected in biased, erroneous surveys

Third, there are several difficulties in data collection. The main problem of data collection in TQM researches is the difficulty in determining whether a company has implemented TQM (and to what extent?) (Davies and Kochhar 2002). There is no straightforward guidance to differentiate between a quality-oriented organization and an ordinary one. Data collection methods vary from mail survey and direct observation to in-depth interview. The danger of biased data always exists; as Sila (2005), Davies and Kochhar (2002), and Stahl (1999) report, respondents are mainly quality managers and like case studies, report positive results; furthermore, direct observation and in-depth interview are time-consuming and produce qualitative data. The tendency to try `to give the "right answers" not the "actual answers" and to exaggerate TQM programs by managers (regarding the time and the scope of implementation) are other sources of bias (ibid). Data accuracy, sample size, and measurements are other sources of challenge in data gathering. Firstly, data inaccuracy does exist, because many studies rely on self-reported data from CEOs (based on their own perceptions) that cannot consider all variables and assess any causal linkage between TQM and

performance. Secondly, in statistical models, the sample size is of paramount importance in obtaining reliable results but a review of previous researches shows that respondent rates are low (often less than 25%). Consequently, sample sizes are small (Dow et al 1999, Singh and Smith 2006). Thirdly, serious measurement problems exist in even standardized indexes of firm performance like market share, profitability and productivity (Stahl 1999). They are insubstantial notions that exist only as concepts with fuzzy definitions. The fuzziness makes it hard to develop an appropriate, measurement yardstick. Most surveys lack operational definition of the terms they use and thus are ambiguous (Stahl 1999).

3.3. Analysis

By analysis, we mean the final stage of research in which the hypotheses are tested or research questions are answered. Some of the most important pitfalls in this category are explained in this section.

The first problem is the existing controversies concerning whether or not TQM has any tangible impact on performance. Because of its importance in companies surviving international competition, the notion of quality and implementation initiatives like QMSs or TQM have attracted the attention of both practitioners and academics and thus, a large number of publications in production and operation management have been devoted to these subjects. Studies of Holoviak (1995), Sebastianelli and Tamimi (2002), Garvin (1988), Seawright and Young (1996), Oh (1995), Reeves and Bendar (1994) and Zairi (2002) are examples of works focusing on the conceptual aspects of quality management. Although there are a very large number of publications supporting the effectiveness of quality initiatives, some scholars like Longbottom et al (2006) and Dayton (2003) raise doubts about the future of TQM. However, the reasons for TQM failure are unknown (Terzioski 2006, Sila 2007, and Singh and Smith 2006). This could, in part, be attributed to difficulties in building causal models between QM and performance (Dow et al 1999). Findings about QM and organization performance are sometimes contradictory and confusing because, on the one hand, most academics and consultants encourage applying QM principles and tools based on the tested hypothesis of a positive link between QM and performance (e.g. Dow 1999 or Wayhan and Balderson 2007). On the other hand, popular press report negative findings and the failure rate of 60-67% (Dow et al 1999, Sila 2007).

Temporal issues between short-term and long-term results are another source of problem. As it was explained before, QM includes some changes in the altitudes, behaviors, and working culture. They do not happen easily in a short time but evolve within a relatively lengthy time span. Therefore, their effects on the organizational performance variables take time to appear. In other words, researching in this area needs longitudinal analysis not snap-shot one. The major problem turns out to be the start date of TQM deployment which is ambiguous (Sila 2005, Davies and Kochhar 2002). Furthermore, sometimes TQM is implemented just in a section of a company and the performance data are not available or cannot be compared with other companies (Davies and Kochhar 2002, Stahl 1999).

The outcome of aforementioned drawbacks is contentious results, seen in the literature and, to ameliorate these problems, it seems that research in QM needs alternative methods. The problems associated with statistical models in QM-Performance relationship, uncovered by means of literature review have posed the first question of this research: is there any research method better than inductive reasoning? If yes, how can QM be researched by means of this new method? It is argued that ABS could play the role of an innovative approach to tackle the problems associated with inductive reasoning that are discussed in the next section.

4. THE PROMISING OUTLOOK OF AGENT-BASED SIMULATION

The ultimate goal of the research on QM-performance studies is to illuminate the causal relationship between QM and performance in a coherent model that preferably prescribes appropriate actions for establishing a quality system in organizations. There are two candidates: deduction and simulation. Deductive reasoning based on mathematics is a static model that either prescribes an optimized solution (e.g. linear programming) or describes a system (e.g. queuing models) based on a set of assumptions but, as Davies and Kochhar (2002) quoted, “in an environment as complex as manufacturing, variables do not remain constant” and thus the complexity and dynamism of QM related problems hinder application of such static models. In other words, there is no appropriate mathematical model to capture the associated complexities of QM. To avoid the pitfalls of induction and deduction, we require a pseudo-experimental environment to control noise and bias-inducing factors and develop abstract models and, as Axelrod (2005) names it the third way of doing science, simulation seems to be an effective tool in this regard. This might be drawn from the widespread application of simulation in natural and physical sciences. “Simulation comes into its own when the phenomena to be studied are either not directly accessible or difficult to observe directly” (Goldspink 2002). It is contemplated as a safe laboratory for testing hypothesis and making predictions (Dooley 2002).

Computer simulation has been emerged as the main part of system modeling especially after conspicuous development of computers. “A simulation model may be considered as a set of rules (e.g. equations, flow charts, state machines, cellular automata) that define how the system being modeled will change in the future, given its present state” (Borshchev and Filippov 2004). Dooley (2002) categorizes simulation in three groups: System

Dynamics (SD), Discrete Event Simulation (DES), and Agent Based Simulation (ABS). The details of each method are not covered in this paper, because they are not within its objectives. However, it is worth noting that SD is equation-based simulation which applies differential equations to express the changes of aggregate variables over time but DES and ABS are entity-based in which the behavior of systems are produced by interaction among system elements. ABS and equation based modeling (EBM) differ in fundamental relationship among entities and the level at which they focus their attention (Parunak et al 1998). In EBM, the relationships between system components are modeled as a set of variables, representing the overall outcome of one-by-one interactions among entities (with restrictive assumptions like equilibrium, constancy of structure, and independence of entity behaviors (Dooley 2002)), and the focus is shifted toward the system performance as a whole. In contrast, the main focus of an ABS is agent-to-agent interactions constituting system variables. In addition, there is no adaptation in EBM (because of the independence assumption), whereas agents could alter their behaviors. There are also other differences that Parunak et al. (1998) summarize as follows:

1. Certain system behaviors (e.g. adaptation) are very difficult (sometimes impossible) to be modeled by mathematical formalism, whereas, within ABS framework, they are easier.
2. ABSs make it easier to distinguish physical space from interaction space, because they "permit the definition of arbitrary topologies for the interaction of agents".
3. ABSs offer an additional level of validation, because you can validate an ABS model both in macro (system) and micro (agent) levels.
4. ABSs support more direct experimentation, because it is easier to examine various what-if scenarios in an ABS than translating them into equations in an EBM.
5. ABSs are easier to translate back into practice, because they are expressed in terms of behaviors that could be simply transferred into the real world.

In addition, individuals are homogenous in EBM but heterogeneous in ABS and equations representing relationships among variables and observables are output of ABM but input for EBM (ibid). It means that opposed to statistical modeling or differential equations, there is no need to make assumptions like linearity, homogeneity, normality and stability in ABS (Bankes 2002). This is not a unique characteristic of ABS but a property of Complex Adaptive Systems (CASs) in general and offers ABS a considerable flexibility to examine problems, freed from within mathematical/statistical modeling framework.

The theory of CAS underlies the principles of ABS. "CASs are systems that have a large number of components, often called agents that interact, adapt or learn" (Holland 2006). If we contemplate QM as a method for enabling organizational change, improvement could be interpreted as an outcome of multiple interacting people. In CAS or complexity theory parlance, organizational performance is an "emergent" behavior (Lissack 2000). "ABS excels at relating the heterogeneous behavior of agents with different information, different decision rules and different situations to the macro behavior of the overall systems" (Lempert 2002). "Agent-based models usually emphasize change in agents schema via learning and adaptation, and also highlight the phenomena of emergent, self-organizing patterns in complex organizational systems" (Dooley 2002).

The key difference between DES and ABS is in entity properties; while entities are passive in DES, they are active and more autonomous in ABS. "The fundamental feature of an agent is the capability of the component to make independent decisions" (Macal and North 2006). Gilbert (2008) describes the features of agents as autonomy, social ability to interact, reactivity to stimuli from environment, and goal-directed proactivity. He adds that agents apply bounded rationality to find local optimums which is very important to find solution for complex problems as is the case for QM-related problems. In Figure 1, the important features of an agent are depicted (Macal and North 2006).

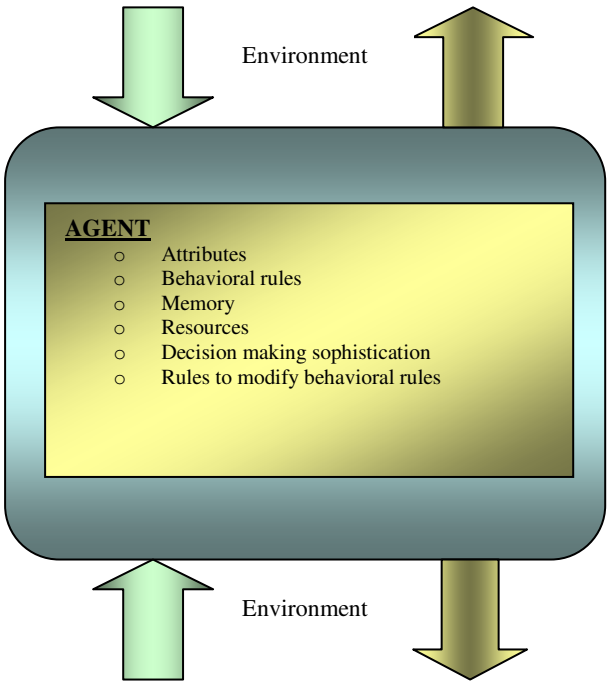


Figure 1: Feature of an agent

Aside from fundamental differences between an entity and an agent, they differ in modeling and

execution. In the words of Nilsson and Darley (2006), "firstly the execution of the agents is based on internal rules not on external and global policies… the focus is on the agent and its adaptiveness within the system being studied… and finally compared to DES models, agent-based models are mostly time driven". These differences make ABS suitable for modeling systems with changing elements. What Macal and North (2006) express as the reasons for using ABS are present in QM-performance researches as well:

1. There are employees, teams, and departments in an organization, naturally representing agents at different levels. The bottom-up approach embedded in ABS is consistent with execution of most organizational processes.
2. In an organization, there are events, decisions, and behaviors related to QM that can be defined discretely (e.g. customer complaints or quality defects). In other words, management of quality is done in response to discrete events by means of processes such that each are composed of discrete events as well.
3. Organizations at different levels adapt and change their behavior, suggesting they are made up of adaptive agents.
4. In organizations, behaviors at micro levels are summed as the organizational performance at the macro level. In ABS terms, these are "emergent" phenomena.
5. There is learning in the organizations (e.g. regarding quality problems)
6. In the organizations, the behaviors and interactions are heterogeneous.
7. Randomness could be embedded in ABS models of organizations
8. In CAS, every element is connected to each other (partially or completely) in a feedback loop, therefore, there is no equilibrium point because changes in any agent might influence negatively or positively other agents and the process of self-organization and emergence due to interactions, to be successful, requires some energy in an open system (Nilsson and Darley 2006). Similarly, organizations are open systems, interacting with their environment; business processes are increasingly nonlinear, self-organizing, changing and rationally bounded (ibid). There is no equilibrium state for quality management initiatives, because, within total quality philosophy, striving for perfection is a permanent task. In addition, QMS like other organizational systems is affected by the environment. Therefore, as McKelvey (1999) and Lissack (2000) conceptualize the basics, application of complexity theory (in our case ABS) to organization science seems plausible.

Borshchev and Filippov (2004) note that agent based modeling is more general and powerful and easier to maintain than other simulation techniques even "in the absence of the knowledge about the global interdependencies". As was mentioned before, ABS are more suitable especially for systems containing active objects like people, business units, vehicles or projects with a type of individual behavior (ibid). Therefore ABS could be an appropriate substitute for the currently pervasive inductive reasoning.

Simulation models of a single phenomenon/system could be various and is largely dependent on modeling objectives. The objective of simulation in general, and ABS in particular, could be prediction, performance, training, entertainment, education, proof, and discovery (Axelrod 2005). Prediction and discovery of hidden causal relationships between QM practices and organization performance and proof of some of general, related hypothesis could be among the research objectives by means of ABS.

5. CONCLUSION

The conflicting results of QM-performance researches suggest that something is wrong with the research method. It is argued that current research method based on inductive reasoning could not capture the complexities of QM principles that are mostly related to the behavior of people, units, processes and departments of an organization. In addition, it is difficult to model synergies, learning and working culture changes with mathematical/statistical models, contemplating that they are always present in QM initiatives. The characteristics of ABS, founded on the principles of CAS theory, make it an appropriate tool for modeling adaptive behaviors. It is argued that there is room for improving research efforts by applying ABS.

REFERENCES

AS/NZS ISO 9000: 2006, Quality management systems-Fundamentals and vocabulary.

Axelrod, R., 2005. Advancing the Art of Simulation in the Social Sciences. In J. P. Rennard, *Handbook of research on Nature Inspired Computing for Economy and Management*. Hersey, PA: Idea Group.

Bankes, S.C., 2002. Agent –based modeling: A revolution?. *Proceedings of the National Academy of Sciences,* pp. 7199-7200. May 14, National Academy of Science.

Bernolak, Imre, 1997. Effective measurement and successful elements of company productivity: the basis of competitiveness and world prosperity. *International Journal of Production Economics,* Vol. 52, 203-213.

Besterfield, D.H., Besterfield, M.C., Besterfield, G.H., Besterfield, S.M., 2003. *Total Quality Managemnet*, third edition, Upper Saddle River, NJ : Prentice Hall.

Bititici, U.S., Turner, T., 2000. Dynamics of performance measurement systems. *International*

Journal of Operations and Production Management, 20 (6), 692-704.

Borshchev, A., Filippov, A., 2004. From system dynamics and discrete event to practical agent based modeling: Reasons, techniques, tools. *Proceedings of 22nd International Conference of the System Dynamics Society*, July 25-29, Oxford, England

Cole, W.E., Mogab, J.W., 1995. *The Economics of Total Quality Management*, Cambridge, MA: Blackwell Business.

Crosby, P.B., 1996. *Quality is still free*. New York, McGraw-Hill.

Das, A., Handfield, R.B., Calantone, R.J., Ghosh, S., 2000. A Contingent View of Quality Management-The Impact of International Competition on Quality. *Decision Sciences*, 31 (3), 649-690.

Dayton, N.A., 2003. The Demise of Total Quality Management. *The TQM Magazine*, 15 (6), 391-396.

Davies, A.J., Kochhar, A.K., 2002. Manufacturing Best practice and Performance Studies: a Critique. *International Journal of Operations and Production Management*. 22 (3), 289-305.

Dooley, K., 2002. Simulation Research Methods. In: J. A. Baum, ed. *Companion to organizations: An introduction. In: The Blackwell Companion to Organizations.* Blackwell, 829-848.

Deming, W.E., 1986. *Out of the Crisis*, published by MIT press.

Dow, D., Samson, D., Ford, S., 1999. Exploding the myth: do all quality management practices contribute to superior quality performance. *Production and Operation Management*, 8 (1), 1-27.

Dreyfus, L.P., Ahire, S.L., Ebrahimpour, M., 2004. The Impact of Just-In-Time Implementation and ISO 9000 Certification on Total Quality Management. *IEEE Transactions on Engineering Management*, 51 (2), 125-141.

Feigenbaum, A.V., 1991, *Total quality control*, third edition. McGraw-Hill, Inc.

Foley, K., 2000. Whither Quality Management. *Seventh National Research Conference on Quality Management*. Available from cmqr.rmit.edu.au.

Garvin, D.A., 1988. *Managing quality*. New York: The Free Press, a division of Macmillan Inc.

Ghobadian, Abby, Speller, Simon, 1994. Gurus of quality: a framework for comparison. *TOTAL QUALITY MANAGEMENT*, 5 (3), 53-69.

Gilbert, N., 2008. agent-based models (Quantitative Applications in the Social Sciences). Los Angeles: Sage Publications, Inc.

Goldspink, C., 2002. Methodological Implications of Complex Systems Approaches to Sociality: Simulation as a Foundation for Knowledge. Journal of Artificial Societies and Social Simulation, 5 (1), available from http://jasse.soc.surrey.ac.uk/5/1/3.html.

Harrington, H.J., Harrington J.S., 1995. *Total Improvement Management*, New York: McGraw-Hill.

Holland, J.H., 2006, studying complex adaptive systems. *Journal of Systems Science and Complexity*, 19, 1-8.

Holoviak, S.J., 1995. Why TQM Fails to Change Behaviors or Attitudes. *Journal for Quality and Participation*. 18 (4), 86-89.

Jackson, M., Petersson, P., 1999. Productivity – an overall measure of competitiveness. *Proceedings of the 2nd Workshop on Intelligent Manufacturing Systems*, pp. 573-81. September 22-24, Leuven.

Juran, J.M., 1989. *Juran on Leadership for Quality*. New York: The Free Press, A division of Macmillan, Inc.

Juran, J.M., 1992. *Juran on Quality by Design*. New York: The Free Press.

Juran, J.M., 2004. *Architect of Quality: the Autobiography of Joseph M. Juran*. New York: McGraw-Hill

Kaydos, W., 1991. *Measuring, Managing and Maximizing Performance*. New York: Productivity Press

Lempert, R., 2002. Agent-based modeling as organizational and public policy simulators. *PNAS*, 99 (3), 7195-7196.

Lissack, M.R., 2000. Complexity: the Science, its Vocabulary, and its Relation to Organizations *Emergence*. 1 (1), 110-126.

Longbottom, D., Osseo-Asare, A.E., Chourides, P., Murphy, W.D., 2006. Real Quality: Does the Future of TQM Depend on Internal Marketing. *Total Quality Management*, 17 (6), 709-732.

Macal, C.M., North M.J., 2006. Tutorial on agent-based modeling and simulation part 2: how to model with agents. In: *WSC '06: Proceedings of the 38th Winter Simulation Conference*, pp. 73–83.

McKelvey, B., 1999. Complexity Theory in Organization Science: Seizing the Promise or Becoming a Fad. *Emergence*, 1 (1), 5-32.

Neely, A., 1999. The performance measurement revolution: Why now and what next. *International Journal of Operations and Production Management,* 19 (2), 205-228.

Neely, A., Mills J., Platts K., Richards H., Gregory M., Bourne M., Kennerley M., 2000. Performance Measurement System Design: Developing and Testing a Process-based Approach. *International Journal of Operations and Production Management,* 20 (10), 1119-1145.

Neely, A., 2004. The Challenge of Performance Measurement. *Management Decision*, 42 (8), 1017-1023.

Nilsson, F., Darley, V., 2006. On Complex Adaptive Systems and Agent-Based Modeling for Improving Decision-Making in Manufacturing and Logistics Settings. *International Journal of Operations and Production Management,* 26 (12), 1351-1373.

Oakland, J.S., 2003. *Total Quality Management*, third edition, Oxford ; Burlington, MA: Butterworth-Heinemann.

Oh, H.L., 1995. A Changing Paradigm in Quality. *IEEE Transactions on Reliability*, 44 (2), 265-70.

Parunak, S.R., Riolo R.L., 1998. Agent-based modeling vs. equation-based modeling: A case study and users' guide. In: *Lecture Notes In Computer Science*. pp. 10–25, Springer-Verlag London, UK.

Popper, K.R., 2005. *The logic of scientific discovery*. sixth impression, London: Hutchinson and Co.

Powell, T.C., 1995. Total quality management as competitive advantage: a review and empirical study. *Strategic Management Journal*, 16, 15-37.

Rathore, A., Mohanty, R.P., Lyons, A. C., Barlow, N., 2005. *International Journal of Advanced manufacturing Technologies*, 25, 1020 1028.

Ray, P.K., Sahu, S., 1990. Productivity measurement through multi-criteria decision making. *Engineering Costs and Production Economics*, 20, 151-163.

Reeves, A.C., Bednar, D.A., 1994. Defining quality: alternatives and implications. *Academy of Management Review*, 19 (3), 419-445.

Seawright, K.W., Young, S.T., 1996. A Quality Definition Continuum. *Interfaces*, 26 (3), 107-113.

Sebastianelli, R., Tamimi, N., 2002. How product quality dimensions relate to defining quality. *International Journal of Quality and Reliability Management*, 19, (4), 442-453.

Shewhart, W.A., 1931. *Economic Control of Quality of Manufactured Product*. New York:Van Nostrand Company Inc.

Sila, I., Ebrahimpour, M., 2005. Critical Linkage among TQM Factors and Business Results. *International Journal of Operations and Production Management*, 25 (11), 1123-1155.

Sila, I., 2007. Examining the Effects of Contextual Factors on TQM and Performance through the Lens of Organizational Theories: an Empirical Study. *Journal of Operations Management*, 25, 83-109.

Singh, P.J., Smith A., 2006. Uncovering the Faultlines in Quality Management, *Total Quality Management*, 17 (3), 395-407.

Singh, P.J., 2008. Empirical assessment of ISO 9000 related management practices and performance relationships. *International Journal of production Economics*, 113, 40-59.

Stahl, M.J., 1999. *Perspectives in Total Quality*, MA: Blackwell Publishers Inc.

Son, Y.K., 1994, Linear Programming-based Productivity and business strategy analyses in the factory automation era. *International Journal of Production Research*, 32 (2), 443-455.

Tangen, S., 2002. Understanding the concept of productivity. *Proceedings of the 7th Asia Pacific Industrial Engineering and Management Systems Conference* (APIEMS20020), Taipei.

Tangen, S., 2004. Performance Measurement: from Philosophy to Practice. *International Journal of Productivity and Performance Management*, 53 (8), 726-737.

Tangen, Stefan, 2005. Demystifying Productivity and performance. *International Journal of Productivity and Performance*, 54 (10), 34-46.

Vaughn, L., 2005. *The power of critical thinking: effective reasoning about ordinary and extraordinary claims*. New York: Oxford University Press,

Wayhan V.B., Balderson E.L., 2007. TQM and Financial Performance: what has Empirical Research Discovered. *Total Quality Management*, 18 (4), 403-412.

White G.P., 1996. A Survey and Taxonomy of Strategy-related performance measures for Manufacturing. *International Journal of Operations and Production Management*, 16 (3), 42-61.

Zairi M., 2002. Beyond TQM Implementation: the New Paradigm of TQM Sustainability. *Total Quality Management*, 13 (8), 1161-72.

AUTHORS BIOGRAPHY

Bizhan Jamshidnezhad is a PhD candidate in Department of Mechanical Engineering. His current research is about analysis of quality management in manufacturing systems by means of agent-based modeling. Application of Axiomatic Design in lean production system design is another area of his interest.

Alan Smith has been an academic in the Department of Mechanical Engineering since 1982. He gained his PhD in the field of thermal analysis of machining. Since then he has divided his research interests between various aspects of the machining of metals and quality management. Alan also has supervised projects in the optimization of assembly lines, business process improvement and technology transfer. The thrust of his work has been on the prediction and quantification of various aspects of manufacturing industry processes. Twenty-eight higher degree by research (masters and PhD) students have been successfully supervised by Alan. His teaching duties cover several areas of engineering management such as economics, organizational behavior and quality management, in addition to manufacturing processes.

RETAIL STORES MODELING: COMPARISON BETWEEN SIMULATION AND VIRTUAL REALITY APPROACHES

Francesco Longo[(a)], Marina Massei[(b)], Enrico Papoff[(a)], Sabrina Rondinelli[(a)]

[(a)]Modeling & Simulation Center, Laboratory of Enterprise Solutions (MSC-LES)
Mechanical Department, University of Calabria
Via P. Bucci, Cubo 44C, 87036 Rende, Italy
f.longo@unical.it, e.papoff@unical.it, srondinelli@unical.it – URL www.msc-les.org

[(b)] Simulation Team
Via Molinero 1, 17100 Savona, Italy
marina.massei@simulationteam.com – URL: www.simulationteam.com

ABSTRACT
In large scale retail supply chain, stores are quite complex systems involving different actors (i.e. managers, customers, distribution centers, suppliers, etc.), each one with specific needs and problems (i.e. store layout, products assortment and inventory management, shelves display policies, marketing and promotional activities, customers' satisfaction, etc.). In this research work the authors describe the development of two architectures for virtual retail stores based on the integration of simulation and three-dimensional virtual environments. The first architecture simulates the evolution over the time of a real store and provides the user with different functionalities (including a 3D visualization). The second architecture, based on the X3D standard (an evolution of VRML), prevalently supports the online commerce in a 3D virtual environment. The paper presents and compares both the architectures, highlighting the main functionalities, differences, advantages, disadvantages and potentialities.

Keywords: logistics and retail stores, simulation, modeling, 3D virtual environments.

1. INTRODUCTION

Retail stores are part of a complex supply chain (large scale retail supply chain) that includes, among others, stores, customers, distribution centers, suppliers, production plants, retail companies headquarters (that usually suggest economic and operative objectives and strategic policies) and other external actors such as banks, insurance companies, State, etc. Therefore retail stores decisions are usually addressed to achieve the business results (decided at headquarters level) respecting needs and requirements of different stakeholders.

A retail store managers deals with different problems that usually require quick decision processes. Among others, the most important are personnel management, marketing management and promotional activities, items inventory management, space allocation on the shelves (items display), e-commerce and virtual stores.

Let us consider the human resources management: alienation and low performances due to repetitive jobs as well as workers turnaround among different stores departments strongly affect retail store efficiency. Store managers usually apply tasks enrichment and enlargement policies in order to obtain major workers (and customers) satisfaction enhancing, at the same time, working conditions. Store workload planning, refurbishment frequencies and refurbishment quantities for each item depend on store departments conditions. Furthermore, some store sectors are more critical than others in terms of frequency, quantity and items display (i.e. the general goods sector involves a huge number of items types and usually requires a greater number of workers). Workload planning requires accurate and real time overviews of the store conditions and the use of specific policies to optimize workers' assignment to store departments. In this case expected results are: down time elimination and higher customers' satisfaction levels.

Consider now marketing activities and decisions; among others, the following elements play a critical role for increasing customers' satisfaction: kindness and assistance within the store, parking facilities comfort, store layout, attractiveness, promotions and discounts, products quality, opening hours and products assortment (Ortega B. 2000; Piotrowski C. and Rogers E. A. 1998).

Retails store managers also deal, every day, with inventory management problems: they have to set correctly the space on the shelves for each item type and select the re-order policies to be used. The inventory management policy should monitor the following parameters: the cost for sale losses (total and for each department), the storage costs, the costs for shelves

refurbishment as well as the revenues from effective sales.

Difficulties and problems also come up during data collection phase for store performance analysis. It is required to collect data about customers number, revenues from effective sales, loss sales, quantities ordered to suppliers, etc. Each store department is considered as a cost center with specific costs, such as employees, maintenance, hygiene and cleaning service. A correct data collection allows to calculate the economic report for the entire retail store as well as for each department. Moreover it is possible to analyze data for department-period, sector-period, family-period, product-period and make different statistic analysis to be used for improving promotional and marketing policies, work planning, logistics activities and, in general, the retail store management.

One of the most recently business for a retail store is the e-commerce. On line retail stores allows customers to interact, order and buy or simply have information about goods. This new form of business involves new critical issues for retail stores, such as information and service quality (Ahn and Ryu, 2007), design of the website, shopping navigation functionalities, (Liang and Lai, 2002), security and privacy of online transactions (Liao and Cheung, 2001). Note that a Modeling & Simulation (M&S) based approach is a powerful methodology for supporting the decision process within retail stores. Different tools based on M&S have been developed to support managers in solving and taking decisions about human resources management, logistic problems, inventory management, e-commerce, etc.

In this research work the authors present and describe the development of two virtual retail stores based on the integration of Simulation and three-dimensional virtual environments. The authors propose two different architectures for recreating a virtual store. The first architecture is capable of simulating the evolution over the time of a real store and provides the user with different functionalities (above all devoted to support the store decision process). The second architecture is based on the X3D standard (based on XML and developed as evolution of VRML) and its functionalities prevalently support the online commerce. The main goal of the paper is to compare both the architectures, highlighting main functionalities, differences, advantages, disadvantages and potentialities.

Before getting into details of the work, in the sequel a brief overview of the paper sections is presented. In Section 2 a state of the art about the use of M&S and Virtual Environment in retail stores design and management is presented. Section 3 proposes the first architecture called *Virtual Dynamic Store* (VDS). Section 4 presents and describes the second architecture called *Shopping Online CReated by Advances Three-dimensional Environments and Simulation* (SOCRATES). Section 5 compares the two virtual stores architectures and finally the last section summarizes conclusions and research activities still ongoing.

2. MODELING & SIMULATION AS SUPPORT TOOL IN THE LARGE SCALE RETAIL SECTOR

As mentioned into the introduction a retail store manager every day has to take in short time critical decisions that strongly affect the store efficiency. The authors believe that a Modeling & Simulation (M&S) based approach can effectively support retail stores design and management.

Consider human resources management within retail stores. Bruzzone et al. (2008) propose an innovative integrated works island model for retail stores: the model uses simulation for evaluating workload and number of workers required in each island (store sector) and it has been implemented in 14 retail stores of a leader brand in the Italian retail business.

Simulation tools are also used for supporting strategic decisions, such as the evaluation of different store management and logistics and distribution alternatives (Bruzzone et al. 2000).

Nielsen (1993) say that effective marketing strategies positively affect customers' satisfaction. Usually elements such as stochastic customers' behavior, inadequate data sources and uncertain market response make marketing strategies planning an exceeding difficult task. Also in this case M&S, by recreating the stochastic evolution of the market is an effective support for marketing strategies development and testing (some application examples of M&S for supporting marketing strategies are reported in Arinze and Burton, 1992; Bruzzone et al., 2000-c).

Moreover M&S is extensively used in retail stores for items inventory management: De Sensi et al. (2008) use M&S for testing different inventory control policies under demand and lead time constraints; Al-Rifai and Rossetti (2007) use M&S for calculating optimal inventory levels while Bertazza et al. (2005) calculates inventory levels that minimize the total costs in a vendor managed inventory system. Additional research works can be found in Chen and Krass (2001), Giannoccaro and Pontrandolfo (2002) Lee and Billington (1993) Lee and Wu (2006).

One of the most recent application fields for M&S (combined with three-dimensional visualization and environments) is the e-commerce (Briano et al., 2004). The main idea is to use simulation for supporting store management and a virtual three dimensional environment for supporting both the store design and management and provide the user with an advanced and interactive online shopping experience (virtual stores). In this last case a virtual store has to recreate exactly the sensation to be inside a real store (Briano et al., 2004; Gong et al., 2004).

In this research work two different architectures to recreating a virtual retail store are presented: the first one is based on the integration of different software

tools (including 3D real time modeling tools, virtual environments and Visual C++) and the second one developed by using X3D standard, a description language of 3D scenes. X3D is an evolution of the VRML standard for web 3D scenes description.

The VRML standard has been extensively used in different domain of applications: from network resources management (Deri and Manikis, 1997) to reconstruction of 3D human organ in the medical field (Zheng-yang, 2005), from visualization of videoconference image sequences (Kompatsiaris and Strintzis, 2006), to design and implementation of virtual shopping systems (Lu et al., 2005).

In the next sections two architectures for recreating a virtual retail store (respectively called VDS and SOCRATES) are presented and described.

3. THE VIRTUAL DYNAMICS STORE

Some years ago, a prototype of a virtual retail store was developed by the authors (Briano et al., 2004). Recently the authors start to work again on the virtual store and implemented a fully functional version called VDS (*Virtual Dynamic Store*). The different steps to implement the VDS are as follows: creation of the 3D store layout, determination of the retail store areas involved in the simulation, definition, design and implementation of the virtual store functionalities (including analytical models combined with genetic algorithms for items inventory management and items optimal space allocation on the shelves). The VDS is able to recreate the evolution of the retail store over the time by using three different integrated modules: the Virtual Environment, the store Geometric Model and the Dynamic Processes Module. The Virtual Environment recreates the real environment that hosts the retail store; the Geometric Model realistically represents the store layout (including building, shelves, warehouses, items, etc.); the Dynamic Processes Module recreates all the interactions that evolve in the hosting environment and modify the store configuration as the time goes by (interactions generated by store workers and customers). Note that the Dynamic Process Module has been developed in Microsoft Visual C++. As a consequence, in order to guarantee a complete integration between the Dynamic Processes Module and the other two modules, the authors decided to use *Vega Prime* by *Presagis* for implementing the Virtual Environment (it provide a full interface with C++) and the modeling toolset *Creator* by *Presagis* (that can be easily interfaced with Vega Prime) for implementing the store Geometric Model.

The VDS recreates all the most important activities of a real retail store: (i) creation and arrangement of different items on the store shelves and items realistic representation; shelves emptying/filling process; (ii) inventory management and items optimal space allocation; (iii) store monitoring and controlling by real-time simulation; (iv) web virtual shop functionality for e-commerce.

To reproduce correctly the retail store activities, it was necessary to collect a set of information about shelves positions and dimensions, items positions and overall dimensions, area allocated on the shelves for each item type, items prices, items on hand inventory and consumption flows over the time.

Such information have been used for developing both the store Geometric Model and the Dynamic Processes Module. The Geometric Model includes 3D shelves (including classic shelves, refrigerators for fresh food, checks-out, etc.) with a dual graphic detail level. The implementation of a dual graphic detail level is strictly connected with the computational workload of the computer graphic card. A real store contains thousands of items that mean thousands of geometric models (i.e. a parallelepiped for each product). Each item geometric model should have minimum one texture mapped on it (in order to have a realistic item representation). Consequently, the computational workload could easily exceed the graphics card capabilities of a low cost hardware platform. To avoid computational overloads, high resolution and low resolution graphic detail levels have been implemented for all store sectors. The levels are activated (activating one excludes the other) according to whether the observer is close or far from the object being observed. In particular, each sector of the store Geometric Model has a dual graphic detail level in order to lighten the computational workload of the graphic card. Bounding boxes define the portions of space – within which the shelves of interest are located – whose confine, if crossed by the observer from inside to outside, leads to a switch from one texture for each item to a single texture for the entire shelf being observed.

The DPM recreates the items geometric models, disposes the items on the shelves and executes the emptying/filling process. The DPM recreates the shelves emptying process during the store business hours: by considering items sale rates, the DPM calculates dynamic stochastic sale flows based on daily/hourly rates. Analogously the DPM manages, outside the business hours, the shelves refurbishment process by considering the inventory costs, profits and optimal space allocation for each item. The items inventory management model and the approach proposed for items optimal space allocation are respectively based on Order-Point, Order-Up-to-Level (s, S) inventory policy and genetic algorithms.

Concerning the store controlling by real time simulation, the DPM has a specific function that modifies the number of items on the shelves. As soon as an item passes through the store checkouts (in the real store), the DPM updates the VDS animation. Therefore, the VDS is capable of monitoring, real time, the store giving an overview of the items inventory on the shelves by using a three-dimensional visualization.

Another additional aspect managed by the DPM regards the virtual shop for e-commerce: as soon as the customer (i.e. a person that use the virtual shop through internet) click on a specific item, the DPM creates a

message box reporting all the information regarding the item: brand, price, best before dates, ingredients, etc. Figure 1 shows a panoramic view of the VDS, figure 2 shows the shelves emptying process managed by the DPM.

Figure 1: panoramic view of the virtual retail store

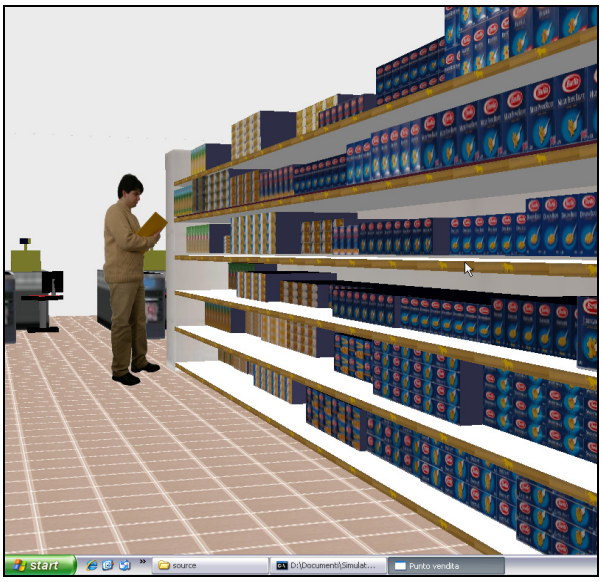


Figure 2: Shelves emptying process managed by the DPM

The DPM also implements two different types of movements: the first one reproduces the customers' movement among the shelves (the movement is controlled by an external input, for instance the keyboard), the second one is similar to the flight of a bird to have a general overview of the retail store.

SHOPPING ON LINE CREATED BY ADVANCED THREE-DIMENSIONAL ENVIRONMENT AND SIMULATION

The second architecture proposed by authors for developing a virtual retail store is based on X3D standard and it is called SOCRATES (*Shopping On Line CReated by Advanced Three-dimensional Environment and Simulation*).

SOCRATES is an useful application to support the e-commerce, indeed it is a virtual retail store in which users can navigate, interact with objects and eventually buy the products.

As for the VDS, the first step is the development of the retail store 3D environment. Data about store layout dimensions, departments number and dimensions, shelves number and dimensions, items position and overall dimensions, items position on the shelves, space allocated for each item type on the shelves, items prices, items general information (i.e. items description, ingredients, best before date, etc) have been collected. Figure 3 shows some information about the store layout, shelves dimensions and disposition.

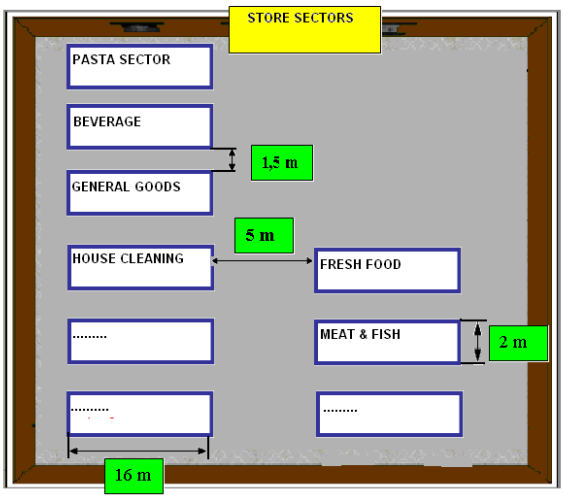


Figure 3: Departments layout and dimensions

The SOCRATES geometric model has been developed by using SketchUp, a Google tool to model and develop 3D environment. The 3D geometric models include the building that host the retail store, shelves, refrigerators, check outs, etc.. As last part of the geometric model development, items have been positioned on the shelves as shown in figure 4.

Figure 4: Virtual shelves with products

The second step after the SOCRATES geometric model creation, was to set-up SOCRATES interactivity. The store geometric model is exported in VRML format

and successively edited by using X3D Vitality Studio editor files. X3D allows users to create the interactivity with the virtual retail store, to position cameras (a cameras positioned on customer overhead represents the starting point for the navigation) and to set light and shadow.

In particular the interactivity is created by using special X3D structures (Anchor node) that allow to link each product with a pop-up web page reporting all the information about the product and the possibility to add the product to a cart. The pop-up web page is activated simply by clicking on the item as shown in figure 5 and 6.

Figure 5: Click on product

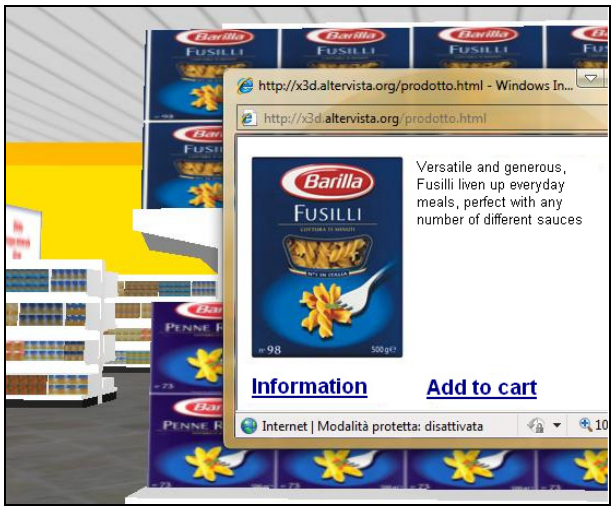


Figure 6: Information Page

The last functionality implemented in SOCRATES is the movement through the virtual environment. To this end the authors use BS Contact (by Bitmanagement Software), a player and a browser web plug-in that allows web visualization of real-time Virtual Reality. BS Contact provides the user with the following functionalities:

- Navigation speed: it possible to select five different navigation speeds, from very slow to very fast;
- Movement in the virtual environment: it is possible to walk, fly, examine and jump in the scene;
- Graphic: this option allow user to set some aspects of the visualization, (i.e. lights position);
- Look my Avatar: if user selects this option, an avatar is visualized within the virtual environment. The user can control avatar movements by using external inputs (i.e. keyboard, mouse).

The software also provides additional functionalities such as volume control, full screen modality, 3D scene saving, etc.

Figure 7 shows the navigation in the virtual environment without avatar, while figure 8 shows the navigation with avatar functionality activated.

Figure 7: Navigation even avatar

Figure 8: Navigation with avatar

An interesting aspect of SOCRATES is the frame recreation. In particular to recreate a frame in X3D, the following steps are required:

- Frame development. To produce a frame the following elements are necessary: a 3D scene and objects, lights and cameras positioned on the scene;
- Imagine visualization;
- External input or simulation;
- Scene modification and camera repositioning;

It is a circular process, repeated about 16 times per second.

The interactivity in X3D it is implemented by using sensor nodes, that acquire users' input and provide appropriate responses (through events):

- Touch Sensor: supervise the actions of pointing device (PD);
- Plane Sensor: convert PD movement on XP plane into objects movement on 3D scene;
- CylinderSensor and SphereSensor: convert PD movement on XP plane to rotate objects in 3D scene;
- KeySensor: detect input by keyboard.

Therefore, the sensors detect different users' actions and produce events, that are send to others nodes. The PD (Pointing Device) in the case of SOCRATES is the mouse, but in general it could be a touch screen, a joystick, a game pad, etc.

The nodes allow to implement complex functions, such as store information, complex animation, geometries generate by algorithms. In this cases Script Nodes have to be used; for such nodes the behavior can be modeled by using java code. The script node contains different script function (java method). The function takes the input (the first argument of the function) from an event, then executes an algorithm and returns new events that modify the scene.

In SOCRATES the script nodes have been used to link the different goods to specific web page that report all the information about the product. Then, the customer can navigate within SOCRATES and by clicking on the item with the PD, can activate different events (information, add to cart, etc.).

4. VDS AND SOCRATES COMPARISON

The previous sections provide the user with a description of the VDS and SOCRATES, two different architectures developed by authors for implementing a virtual retail store. In the present section, the two architectures are compared and analyzed to point out advantages and disadvantages.

The first difference between the VDS and SOCRATES is in their scope. In effect, even if both architectures provide the user with a three-dimensional visualization of the retail store layout, the VDS has to be regarded as a decision support tool for retail store design and management, while SOCRATES is more devoted to shopping on-line and e-commerce functionalities.

The VDS simulator (implemented within the DPM) recreates the main activities of the retail store: shelves emptying/filling process, shelves refurbishment, items optimal space allocation on the shelves, real time store controlling for store departments workload planning. Moreover the DPM implements specific inventory control policies and genetic algorithms for optimal space allocation of items on the shelves. The VDS can also be used as support tool for store layout optimization: in fact the 3D visualization allows user to test and analyze different store layout (from grid layout to free form layout.

The second architecture, SOCRATES, recreates with high accuracy the retail store environment, the layout, the shelves disposition, the products display and allows users to interact with the 3D environment, to select products and obtain information or to buy products. Then SOCRATES is mostly devoted to e-commerce.

To provide SOCRATES with additional potentials in terms of web application, the authors decided to develop SOCRATES by using VRML language. The VRML is a standard language for internet applications; VRML files are machine readable independently from the platform (each user can visualize SOCRATES without technical problems). On the contrary, the VDS can be put on-line (i.e. by a remote desktop) but latency problems can easily occur.

However both VDS and SOCRATES present some disadvantages. The VDS scalability could be a problem in the case of very large number of items (low performances on low-cost hardware platforms). Some SOCRATES disadvantages are related to get used to interact with a virtual environment by using 2D tools (mouse or keyboard). Another problem could be the time for downloading the VRML file. A single VRML file contains the description of the entire 3D scene, then if the scene is complex the dimension file will grows as well as the downloading time (the VDS presents the same problems if the user has to download the VDS databases before using it). In both cases care must be taken when implementing the 3D geometric models (i.e. by using different graphic detail levels).

5. CONCLUSIONS

In the present research work two different architecture for virtual retail store have been presented. In the first part of the paper, the most important retail store problems have been described and then a state of the art about Modeling and Simulation and Virtual Environment tools to solve these problems have been presented.

In the second part of the paper the two architectures, respectively called VDS and SOCRATES have been presented and described. VDS is an advanced tool for supporting retail store design and management. SOCRATES, based on VRML standard successively edited by using X3D standard, is mostly devoted to support e-commerce. Some advantages and disadvantages of both architectures are finally presented and discussed.

REFERENCES

Ahn, Tony, Seewon Ryu and Ingoo Han (2007). "The impact of Web quality and playfulness on user acceptance of online retailing". *Information & Management*, 44(April): 263–275.

Al-Rifai, M.H. and Rossetti, M.D. (2007) 'An efficient heuristic optimization algorithm for a two-echelon

(R, Q) inventory system', Production Economics, Vol. 109, pp.195–213.
Arinze B. and Burton J. A simulation model for industrial marketing. Article provided by Elsevier in its journal Omega. Volume 20, year 1992, Issue 3, Month May, Pages 323-335.
Bertazzi, L., Paletta, G. and Speranza, M.G. (2005) 'Minimizing the total cost in an integrated vendor-managed inventory system', Journal of Euristics, Vol. 11, pp.393–419.
Briano C., Bruzzone A.G., Longo F., Papoff E., Viazzo S., Simulation and virtual reality applied to modelling retail and store facilities, Proceedings of SCSC, July 25th – 29th 2004, S. Jose, California, USA.
Bruzzone A.G. Giribone P., Revetria R., Brandolini M., Briano C. (2000-a) " Modelling & Identification as a Decision Support System for Retail Chains", Proc.of MIC, Innsbruck, February;
Bruzzone A.G., Mosca R., Briano C., Brandolini M. (2000c) "Models for Supporting Customer Satisfaction in Retail Chains", Proceedings of HMS2000, Portofino, October 5-7;
Bruzzone A.G., Poggi S., Bocca E., Longo F., Madeo F., Rondinelli S., Workload Forecast Algorithm optimization for re-organizing retail network, Proc. of MAS 2008, Campora S. Giovanni (CS), Italy;
Chen, F. and Krass, D. (2001) 'Inventory models with minimal service level constraints', Operational Research, Vol. 134, pp.120–140.
Corstjens J. , Corstjens M. (1999) "Store Wars : The Battle for Mindspace and Shelfspace", John Wiley and Sons, NYC;
Curcio, D., Longo F., 2009. Inventory and Internal Logistics Management as Critical Factors Affecting the Supply Chain Performances. *International Journal of Simulation & Process Modelling*, 5(2), 127-137.
De Sensi, G., Longo, F. and Mirabelli, G. (2008) 'Inventory policies analysis under demand patterns and lead times constraints in a real supply chain', International Journal of Production Research, Vol. 46, pp.6997–7016.
Deri L., Manikis D. VRML: Adding 3D to Network Management. Intelligence in Services and Networks: Technology for Cooperative Competition, Volume 1238/1997, pages 469-477, 1997.
Giannoccaro, I. and Pontrandolfo, P. (2002) 'Inventory management in supply chains: a reinforcement learning approach', Production Economics, Vol. 78, pp.153–161.
Gong J., Li W., Li Y., Huang M., Network Collaborative Virtual Geographic Environments: Design and Implementation, Geoscience and Remote Sensing Symposium, 2004, Volume 7, Issue , 20-24 Sept. 2004 Page(s):4866 - 4869 vol.7.
Gousty Yvon (1998) "Le Genie Industriel", Presse Universitaire, Paris;
Hatch M.L. (1997) "Organization Theory: Modern, Symbolic, and Postmodern Perspectives", Oxford Press, Cambridge;
Kompatsiaris I. and Strintzis M. G.. 3D Representation of Videoconference Image Sequences Using VRML. Lecture Notes in Computer Science, Volume 1425/1998, pages 1-12, 2006.
Lee, H. and Billington, C. (1993) 'Material management in decentralized supply chains', Operations Research, Vol. 41, No. 5, pp.835–847.
Lee, H.T. and Wu, J.C. (2006) 'A study on inventory replenishment policies in a two-echelon supply chain system', Computers and Industrial Engineering, Vol. 51, pp.257–263.
Liang, Ting Peng and Hung-Jen Lai (2002). "Effect of store design on consumer purchases: an empirical study of on-line bookstores". *Information & Management*, 39(May): 431-444.
Liao, Ziqi and Michael Tow Cheung (2001). "Internet-based e-shopping and consumer attitudes: an empirical study". *Information & Management*, 38(April): 299-306.
Longo, F., Mirabelli, G., 2008. An Advanced supply chain management tool based on modeling and simulation. *Computer and Industrial Engineering*, 54/3: 570-588.
Lozinsky S. 1998 "Enterprise-Wide Software Solution", Addison Wesley, Reading;
Lu Y., Bing, X., Qingge J. Zhigeng P., Hongwei Y., 2005. Design and implementation of a collaborative virtual shopping system. IEEE International Conference Volume 1, Issue , 29 July-1 Aug. 2005 Page(s): 315 - 320 Vol. 1.
Monks J.G. (1996) "Operations Management", McGraw Hill, NYC;
Nielsen Marketing Research (1993) "Category Management" McGraw-Hill Trade;
Ortega B. (2000) "In Sam we Trust", Times Books;
Piotrowski C. Rogers E.A. (1998) "Designing Commercial Interiors", John Wiley & Sons, NYC;
Reznikoff S. C. (1989) "Specifications for Commercial Interiors", Whitney Library of Design;
Varley R., Gillooley D. (2001) "Retail Product Management: Buying and Merchandising", Routledge.
YU Zheng-yang, ZHENG Shu-sen, CHEN Lei-ting, HE Xiao-qian, WANG Jian-jun. Dynamic concision for three-dimensional reconstruction of human organ built with virtual reality modeling language (VRML). Journal of Zhejiang University - Science B, Issue Volume 6, Number 7 / July, 2005, pages 611-616.

AUTHORS BIOGRAPHY

FRANCESCO LONGO took the degree in Mechanical Engineering from University of Calabria (2002) and the PhD in Industrial Engineering (2005). He is currently scientific responsible of the MSC-LES (Modeling & Simulation Center – Laboratory of Enterprise Solutions). He is also Associate Editors of the international journal Simulation: Transaction of

SCS. He supports the organization of the most important conferences on M&S (SCSC, EMSS, MAS, etc.). His research interests include modeling & simulation of manufacturing systems and supply chain management, supply chain vulnerability, resilience and security, DOE, ANOVA.

MARINA MASSEI completed her thesis in Genoa University, Humanistic Area on File Structures and Data Warehousing. She had attended Postgraduate courses in Project Management, Modelling and Simulation, Operational Management, Total Quality Management, Logistics. She operates with the simulation team in DIPTEM, University of Genoa as a Project Controller. She is involved in the organisation of International Scientific Events and she is a Member of the International Programme Committee of major M&S Conferences. She is an Associate Director of the MISS Center located in Perugia University.

ENRICO PAPOFF was born in Naples (Italy) on February the 03rd, 1948. He took the degree in Mechanical Engineering from University of Napoli Federico II, in 1973. He is currently Associate Professor at the Mechanical Department (Industrial Engineering Section) of the University of Calabria. His research interests regard project management and business plans and Modeling & Simulation applied to production and logistics systems.

SABRINA RONDINELLI was born in Tropea (Italy) in May the 9th, 1984. She took her degree in Management Engineering, summa cum Laude, in September 2008 from the University of Calabria. She is currently PhD student at the Mechanical Department of University of Calabria. Her research activities concern the development of simulation models, based on general purpose programming languages and artificial intelligence techniques, for production and Supply Chain planning and management.

SIMULATING THE INTEGRATION OF ORIGINAL EQUIPMENT MANUFACTURER AND SUPPLIERS IN FRACTAL ENVIRONMENT

Sameh M. Saad[(a)] and Julian C. Aririguzo[(b)]

[(a)(b)] Department of Engineering and Mathematics,
Sheffield Hallam University, City campus Howard street Sheffield S1 1WB

[(a)]S.Saad@shu.ac.uk, [(b)]Julian.C.Aririguzo@student.ac.uk

ABSTRACT
Partnerships and collaborations between companies (OEMs) and suppliers are not new. Many companies rely on partnerships with key suppliers to improve operational effectiveness through minimized inventory, information and culture integration to boost their lean/ agile credentials. The Fractal Manufacturing Partnership (FMP) is new manufacturing strategy, whereby OEMs form relationship with key suppliers. The former cede autonomy of non-core activities to tried and tested suppliers, while concentrating on their core competencies. The latter become assemblers of their components with heightened sense of responsibility while co-owning the OEMs' facility. The objective of this relationship is maximization of flexibility - ability to respond swiftly and robustly to changes in environment, requiring physical network and more importantly, cultural network linking the people rather than their machines. In this paper, integration of OEM and suppliers is modeled and simulated to highlight critical factors in this partnership and quantifying and harnessing benefits of this new approach.

Keywords: OEM and Supplier partnerships, FMP, Supply chain managements.

1. INTRODUCTION

Partnerships and close relationships between OEMs and key suppliers and customers are not new. OEMs increasingly outsource the manufacture of auto parts and this purchasing practice not only affect direct costs, but also impact quality, lead-time, technology, over head costs and most importantly, market success (Cross and Gordon 1995; Lewis *et al*. 1993). Many companies especially in the automotive industries rely extensively on important partnerships with key, time tested suppliers. It has been established that the cost of purchased parts and products make up to 30% to 50% of the final selling price of finished product (depending on the firm's vertical integration strategy) (Dyer *et al*. 1998; Dyer 1996; Cross and Gordon 1995). Close collaboration with suppliers from initial product design to final assembly, reduces product development time, manufacturing expense and improves quality (Noori and Lee 2000; Lewis *et al*. 1993). This logical and more recent progression from single sourcing has been the development of long-term supply agreements (LTSAs) between OEMs and their key suppliers. The partnership is marked by great motivation and synergism and requires cooperation, commitment, trust, teamwork and information sharing between parties and complete integration of parties involved to facilitate effective product lunches and competitive pricing (Simonian 1996; Cross and Gordon 1995). FMP is a revolutionized manufacturing method whereby OEMs go into close relationship with their key suppliers. Conceptually from the fractal system, it elevates the operation of sub-factory within a factory and enhances close links within members. This practice is necessitated by swift technological developments and by the need to take cost out of their operations. Companies examine their internal strengths, focusing their efforts towards achieving excellence in their core capabilities (Noori and Lee 2000; Dyer *et al*. 1998). These trusted suppliers then take responsibility for non-core activities. They design, manufacture, and assemble their parts on the assembly line directly to the product while sharing and co-owning the OEMs' facility. In the case of automotive companies, the OEM concentrate on the vehicle concept which includes envelop size and weight and assembly, relinquishing parts and components that have been undertaken by them in the past to trusted suppliers in a long term relationship (Cross and Gordon 1995). An increasing shift to modular component purchasing e.g. seats, belts, instruments panel and headliner may be integrated into an interior module that is undertaken by a single supplier - such as a tier-one supplier (Dorrell 1996). This results in fewer, but larger tier-one suppliers that are taking responsibility for the system design, development, assembly and management of the supply chain (Simonian 1996). OEMs need to consider which core competencies they are maintaining and which ones they will need for the future and ensure that sufficient investment in these continue. Given the long lead-term in development, failure to invest in a key area now may make it difficult later. However, de-integrating certain functions out of the organization does not have benefits for the OEM, instead capital investment requirements ,

operational costs and the logistical costs of maintaining product balances are all transferred to the supplier, while flexibility and the ability to concentrate on core competencies is enhanced (Cross and Gordon 1995). FMP is designed to maximize the logistical attribute of a lean production system and configured to provide strategic merging of engineering network capabilities (Phelan 1996). It combines logistical attributes of lean production methods with strategic configuration of agile network capabilities (Dyer *et al.* 1998; Phelan 1996; Noori and Lee 2000). The organizational structure of the FMP is based on series of highly coordinated production silos arranged side by side to each other to promote high degree of cooperation, communication and integration of operation and managerial activities, culminating in further reduction in work in process (WIP) inventory and instantaneous communication amongst parties involved. The communication and 'open book' information system present allows complete flexibility and an information enriched manufacturing atmosphere. There is also better service and product quality especially when suppliers feel part of the team. The degree of integration between OEM and these key suppliers is of great significance. Studies carried out by (Dyer 1996; Dyer *et al*. 1998; Lewis *et al.* 1993; and Cross and Gordon 1995) highlighted that this integration leads to improved operational effectiveness through reduced inventory, improved communication, quality, faster product development, design for manufacture and productivity. All parties must weigh the costs against the relative benefits in establishing their integration policy. Cost, control, communication, organizational climate, operations management and competitive differentiation must be analyzed exhaustively (Dyer *et al.* 1998; Cross and Gordon 1995). It is imperative to point out and highlight how OEM - supplier partnerships have evolved in recent years from an arms length relationship - just-in-time or bulk delivery, JIT (11) (Issacson 1994), through modular sequencing (Dinsdale 1996) and supplier parks (Feast 1997; Kochan 1996) to a 'hands on', proximate FMP (Friedland 1996; McElroy 1996). As the evolution progresses, there is increased responsibility on the part of the supplier for design, assembly, higher value added contribution and increased integration. However, FMP has both higher degree of integration as well as complex supplier responsibility. The focus of this paper is the determination of an optimal representation of the FMP. This facilitates achievement of flexibility and swift response to uncertainties in the manufacturing environment, the realization of a host of other benefits as listed in (Noori and Lee 2000) and most importantly a harmonious cultural and technological integration of the parties involved in the long-term FMP relationship. However, culture integration, union philosophy that is resistant to radical changes and costs all pose a challenge in implementation of the FMP configuration. The rest of the paper is organized as follows; section two details the methodology employed for the modeling and experimentation and software used in the study. Section three elaborates on the model development, including the tricks and turns involved in such exercises. Section four studies and discusses the results and section five concludes the paper.

2. METHODOLOGY

A comprehensive computational representation of the FMP is made using modeling and simulation. This aids in evaluating its performance in dynamic conditions. The structure, resources, behavior, strategic objective, values and constraints of the FMP is captured in Arena software (Kelton *et al.* 1998) through enterprise design, analysis, and operation. Understanding of the nature and working of FMP before conducting statistical experiments is also crucial in the final results of the modeling. The output data of the simulation is used to identify system bottlenecks and to generate alternative states that may provide the desired performance improvements for the system. Arena is designed to describe, model and analyze an existing or proposed application accurately and gives maximum flexibility to systems. It integrates all simulation related functions; animations, input data analysis, model verification, and output analysis into a single simulation modeling environment (Kelton *et al.* 1998). Its flexible flowcharting objects will be used in this project to capture the essence of the FMP system being considered and compare different competing manufacturing scenarios, so as to select one that best meets the objectives. Visual Basic for applications (VBA) is a technology used to write custom program codes that argument Arena model logic. VBA is embedded directly in Arena to enable writing codes (via the visual basic editor) that automate other applications such as excel, auto cad or Visio. VBA code will be used in this project to automate Arena, such as to get values of a simulation output statistics, change values of module Operands or add animation variables (Kelton *et al*., 2004). Opt quest for Arena is an optimization tool and will be used to analyze the results of the simulation runs. It includes sampling techniques and advanced error control to find better answers faster (Rathmell *et al.* 2002). This package combines the metaheuristics of Tabu search, neural networks, and scatter search into a single, composite search algorithm to provide maximum efficiency in identifying new scenarios (Kelton *et al*., 2004; Kelton *et al*. 1998). Finally the Arena Output analyzer will be used in fitting confidence intervals on expected output performance measures, and statistical comparison of alternatives (Sweet and Grace-Martin 2003). These applications will be used in; (i) building and developing a virtual scenario for the proposed FMP. (ii) finding the best fit and balance for the OEM/ supplier partnership to ensure a harmonious collaboration (iii) calculating the best mix of resource capacities to maximize throughput in the integration of lean production / agile network capabilities (iv) finding the optimal balance for the system in a volatile environment while meeting the conceptual benefits of the FMP. An organized set of procedures and guidelines

are used for specifying the structural and quantitative parameters and relationship between the factors affecting the output performance. These factors are varied systematically with a view to finding and identifying the optimal conditions that most influence the results. Important variables are identified and investigated. These are defined, measured and controlled during the simulation with a view to tracking their level of variation.

3. MODEL DEVELOPMENT

3.1. System Description

The system under studies is a truck assembly plant. To keep things simple, only major modular components have been represented in this model. In total there are eight sub-models that represent eight distinct operative activities. These include; Body in white, Chassis Trim supplier, Motor Engine builder, Electrical / Electronics supplier, Motor Transmission supplier, Paint supplier/ shop, OEM (Dealership) Inspection, and the Exit logic (see figure 1). As mentioned earlier, these suppliers have been vested with the responsibility of designing, building and assembling their modular components in close proximity to the OEM's assembly line. The suppliers rent production silos side by side to each other on the assembly line in a highly coordinated arrangement. The layout of the FMP assembly line allows complete flexibility in its operation and essentially shows the physical link with the different suppliers involved in this partnership. The OEM concerns with the brand concept which includes the envelope size and the weight of the finished truck, and is fully represented on the assembly line, eyeballing these different suppliers and supervising the overall assembly process.

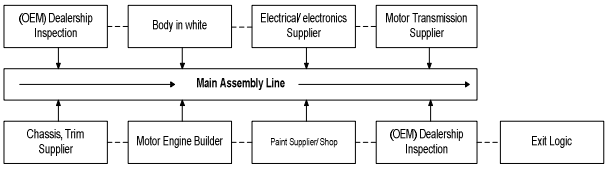


Figure 1: FMP assembly line (Noori and Lee 2000)

3.2. Sub-factory within a factory

The FMP operates on the conceptual philosophy of the Fractal Manufacturing system (Ryu and Jung 2003 & 2004). The fractal is an independent acting corporate entity whose goals and performance can be described precisely (Warnecke 1993). The idea of 'assembly within assembly' is applicable to organizational structuring of distributed manufacturing systems (Shin *et al*. 2008). (Strauss and Hummel 1995) in their work on industrial engineering, say that a fractal is a partial system of an enterprise which offers opportunities for entrepreneurship to all employers, and it has a relation with other fractal units as a service centre. Each fractal is a customer as well as a supplier within the enterprise, and plays the role of an individual service centre within other service centre, i.e. 'a design within design' or 'pattern within pattern'. Each business unit of the factory acts as an autonomous factory which is integrated within a communication network (Sihn and Von 1999). In the FMP, the suppliers are incorporated as assemblers, working within the manufacturing facility alongside the OEMs' employees. Every fractal unit has or is inherently equipped with the fractal specific characteristics. This include; self-similarity, self-organization, self optimization, goal orientation and dynamics (Warnecke 1993). These are congenital attributes of fractals.

3.3. Decentralized hierarchical structure

The fractal structure is characterized by constant evolution with respect to its partners and environment (Tharumarajah *et al.* 1996). The administrative functions in the FMP are distributed over a less concentrated area, decentralizing the structure and highlighting the evolution from a vertically integrated enterprise to a network of integrated core competencies (Noori and Lee 2000). This structure is subject to a constant dynamical process of change, making them more suitable and adaptable to turbulent environment. It is also more flexible because it is susceptible to modification or adaptation and more responsive to change. Every fractal in the FMP has the same functional modules which are well-defined interfaces to the other components. In terms of job processing, this is carried out through the goal-formation process. Component relationship also exists, whereby there is a coordinative higher fractal and an active lower fractal. The fractal model manages the structural complexity and coordination of a flexible manufacturing system by maximizing local functionality and minimizing global control (Tirpak *et al.* 1992).

3.4. Modeling of FMP

The top-level model for the layout of the FMP is shown in figure two. The system to be modeled is a truck assembly facility. Shots of 'body-in-white', dealership (OEM) inspection and paint shop sub-models during the simulation have been included in figures three, four, and five respectively. It consists of part arrivals, manufacturing cells with different machines and part departures, eight major sub-factories represented by sub-models located adjacent to each other. The suppliers design, build and assemble their modular components while residing on highly coordinated production silos. This representation not only allows flexibility and ease of organization but also shows the physical link with the participants. Transfer of parts and components is by a loop conveyor system following the concept of pre-defined entity-dependent sequences. The time between a part's arrival and that of the next part is called inter-arrival time of parts. The assembly operation starts at the 'body in white' sub-model where the metal frame arrives and within which threads and supports, doors, hoods and deck lids are assembled. On completion, this is transported by the loop conveyor to the chassis, trim supplier where seats, upholstery and

windshield are coupled. After undergoing a quality check this is conveyed to the electrical and electronics supplier where the electrical aspects of the assembly operation are done, including the airbags and sensors. The motor engine builder is next on the assembly line, and he mounts the engine which was pre-built at his sub-factory. The transmission supplier follows, and here both the gear box and crank case are assembled and coupled on, followed by elaborate greasing of different movable parts. From this sub-model, nearly completed truck is conveyed to the paint shop which is manned by the paint supply who organizes the priming, initial coating and finishing of the painting. Trucks that pass the painting quality check proceeds to the Dealership (OEM) inspection. Here there is continuous eyeballing of the entire assembly progression and trucks undergo an elaborate inspection for overall envelop size and weight. There is also room for rework for trucks that don't make it through the inspection. This final inspection rolls the fully built truck off the loop conveyor and production line. All process times (the time a part spends processing in a particular cell) are triangularly distributed, inter-arrival times between successive parts arrival are exponential distributions. Load and unload times unto the loop conveyor are 2 minutes each. Information is considered from the output performance measure of 10 statistically independent and identically distributed (IID) replications, of length 480 hours, to study the system's average Work in Process (WIP) and to get statistics on the system's behavior, utilization and turnarounds. Statistics is gathered from the long run (steady state) behavior of the system, hence there is a warm-up period of 240 hours to clear the statistical accumulators from

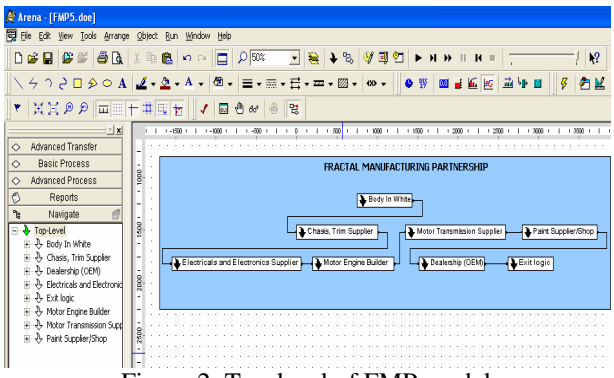


Figure 2: Top-level of FMP model

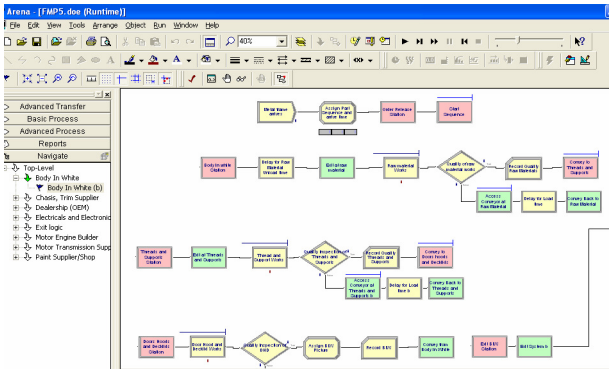


Figure 3: The 'Body-in-White sub-factory

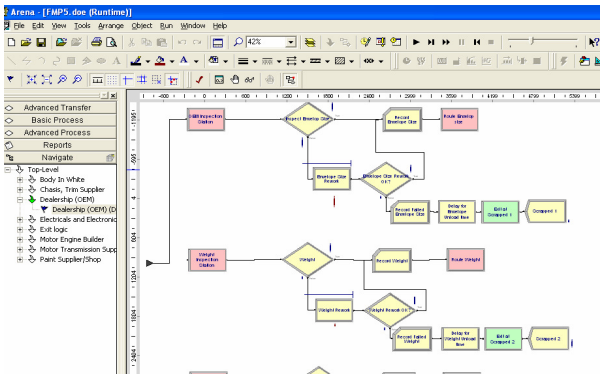


Figure 4: Dealership (OEM) Inspection

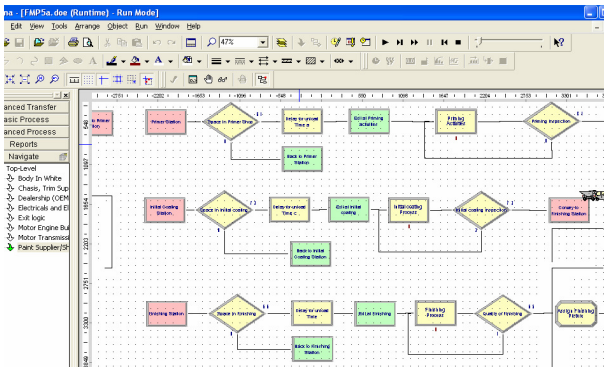


Figure 5: Paint Supplier/Shop sub-factory (taken at simulation time, 1243 minutes)

biasing initial conditions. The steady state is tracked from the plot of the curve of WIP vs. Time when the effect of the empty-and-idle initial conditions appear to settle or wear out. The base time unit is in minutes. We will be interested in collecting statistics in each area on resource utilization, number in queue, time in queue, and the cycle time (total time in system). The work in process (WIP) inventory is very important to the FMP. Obviously the OEM eyeballs the entire assembly process, but to establish a single overall output performance measure for the WIP, we tracked the history of the WIP over time and summed this for the individual activities in the different sub-models and found an average. We also created an entry (Figure 6) labeled Total WIP in the statistic module which shows in the category overview as 'user defined', giving the time average and maximum of the total number of parts processing in the system.

Statistic - Advanced Process

	Name	Type	Expression	Report Label	Output File	Categories
1	Total WIP	Time-Persistent	EntitiesWIP(Metal Frame)+EntitiesWIP(Seat)+EntitiesWIP(Upholstery)+EntitiesWIP(	Total WIP	Total WIP History.dat	0 rows

Figure 6: The Total WIP Entry in the statistical data module

The significance of this is to show the compatibility of the different partners and their activities and the harmony in their intra- and inter- operations. The model has taken into account the similarity requirement in organization and orientation of different sub-assemblies present. This has been built from bottom up. The sub-factories are similarly organized both internally and in their goal system. Similarity of goals means conformity of objectives in each organizational unit to the overall corporate goal (Shin *et al*. 2008; Sihn and Briel 1997).

3.5. Model verification and validation

The validity of the developed simulation model was evaluated by comparing the performance of the model to the conceptual system. Separate experiments were carried out, to investigate how robust the system is and how it can recover from uncertainties like equipment breakdown and unforeseen delays from sub-factories. The output values obtained (Tables 1, 2, &3) from the simulation model were found to be very similar to the estimated values of the conceptual system, differing at most 7.9%. Therefore the tests are suitable for system analysis and experimentation. The Output Analyzer provides one way of quantifying the imprecision in the parameter estimates through a 95% confidence interval. This is achieved by forming intervals with endpoints that cover the target with high probability. Half width of the output performance is the half width of a (nominal) 95% confidence interval on the expected value of the output result. These resulted in reliable and precise statistical conclusions.

3.6. Model Debugging

The model of the FMP is a particularly large model, going into great depth on the lower-level modelling constructs as well as correspondingly detailed statistical requirements, comprising essentially eight sub-models as has been established. The sub-models were ran separately for a start and huge amount of time was spent debugging the model and making sure that it runs without errors.

4. EXPERIMENTATIONS AND RESULTS

The fractal concept advocates adaptability and the ability of the system to recover from failures and uncertainties. To study this capability of the system, we looked at three key scenarios. First we observed the system under normal conditions. Then we watched how the system managed without grinding to a halt to cope or adapt when;

- there is surge/ drop in demand of the product.
- when a machine breaks down or there is delay in meeting with a pre-scheduled operation in a sub-factory.

We managed the practical mechanics of making the model changes for these alternatives, and that involves lots of parameter changes in the model especially the process times for different machines at different sub-factories. These model variants from changes in the model's input parameters were ran in an efficient and organized way using Arena Output Analyzer.

4.1. Output Statistics

Table 1: Surge in demand

Performance measures	Conveyor velocity			
	15 Feet/m	20 Feet/m	25 Feet/m	30 Feet/m
Throughput	834	844	857	867
Cycle time	20708.75	20699.53	20679.86	20676.19
WIP	224880.80	224902.85	224889.93	224923.00
Scheduled Utilization	0.700	0.699	0.700	0.701
Wait time in queue	19909.46	19906.51	19897.09	19901.97
Number in queue	6307.01	6306.24	6308.10	6306.88

Table 2: Drop in demand

Performance measures	Conveyor velocity			
	15 Feet/m	20 Feet/m	25 Feet/m	30 Feet/m
Throughput	190	192	191	190
Cycle time	20832.66	20916.64	20939.78	20904.20
WIP	83981.26	83873.78	83823.29	83860.25
Scheduled Utilization	0.626	0.624	0.625	0.625
Wait time in queue	17434.44	17579.11	17587.30	17491.06
Number in queue	2194.61	2192.99	2192.35	2192.25

Table 3: Equipment breakdown in sub-factory

Performance measures	Conveyor velocity

	15 Feet/m	20 Feet/m	25 Feet/m	30 Feet/m
Throughput	823	827	829	835
Cycle time (mins)	19646.98	19632.93	19639.06	19621.39
WIP	112667.61	112657.91	112642.00	112654.70
Scheduled Utilization	0.685	0.685	0.685	0.685
Wait time in queue	18578.52	18567.86	18575.01	18564.19
Number in queue	2977.84	2977.77	2977.35	2977.93

4.2. Discussions

Comparing different versions or alternatives of FMP model, there isn't huge differences in the output statistics between different replications. What makes the alternatives differ more significantly is more of a fundamental change in logic rather than simple parameter variations. During a surge in demand, the number of trucks produced (Figure 7), after 480 hour long replication, increased directly with increase in conveyor velocity and peaks at 867 trucks for conveyor velocity of 30 Feet/minutes.

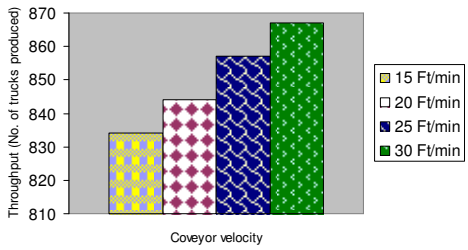


Figure 7: Average number of trucks produced during a surge in demand

Conversely, the average cycle time i.e. the total time parts spend servicing in system (figure 8) dropped with increase in conveyor velocity. This figure was maximum at just above 20708 minutes at velocity, 15 Ft/min and least at about 20676 minutes.

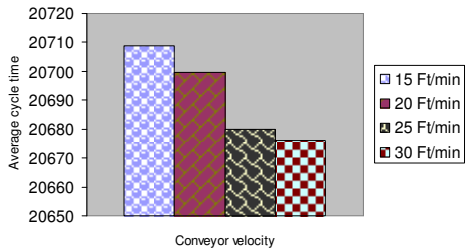


Figure 8: Average cycle time (in minutes) during a surge in demand

Figure 9 shows the amount of parts servicing in the system or work in process (WIP) during a drop in demand of the product. This was least at a conveyor velocity of 25 Ft/min at just above 83823 parts and most at 15 Ft/min conveyor velocity.

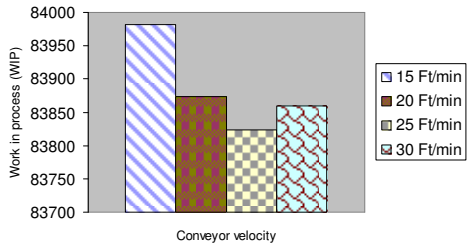


Figure 9: Average number of parts in system (WIP) during a drop in demand

The amount of queue seen in the system during a drop in demand (Figure 10) dropped with increase in conveyor velocity. There were at least 2192 parts at velocity of 30 Ft/min. Expectedly, the system was not exploding with parts in service since there weren't too much activities going on.

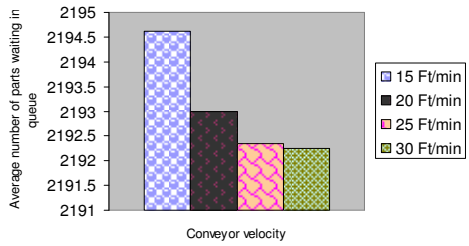


Figure 10: Average number of parts waiting in queue during a drop in demand

The system's behavior was investigated during some five hour equipment breakdown in two sub-factories. Parts spent the least time on average (figure 11) at the 30 Ft/min conveyor velocity at 18564 minutes.

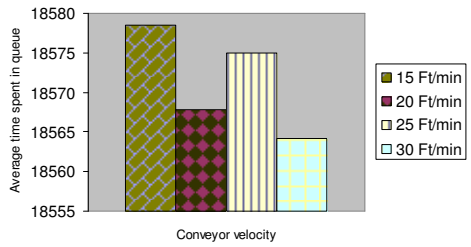


Figure 11: Average waiting time in queue (in minutes) during equipment breakdown

The average scheduled utilization during equipment break down (figure 12) stayed marginally displaced at just under 69% throughout, not minding an increase in conveyor velocity.

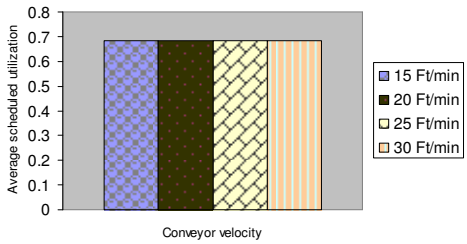


Figure 12: Average scheduled utilization during equipment breakdown

5. CONCLUSION

The paper has reported on the simulation model development of the integration of automotive OEM and their key suppliers. The modeling and simulation focus was on harmonizing as well as synchronizing the operations of these different parts suppliers, who have now become assemblers of their modular components while residing side by side with each other on the assembly line, and harnessing the synergic effects of such 'hands on' collaboration to boost lean production and provide agile capability for rapid response to competitive markets. Hence the truly agile manufacturing framework/ structure formed in the FMP is ultimately used to carry out production with a sense of shared or mutual dependency, motivation and a heightened sense of responsibility between OEMs and this web of suppliers that provide all the elements required in the production process perhaps under one roof.

REFERENCES

Cross, B., and Gordon, J., (1995), Partnership strategies for market success, *Business Quarterly, Autumn*, pp. 91-6

Dorrell, K., (1996), Auto industry prepares to weather the storm, Plant, Canada's Industrial Newspaper

Dyer, J., Cho, D. S. and Chu, W, (1998), Strategic supplier segmentation: the next 'best practice' in supply chain management, *California management Review*, Vol.4, No. 2, pp.57 - 77.

Kelton, W.D., Sadowski, R. P., Sadowski, D. A., (1998), Simulation with Arena, The McGraw-Hill companies, inc. USA.

Kelton W., Sadowski R., Sturrock D., (2004), Simulation with Arena. 3rd Ed. New York: McGraw-Hill Companies.

Noori, H., and Lee, W. (2000): Fractal manufacturing Partnership: Exploring a new form of strategic alliance between OEMs and suppliers, logistics information management, Volume 13, No. 5 pp.301-311.

Phelan, M., (1996), Stalking the elusive 5-day car, Automotive industries, November, Vol. 176, p.62.

Rathmell, J., and Sturrock, D. T., (2002), The Arena product family: Enterprise modeling solutions, proceedings of the 2002 winter simulations conference.

Ryu, K. and Jung, M., (2004), Goal-orientation mechanism in a fractal manufacturing system. International Journal of Production Research, 42, 11, 2207-2225.

Ryu, K. and Jung, M., (2003), Agent-based fractal architecture and modeling for developing distributed manufacturing systems. International Journal of Production Research, Vol. 41, No. 17, pp. 4233-4255.

Shin, M., Mun, J., Lee, K., and Jung, M., 2008, r-FrMS: a relation-driven fractal organization for distributed manufacturing systems, *International Journal of Production Research*, pp. 1-24.

Sihn, W. and von Briel, R., (1997), Process cost calculation in a fractal company. *International Journal of Technology Management*, 13, 68-77.

Simonian, H., (1996), Alliances forged in the factory binding carmakers to parts firms. The financial post, 9 November, pp.102. (The Financial Times).

Simonian, H., (1997), Car making joint ventures takes radical route. The financial post daily, 11 July, pp.49. (The Financial Times).

Strauss, R., and Hummel, T., (1995), The new industrial engineering revisited - information technology, business process reengineering, and lean management in the self-organizing fractal company, proceedings of the 1995 IEEE Annual International Engineering Management Conference, Theme "Global Engineering Management: Emerging Trends in the Asia Pacific", ed. F.S. Wei, pp. 287-292

Sweet, S., and Grace-Martin, K., 2003, Data analysis with SPSS, a first course in applied statistics, 2nd ed., *Pearson education*, Inc. USA.

Tirpak, T.M., Daniel, S.M., Lalonde, J.D., and Davies, W.J., (1992), A note on a fractal architecture for modeling and controlling flexible manufacturing systems. IEEE Transactions on Systems, Man and Cybernetics. Vol. 22, No. 3.

Warnock, I., (1996), Manufacturing and business excellence, Strategies, techniques and technology, Prentice hall, Europe, pp.120.

Warnecke, H.J., (1993), The fractal company- A revolution in corporate culture, Springer- Verlag.

AUTHORS BIOGRAPHY

Professor Sameh M. Saad, BSc (Honours), MSc, PhD, CEng, MIET, MILT, is Professor of Enterprise Modelling and Management, Postgraduate Research Coordinator and MSc/MBA Course Leader, in the Department of Engineering, Faculty of Arts, Computing, Engineering and Sciences, Sheffield Hallam University, UK. His research interests and experience include modelling and simulation, design and analysis of manufacturing systems, production planning and control, reconfigurable manufacturing systems and next generation of manufacturing systems including fractal and biological manufacturing systems. He has published over 130 articles in various national

and international academic journals and conferences, including keynote addresses and a book.

Julian Aririguzo C. BEng, MSc**,** MIET**,** is final year PhD. student in Manufacturing Systems at Sheffield Hallam University. He has an MSC. in Automatic Control and Systems Engineering from University of Sheffield and a BEng. Degree in Mechanical/ Production Engineering from Enugu state University of Science and Technology, Nigeria. He has published various research papers based on his research. He is passionate about resource efficiency in facility developments, innovative technologies in efficient use of resources and his research interests also include sustainability, sustainable developments and environmental protection/ clean energies.