THE 9TH INTERNATIONAL CONFERENCE ON MODELING & APPLIED SIMULATION

OCTOBER 13-15 2010
FES, MOROCCO

EDITED BY
ROSA M. AGUILAR
AGOSTINO BRUZZONE
CLAUDIA FRYDMAN
FRANCESCO LONGO
MARINA MASSEI
KHALID MEKOUAR
MIQUEL ANGEL PIERA

PRINTED IN FES, OCTOBER 2010

ISBN 2-9524747-5-3
EAN 9782952474757

THE 9TH INTERNATIONAL CONFERENCE ON MODELING & APPLIED SIMULATION

OCTOBER 13-15 2010, FES, MOROCCO

ORGANIZED BY

DIPTEM – UNIVERSITY OF GENOA

LIOPHANT SIMULATION

ECOLE SUPERIEURE D'INGENIERIE EN SCIENCES APPLIQUEES

UNIVERSITY OF AIX-MARSEILLE - LSIS MARSEILLE

MECHANICAL DEPARTMENT, UNIVERSITY OF CALABRIA

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

MSLS – AUTONOMOUS UNIVERSITY OF BARCELONA

SPONSORED BY

ACADEMIC, INSTITUTES, INDUSTRY AND SOCIETIES SPONSORS

SIMULATION TEAM

MISS – MCLEOD INSTITUTE OF SIMULATION SCIENCE

M&SNET – MODELING & SIMULATION NETWORK

IMCS – INTERNATIONAL MEDITERRANEAN & LATIN AMERICAN COUNCIL OF SIMULATION

UNIVERSITY OF LA LAGUNA

MISS LATVIAN CENTER – RIGA TECHNICAL UNIVERSITY

MISS SPANISH CENTER - UAB

UNIVERSITY OF PERUGIA

MAST – MANAGEMENT AND ADVANCED SOLUTIONS AND TECHNOLOGIES

LAMCE, UFRJ

CIFASIS: CONICET - UNR - UPCAM

EDITORS

Rosa M. Aguilar
University of la Laguna
Camino San Francisco de Paula, s/n
38271, La Laguna, Tenerife, Spain
raguilar@ull.es
www.isaatc.ull.es

Agostino Bruzzone
MISS-DIPTEM, University of Genoa
Via Opera Pia 15
16145 Genova, Italy
agostino@itim.unige.it
http://st.itim.unige.it

Claudia Frydman
LSIS - Laboratoire des Sciences de l'Information et des Systemes
Universite Aix-Marseille (U3)
Domaine Universitaire de Saint-Jerome, Av Escadrille Normandie-Niemen, 13397 Marseille Cedex 20
claudia.frydman@lsis.org
www.lsis.org

Francesco Longo
MSC-LES, Modeling & Simulation Center – Laboratory of Enterprise Solutions
Mechanical Department
University of Calabria
Via P. Bucci – 87036, Rende (CS), Italy
flongo@unical.it
www.msc-les.org

Marina Massei
SimulationTeam
Via Molinero 1,
Savona 17100, Italy
marina.massei@simulationteam.com
www.simulationteam.com

Khalid Mekouar
Ecole Superieure d'Ingenierie en Sciences Appliquees
29 bis Av Ibn Khatib Route d'Immouzzer
Fes, Morocco
k.mekouar@esisa.ma
http://esisa.ac.ma/

Miquel Angel Piera
MSLS, Department of System Engineering
University Autonomous of Barcelona
08193- Bellaterra- Spain
MiquelAngel.Piera@uab.es
www.uab.es

THE INTERNATIONAL MEDITERRANEAN AND LATIN AMERICAN MODELING MULTICONFERENCE, I3M 2010

GENERAL CO-CHAIRS

AGOSTINO BRUZZONE, *MISS DIPTEM, UNIVERSITY OF GENOA, ITALY*
CLAUDIA FRYDMAN, *LSIS, UNIVERSITY OF AIX-MARSEILLE, FRANCE*

PROGRAM CHAIR

KHALID MEKOUAR, *ESISA, MOROCCO*
MIQUEL ANGEL PIERA, *AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN*

THE 9TH INTERNATIONAL CONFERENCE ON MODELING & APPLIED SIMULATION, MAS 2010

GENERAL CO-CHAIRS

FRANCESCO LONGO, *MSC-LES, UNIVERSITY OF CALABRIA, ITALY*
ROSA M AGUILAR, *UNIVERSITY OF LA LAGUNA, SPAIN*

PROGRAM CO-CHAIRS

MARINA MASSEI, *SIMULATIONTEAM, ITALY*
KHALID MEKOUAR, *ESISA, MOROCCO*

MAS 2010 International Program Committee

Rosa Aguilar, University of la Laguna, Spain
Naamane Aziz, *LSIS, France*
Enrico Bocca, *Liophant Simulation, Italy*
Philippe Bouche, *INSA Strasbourg, France*
Agostino Bruzzone, *University of Genoa, Italy*
Ivan Castilla*, University of la Laguna, Spain*
Antonio Cimino, *MSC-LES, University of Calabria, Italy*
Duilio Curcio, *University of Calabria, Italy*
Fabio De Felice, *University of Cassino, Italy*
Priscilla Elfrey, *Kennedy Space Center, USA*
Claudia Frydman, *LSIS France*
Claudio Garuti, *Fulcrum Engineering Ltd., Cile*
Norbert Giambiasi, *LSIS, France*
Francesco Longo, *MSC-LES, University of Calabria, Italy*
Roberto Lu, *Boeing Company, USA*
Marina Massei, *Simulationteam, Italy*
Khalid Mekouar, *ESISA, Morocco*
Giovanni Mirabelli, *University of Calabria, Italy*
Istvan Molnar, *Bloomsburg University, USA*
Elio Padoano, *University of Trieste, Italy*
Elisabetta Pedace, *University of Catanzaro, Italy*
Jeroen Schuddebeurs, *Energy research Centre, Netherlands*
Adriano Solis, *York University, Canada*
Gregory Zacharewicz, *University of Bordeaux, France*
Cecilia Zanni, *INSA Strasbourg, France*

Tracks Chairs

Decision Support Systems Applications
Chair: Fabio De Felice, *University of Cassino (Italy)*

Inventory Management Simulation
Chair: Adriano solis, York University (Canada), Duilio Curcio, University of Calabria (Italy)

Enterprise and Supply Chain Modeling & Interoperability
Chairs: Cecilia Zanni, INSA Strasbourg (France); Philippe Bouche INSA Strasbourg (France); Gregory Zacharewicz, IMS Universite Bordeaux 1 (France)

Decision Support Systems Applications
Chair: Fabio De Felice, University of Cassino (Italy)

Production Simulation
Chair: Roberto Lu, The Boeing Company (USA)

Prototyping complex systems using Hardware In the Loop (HIL) Simulations
Chair: Schuddebeurs Jeroen, Energy research Centre of the Netherlands (Netherlands)

Production Systems Design
CHAIR: Antonio Cimino, University Of Calabria (Italy)

Chairs' Message

Welcome To MAS 2010

The Annual Modeling and Applied Simulation Conference is the ideal International Framework for presenting the very best research results, problem solutions and insight on new challenges facing the field of M&S Applications; MAS is the youngest pillar of I3M, even if, 2010 edition represents the 9th time this event is organized in the Simulation International Community.

Nowadays, Modeling & Simulation play a significant role in many critical areas and new Simulation Applications are strongly required by many different sectors that expect to use M&S as enabling technology for introduction break-trough innovation.
Obviously the use of simulation into new and complex problems provides researchers and scientists with big challenges; such challenges result from market evolution, sustainability issues, product innovation, interactive communications, globalization, mass customization, intensification of competition, and many others. So it becomes more and more important for succeeding in creating effective M&S applications to develop multidisciplinary teams and networks involving different skills and expertise including both academicians, technicians, subject matter experts and practitioners. MAS 2010 is focusing exactly on this goal and it represents a very effective moment for the International Community to discuss up-to-date initiatives, opportunities provided by enabling technologies and scenario evolutions.

Indeed MAS 2010 provides us with a timely opportunity to reflect and discuss problems and solutions, to identify new issues and to shape future directions for research and industry users. MAS 2010 is also an opportunity for researchers and practitioners from industry, academia and government to come together to share and advance their knowledge, discussing and working together for setting up new R&D proposals and projects based on state of art of M&S techniques and technologies.

The submitted papers were subjected to hard review not only to select best manuscripts, but even to feedback authors with comments and suggestions to improve the overall quality of the contributions. As results, the selection shows that all the most innovative themes and issues in this area are under consideration at the MAS 2010 and all mainstream features of Applied Simulation are addressed.

Last but not least, Fes is an unique combination of culture and tourists attractions in a fast evolving framework: so we hope that this year edition will leave you with a cultural, scientific and networking experience that is uniquely MAS 2010 in Fes.

Francesco Longo, MSC-LES, University of Calabria, Italy

Khalid Mekouar, *ESISA, Maroc*

ACKNOWLEDGEMENTS

The MAS 2010 International Program Committee selected the papers for the Conference among many submissions and we expected a very successful event based on their efforts; so we would like to thank all the authors as well as the IPCs and reviewers for their review process.
A special thank to the organizations, institutions and societies that are supporting and technically sponsoring the event: University of Genoa, Liophant Simulation, Ecole Supérieure d'Ingénierie en Sciences Appliquées, University of Aix-Marseille, Autonomous University of Barcelona, University of Calabria, Modeling & Simulation Center – Laboratory of Enterprise Solutions (MSC-LES), McLeod Institute of Simulation Science (MISS), Modeling & Simulation Network (M&SNet), International Mediterranean & Latin American Council of Simulation (IMCS), Management and Advanced Solutions and Technologies (MAST). Finally, we would like to thank all the Conference Organization Supporters.

Table of Contents

INVENTORY CONTROL PERFORMANCE OF VARIOUS FORECASTING METHODS WHEN DEMAND IS LUMPY

Adriano O. Solis[(a)], Somnath Mukhopadhyay[(b)], Rafael S. Gutierrez[(c)]

[(a)] Management Science Area, School of Administrative Studies, York University, Toronto, Ontario M3J 1P3, Canada
[(b)] Department of Information & Decision Sciences, The University of Texas at El Paso, El Paso, TX 79968-0544, USA
[(c)] Department of Industrial & Systems Engineering, The University of Texas at El Paso, El Paso, TX 79968-0521, USA

[(a)] asolis@yorku.ca, [(b)] smukhopadhyay@utep.edu, [(c)] rsgutier@utep.edu

ABSTRACT
This study evaluates a number of methods in forecasting lumpy demand – single exponential smoothing, Croston's method, the Syntetos-Boylan approximation, an optimally-weighted moving average, and neural networks (NN). The first three techniques are well-referenced in the intermittent demand forecasting literature, while the last two are not traditionally used. We applied the methods on a time series dataset of lumpy demand. We found a simple NN model to be superior overall based on several scale-free forecast accuracy measures. Various studies have observed that demand forecasting performance with respect to standard accuracy measures may not translate into inventory systems efficiency. We simulate on the same dataset a periodic review inventory control system with forecast-based order-up-to levels. We analyze resulting levels of on-hand inventory, shortages, and fill rates, and discuss our findings and insights.

Keywords: lumpy demand forecasting, neural networks, inventory control, simulation

1. INTRODUCTION

When there are intervals with no demand occurrences for an item, demand is said to be *intermittent*. Intermittent demand is also *lumpy* when there are large variations in the sizes of actual demand occurrences. Intermittent or lumpy demand has been observed in both manufacturing and service environments (Willemain, Smart, Schockor, and DeSautels 1994; Bartezzaghi, Verganti, and Zotteri 1999; Syntetos and Boylan 2001, 2005; Ghobbar and Friend 2002, 2003; Regattieri, Gamberi, Gamberini, and Manzini 2005; Teunter, Syntetos, and Babai 2010). In proposing a theoretically coherent scheme for categorizing demand into four types (smooth, erratic, intermittent, and lumpy), Syntetos, Boylan, and Croston (2005) suggest $CV^2 > 0.49$ and $ADI > 1.32$ for characterizing lumpy demand (where CV^2 represents the squared coefficient of variation of demand sizes and *ADI* is the average inter-demand interval).

We apply a number of forecasting methods to actual demand data from an electronic components distributor operating in Monterrey, Mexico, involving 24 stock keeping units (SKUs) each with 967 daily demand observations exhibiting a wide range of demand values and intervals between demand occurrences. Values of CV^2 range between 9.84 and 45.93 while values of *ADI* range between 3.38 and 5.44 (see Table 1) – all well over the cutoffs for lumpy demand as specified above.

Table 1: Basic Dataset Statistics

Series	1	2	3	4	5	6
% Nonzero Demand	30.4	32.8	32.7	34.1	35.7	36.2
Mean Demand	251.02	262.08	271.60	274.43	278.01	324.84
Std Dev	1078.80	985.19	1305.36	1221.31	1191.04	1387.20
CV^2	18.47	14.13	23.10	19.81	18.35	18.24
ADI	4.51	4.25	4.78	3.97	3.77	3.73

Series	7	8	9	10	11	12
% Nonzero Demand	32.4	33.3	34.4	33.8	35.0	35.2
Mean Demand	237.09	274.31	253.77	346.04	303.11	321.61
Std Dev	743.88	1134.55	959.19	1710.19	1229.80	1149.70
CV^2	9.84	17.11	14.29	24.43	16.46	12.78
ADI	5.21	4.73	4.03	4.83	5.14	4.83

Series	13	14	15	16	17	18
% Nonzero Demand	33.6	34.1	35.2	35.0	33.8	36.3
Mean Demand	299.15	296.07	288.78	305.81	228.74	352.32
Std Dev	1425.87	1321.28	1090.65	1257.98	889.07	1480.69
CV^2	22.72	19.92	14.26	16.92	15.11	17.66
ADI	5.44	4.68	4.39	4.41	4.30	4.09

Series	19	20	21	22	23	24
% Nonzero Demand	38.1	34.7	35.8	33.0	35.7	32.7
Mean Demand	322.98	355.48	328.70	394.84	314.33	410.00
Std Dev	1054.75	1609.05	1390.67	2675.95	1438.57	1929.56
CV^2	10.66	20.49	17.90	45.93	20.95	22.15
ADI	3.90	4.86	4.09	4.37	3.38	3.39

Seven forecasting methods were initially evaluated, namely:

- single exponential smoothing (SES)
- Croston's method
- Croston's method with two separate smoothing constants
- the Syntetos-Boylan approximation
- the Syntetos-Boylan approximation with two separate smoothing constants
- a five-period weighted moving average with optimized weights
- neural networks.

1.1 Well-Referenced Methods for Forecasting Lumpy Demand

Croston (1972) noted that SES, frequently used for forecasting in inventory control systems, has a bias that places the most weight on the most recent demand occurrence. He proposed a method of forecasting intermittent demand using exponentially weighted moving averages of nonzero demand sizes and the intervals between nonzero demand occurrences to address the bias problem. Leading application software packages for statistical forecasting incorporate Croston's method (Syntetos and Boylan 2005; Boylan and Syntetos 2007).

While Croston assumed a common smoothing constant α, Schultz (1987) suggested that separate smoothing constants, α_i and α_s, be used for updating the inter-demand intervals and the nonzero demand sizes, respectively. Eaves and Kingsman (2004) provide a clear formulation of Croston's method with 'two alpha values'. In the current study, for each demand series, we identify the combination of two alphas corresponding to the best forecast in the calibration sample. We then apply the best combination of α_i and α_s for each series to forecast the test sample.

Syntetos and Boylan (2001, 2005) reported an error in Croston's mathematical derivation of expected demand, leading to a positive bias. Syntetos and Boylan (2005) proposed what is now referred to in the literature as the Syntetos-Boylan approximation (SBA) – which involves multiplying Croston's estimator of mean demand by a factor of $(1-\alpha_i/2)$, where α_i is the exponential smoothing constant used in updating the inter-demand intervals.

We note, however, that Syntetos and Boylan (2005) used the same smoothing constant for updating demand sizes as for updating inter-demand intervals in applying SBA to monthly demand histories over a two-year period of 3000 stock-keeping units (SKUs) in the automotive industry. As we do with Croston's method in the current study, we likewise consider SBA with separate smoothing constants, α_i and α_s, for updating the inter-demand intervals and the nonzero demand sizes. Other than Schultz (1987), only Syntetos, Babai, Dallery, and Teunter (2009) and Teunter, Syntetos, and Babai (2010) have to-date reported using two separate smoothing constants on inter-demand intervals and demand sizes in empirical investigation – in the two latter studies, applied to the SBA demand estimator.

The use of low α values in the range of 0.05-0.20 has been recommended in the literature on lumpy demand (Croston 1972; Johnston and Boylan 1996). Syntetos and Boylan (2005) used the four α values of 0.05, 0.10, 0.15, and 0.20 for the SES, Croston's, and SBA methods. We use these same four values in the current study.

1.2 'Non-Traditional' Methods for Forecasting Lumpy Demand

Sani and Kingsman (1997) observed that less sophisticated (e.g., moving average) methods can prove superior to Croston's method in practice. Eaves (2002) also found that forecasting methods simpler than Croston's or SBA method can provide better forecasting results for intermittent and slow-moving demand. Regattieri, Gamberi, Gamberini, and Manzini (2005) studied monthly demand data pertaining to spare parts for Alitalia's fleet of Airbus A320 aircraft in 1998-2004. They found weighted moving average (WMA) forecasts, based on selecting the best sets of weights for three, five, and seven-month periods, to perform generally better than Croston's, SES, and other smoothing methods (SBA was not considered).

In the current study, we applied a five-day weighted moving average method with optimized weights (WMA5) – to correspond to weekly demand over a five-day work week. The method averages the last five lagged values of lumpy demand through optimized weights. The lagged value 1 means the demand during the last time period and so on. To determine the optimized weights, the method runs a standardized linear ordinary least square (OLS) regression on current period demand as target variable and the five most recent lagged period demands as predictor variables. The beta values of the lagged demands are normalized so that the values add up to 1.000. The normalized values (see Table A.1 in the Appendix) are used as the moving average weights. The method determines the weights from calibration data (as discussed in Section 2.1) only.

Researchers have used neural network (NN) models in various forecasting applications. NN models can provide reasonable approximations to many functional relationships (e.g., White 1992; Elman and Zipser 1987), with flexibility and nonlinearity cited as their two most powerful aspects. Hill, O'Connor, and Remus (1996) compared forecasts produced by NN models against forecasts generated using six time series methods from a systematic sample of 111 of the 1001 time series in a well known 'M-competition' (Makridakis, Andersen, Carbone, Fildes, Hibon, Lewandowski, Newton, Parzen, and Winkler 1982). They found NN forecast models to be significantly more accurate than those of the six traditional time series models for monthly and quarterly demand data across a number of selection criteria. Very few previous studies have used NN to forecast irregular or lumpy demand (e.g., Carmo and Rodrigues 2004; Gutierrez, Solis, and Mukhopadhyay 2008).

We used a multi-layered perceptron (MLP) trained by a back-propagation (BP) algorithm (Rumelhart, Hinton, and Williams 1988). We followed guidelines proposed by a fairly recent study on MLP architecture selection (Xiang, Ding, and Lee 2005) which suggests that one should first try a three-layered MLP. One should also start with the minimum number of hidden units required to approximate the target function.

Functions learned by a minimal net over calibration sample points work well on new samples. We used three layers of network:

- one input layer for input variables
- one hidden unit layer
- one output layer of one unit.

We chose three hidden units, which is a reasonably low number required to approximate any complex function. The network connects all hidden nodes with the input nodes representing the last time period's demand value and cumulative number of time periods with zero demand. The output node representing the current period's demand value connects to all hidden nodes. We used 0.1 for the learning rate and 0.9 for the momentum factor, as recommended by seminal research (Rumelhart, Hinton, and Williams 1988).

NN usually can approximate any function with the proper choice of parameters and a specific network structure (Lippmann 1987). Eventually, after a repeated change of network structure and parameter values, one can find a "successful" combination of calibration and validation samples which provides a false impression of model generalization. In this study, we choose a simple network structure with the same parameter values across all 24 lumpy demand series. We validate once and report the results without going back to improve upon them. If, accordingly, the NN model with this restriction outperforms other methods on the test sample, we are able to conclude the model to be superior. We do not change the parameter values of NN across all the 24 time series. On the other hand, we relax the restriction on other methods by trying out different parameter values as recommended in the literature.

2. DATA SET PARTITIONING AND FORECAST ACCURACY MEASURES

2.1 Data Set Partitioning

We initially used the first 624 observations of the 967 daily demand observations in each of the 24 time series to "train" and validate the models (the *training* sample). We then tested, at each of the four values of α, the other forecasting models under consideration on the final 343 observations (the *test* sample). This generated an approximately 65:35 (65% training data and 35% test data) partitioning. Researchers typically use an 80:20 split to validate models (Bishop 1995). To compare the forecasting methods further we have also ran the models on 50:50 and 80:20 data partitions. Due to space limitations, however, we report results only for the 65:35 data partitioning in this paper.

2.2 Forecast Accuracy Measures

Mean absolute percentage error (MAPE) is the most widely used accuracy measure for ratio-scaled data. The traditional definition of MAPE involves terms of the form $|E_t|/A_t$ (where A_t and E_t, respectively, represent actual demand and forecast error in period t). Since lumpy demand involves periods with zero demands, the traditional MAPE definition fails. We used an alternative specification of MAPE as a ratio estimate (Gilliland 2002), which guarantees a nonzero denominator:

$$\text{MAPE} = \left(\sum_{t=1}^{n} |E_t| \Big/ \sum_{t=1}^{n} A_t \right) \times 100 \,. \tag{1}$$

Willemain, Smart, Schockor, and DeSautels (1994) conducted a study comparing performance of SES and Croston's method in intermittent demand forecasting, using (i) MAPE based on the above ratio estimate, (ii) median absolute percentage error (MdAPE), (iii) root mean squared error (RMSE), and (iv) mean absolute deviation (MAD) as forecast accuracy measures. However, they reported only MAPEs, noting that relative results were the same for all four measures. Eaves and Kingsman (2004) applied MAPE, RMSE, and MAD in comparing the performance of several methods (SES, Croston's, SBA, 12-month simple moving average, and the previous year's simple average) in forecasting demand for spare parts for in-service aircraft of the Royal Air Force (RAF) of the UK. Using demand data over a six-year period for 18750 SKUs randomly selected out of some 685000 line items, they found SBA to provide the best results overall using MAPE, but the 12-month simple moving average yielded the best MADs overall.

Armstrong and Collopy (1992) did an extensive study for making comparisons of errors across time series. For selecting the most accurate method, they recommend the median RAE (MdRAE) when few time series are available. The relative absolute error (RAE) is calculated for a given series, at a given time t, by dividing the absolute error under method m, $|F_{m,t} - A_t|$, by the corresponding absolute error for the random walk, $|F_{rw,t} - A_t|$. We compute the random walk forecast by simply adding one unit to the actual demand in the immediately preceding period. Hence,

$$\text{RAE}_t = |F_{m,t} - A_t| \big/ |(A_{t-1} + 1) - A_t| \,. \tag{2}$$

MdRAE is simply the median of all RAE_t values across the entire test sample.

Syntetos and Boylan (2005) employed two accuracy comparison measures: relative geometric root-mean square error (RGRMSE) and percentage best (PB). The first measure is as follows:

$$\text{RGRMSE} = \left(\prod_{t=1}^{n} (A_{a,t} - F_{a,t})^2 \right)^{1/2n} \Bigg/ \left(\prod_{t=1}^{n} (A_{b,t} - F_{b,t})^2 \right)^{1/2n} \tag{3}$$

where the symbols $A_{m,t}$ and $F_{m,t}$ denote actual demand and forecast demand, respectively, under forecasting method *m* at the end of time period *t*. PB, another scale-free accuracy measure, is the percentage of time periods that one method outperforms all the other methods. We use absolute error as the criterion to assess alternative methods' performance under the PB approach. Gutierrez, Solis, and Mukhopadhyay (2008) used MAPE as well as RGRMSE and PB to assess performance of the SES, Croston's, SBA, and NN forecasting methods.

In the current study, we assess and compare the performance of the seven forecasting methods – specified in Section 1 – as applied to the test samples in the 24 time series in the dataset, using four scale-free error criteria: (i) MAPE, (ii) MdRAE, (iii) RGRMSE, and (iv) PB. We used SAS software release 9.1 for our empirical investigations of both forecasting performance (reported in Section 3) and inventory control performance (reported in Section 4).

3. EMPIRICAL INVESTIGATION OF FORECASTING PERFORMANCE

Like Syntetos and Boylan (2005), Gutierrez, Solis, and Mukhopadhyay (2008) applied four α values: 0.05, 0.10, 0.15, and 0.20. The latter study found the SES, Croston's and SBA methods to work best with $\alpha = 0.05$ for all 24 time series considered, which appears consistent with the lumpiness observed in the dataset.

For the Croston's and SBA methods with separate smoothing constants, we identified in the current study – for each demand series – the combination of α_i and α_s corresponding to the best forecast in the training sample, based upon a minimum MAPE criterion. We then use the best combination for each series (see Table A.2 in the Appendix) to generate forecasts on the test sample. (Eaves and Kingsman (2004), in applying the SES, Croston's and SBA methods, likewise optimized smoothing constants using MAPEs only, but cautioned that the smoothing methods may yield better results if smoothing constants were optimized using a different forecast accuracy criterion.)

Figure 1 shows the relative performance of all the seven methods with respect to MAPE under the 65:35 data partitioning. NN MAPEs are superior for 20 of the 24 time series. WMA5 is clearly the worst performer in all series. For four series (4, 22, 23, and 24), NN, Croston, SBA, and SES perform quite closely. If MAPE is the criterion to select the best method, a simple NN model is clearly the best performing method overall.

In the current study, we did not observe any substantial improvement in forecast accuracy arising from using separate smoothing constants, α_i and α_s. To execute forecasting and demand management, calibration of two-alpha combinations will add more complexity to the process. In light of practical implications, we decided to drop the two-alpha Croston and SBA methods. Moreover, because the SBA method is consistently superior to Croston's method, we proceed to investigate only four methods – SES, SBA, WMA5 and NN.

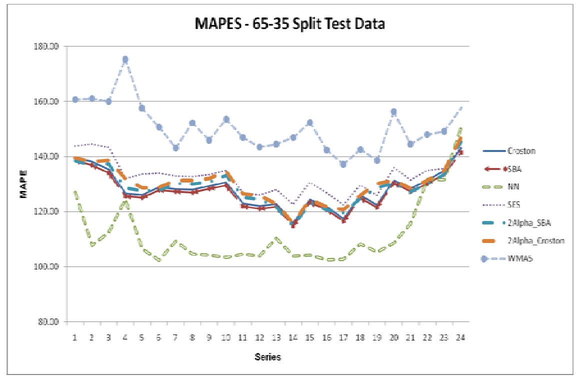


Figure 1: Comparison of MAPEs

Figure 2 shows the performance of the four remaining methods with respect to PB. NN is again the superior method overall, while WMA5 ranks second.

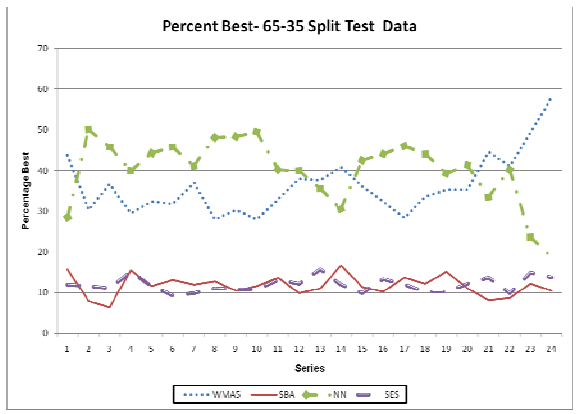


Figure 2: Comparison of Percentage Bests

Table 2 shows, for the 65:35 data partitioning, the best performing method across the 24 series for each accuracy measure. NN is the best method overall with respect to MAPE, MdRAE, and PB, while NN and WMA5 perform equally well with respect to RGRMSE. However, WMA5 performs poorly when MAPE is the criterion for selecting the best method. The other two methods, SES and SBA, which were developed and heavily researched for forecasting of intermittent/lumpy demand, did not perform as well as NN and WMA5.

4. EMPIRICAL INVESTIGATION OF INVENTORY CONTROL PERFORMANCE

Demand forecasting and inventory control have traditionally been examined independently of each other (Tiacci and Saetta 2009; Syntetos, Babai, Dallery, and Teunter 2009). In reality, demand forecasting performance with respect to standard accuracy measures may not translate into inventory systems efficiency (Syntetos, Nikolopoulos, and Boylan 2010). In an intermittent demand setting, a periodic review inventory control system has been recommended (Sani and Kingsman 1997; Syntetos, Babai, Dallery, and Teunter 2009). A number of recent studies that address both forecasting and inventory control performance for intermittent demand (e.g., Eaves and Kingsman 2004; Syntetos and Boylan 2006; Syntetos, Babai, Dallery, and Teunter 2009; Syntetos, Nikolopoulos, and Boylan 2010; Teunter, Syntetos, and Babai 2010) have employed the order-up-to (*T*,*S*) periodic review system (see, for example, Silver, Pyke, and Peterson 1998) – where *T* and *S* represent the review period and order-up-to level, respectively.

Table 2: Best Method by Forecast Accuracy Measure

	65:35 Data Partitioning			
Series	MAPE	MdRAE	RGRMSE	PB
1	NN	NN	WMA	NN
2	NN	NN	NN	NN
3	NN	NN	WMA	NN
4	NN	NN	WMA	NN
5	NN	NN	NN	NN
6	NN	NN	NN	NN
7	NN	NN	WMA	NN
8	NN	NN	NN	NN
9	NN	NN	NN	NN
10	NN	NN	NN	NN
11	NN	NN	NN	NN
12	NN	NN	NN	NN
13	NN	NN	WMA	WMA
14	NN	NN	WMA	WMA
15	NN	NN	NN	NN
16	NN	NN	NN	WMA
17	NN	NN	NN	NN
18	NN	NN	NN	NN
19	NN	NN	NN	NN
20	NN	NN	WMA	NN
21	NN	NN	WMA	WMA
22	SBA	NN	WMA	WMA
23	NN	WMA	WMA	WMA
24	SBA	WMA	WMA	WMA
Overall	**NN**	**NN**	**N/W**	**NN**

In the study by Eaves and Kingsman (2004) earlier discussed in Section 2.1, simulations of a (*T*,*S*) system were performed on actual demand data, aggregated quarterly, for the 18750 randomly selected SKUs. Forecast-based order-up-to levels *S* were determined as the product of the forecast demand per unit of time and the “protection interval”, *T*+*L* (where *L* is the reorder lead time). Implied average stockholdings were calculated using a backward-looking simulation assuming a common fill rate (or percentage of total demand filled by on-hand inventory) of 100%. SBA yielded the lowest average stockholdings among the five forecasting methods evaluated.

Syntetos and Boylan (2006) used a dataset consisting of monthly demand observations over a two-year period for 3000 SKUs in the automotive industry. They modeled demand over *T*+*L* in a (*T*,*S*) system by way of a negative binomial distribution – a compound Poisson distribution whose variance is greater than its mean. Two target fill rates were considered: 90% and 95%. Using two cost policies in simulation comparisons, they demonstrated the superior inventory control performance of the SBA forecasting method relative to Croston’s, SES, and 13-month simple moving average methods.

Two recent studies (Syntetos, Babai, Dallery, and Teunter 2009; Teunter, Syntetos, and Babai 2010) used a large dataset from the RAF of the UK, involving 84 monthly observations of demand for 5000 SKUs over seven years (1996-2002). The first 24 observations of each time series were used to initialize estimates of demand level and variance, and the second 24 observations were used to optimize separate smoothing constants, α_i and α_s, on inter-demand intervals and demand sizes, respectively. Simulation of inventory control performance in applying the SBA method was then performed over the final 36 observations. In the 2009 article, the authors noted that in many intermittent-demand situations the *ADI* is larger than the lead time (or the lead time plus one review period). They accordingly excluded those SKUs in the RAF dataset with *ADI* less than *T*+*L* (with *T* = 1 month in this case), resulting in 2455, or 49% of the original 5000 SKUs, actually being considered. In the 2010 article, lead time demand was modeled as a compound binomial process, with demands in successive periods being identically and independently distributed. Both studies introduced a new approach to determine, in a (*T*,*S*) inventory control system, order-up-to levels utilizing both inter-demand interval and demand size forecasts explicitly whenever demand occurs. Using various service-oriented and cost-oriented criteria, the two studies observed the superiority of the new approach compared to the classical approach which uses only the SBA estimate of average demand size.

Sani and Kingsman (1997) have earlier applied simulation of real data (consisting of 30 long series of daily demand data over five years for low demand items), involving a single run for each data series, as a form of empirical evaluation. In the current study, our simulations have also taken the form of a single run performed on the test sample consisting of the final 343 daily demand observations for each of the 24 series in the dataset. Simulation experiments involving multiple runs have not been attempted owing to the difficulty of mathematically modeling the degree of demand lumpiness observed in our dataset.

We assume in the current study a (*T*,*S*) periodic review inventory control system with full backordering. For initial simulation runs, *T* is five days (one week) and we assume a deterministic reorder lead time, *L*, of 10 days (two weeks). Let I_t and B_t, respectively, denote on-hand inventory and inventory shortage/ backlog at the time of review *t*, and F_j represent the forecast demand for period *j* (*j* = *t*+1, …, *t*+*T*+*L*). Without providing a safety stock component, the replenishment quantity based upon a forecast-based order-up-to level is

$$Q_t = \sum_{t+1}^{t+T+L} F_j - I_t + B_t \,. \qquad (4)$$

In our simulation studies, we continue to investigate only the four methods remaining under consideration – SES, SBA, WMA5 and NN – as identified in Section 3. Figure 3 shows the mean inventory on-hand for each of the 24 SKUs throughout the test sample. We excluded WMA5 from this figure, because means for most series are well over those computed when using SBA, SES, and NN. We find that the mean inventory on-hand arising from the use of NN is lower, in most instances, than when SBA or SES is used.

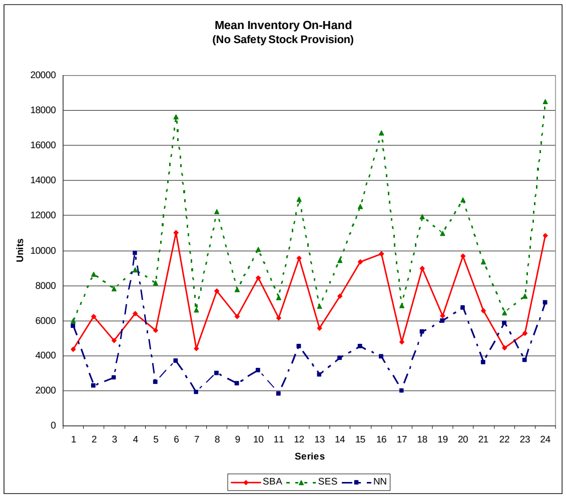


Figure 3: Mean On-Hand Inventory with No Safety Stock Provision

In Figure 4, however, we observe average backorders to be higher with NN than with SBA or SES. Mean shortages are much lower with WMA5, consistent with the much higher average on-hand inventory levels observed in Figure 3 for this method. Figure 5 shows that the percentage of time when inventory shortages occur is generally highest when NN is used. In like manner, we see in Figure 6 that the average fill rate is lowest overall when NN is used.

We reiterate that Figures 3-6 pertain to the case where there is no safety stock provided. The literature on inventory control suggests a safety stock component in order-up-to levels to compensate for uncertainty in demand during the "protection interval" $T+L$. For each demand series, we calculated the standard deviation s_{tr} of daily demand during the training sample. Initially, we set the safety stock level to be k standard deviations of daily demand during the training sample – i.e., $k \cdot s_{tr}$ – with k = 4, 6, 8, 10, and 12. We then proceeded to conduct single run simulations over the 343 observations in the test sample for each of the 24 series.

The values of k we have thus far tested give rise to safety stocks which are, more or less, comparable with the $z \cdot \sqrt{T+L} \cdot \sigma_d$ suggested when daily demand during the protection interval is assumed to be identically and independently normally distributed with standard deviation σ_d (e.g., Silver, Pyke, and Peterson 1998). With the safety stock component, the replenishment quantity to order is

$$Q_t = \sum_{t+1}^{t+T+L} F_j + k \cdot s_{tr} - I_t + B_t \,. \qquad (5)$$

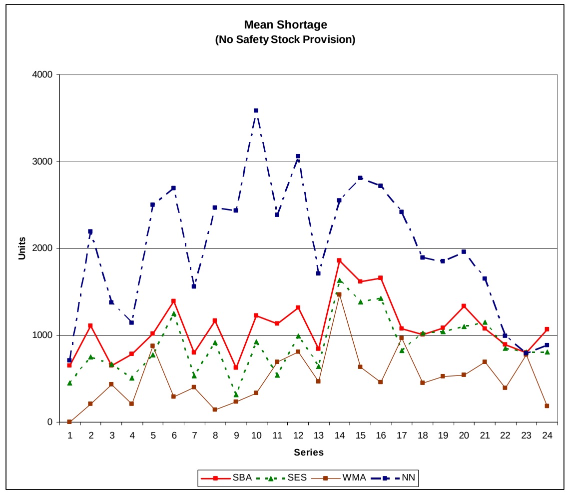


Figure 4: Mean Shortage with No Safety Stock Provision

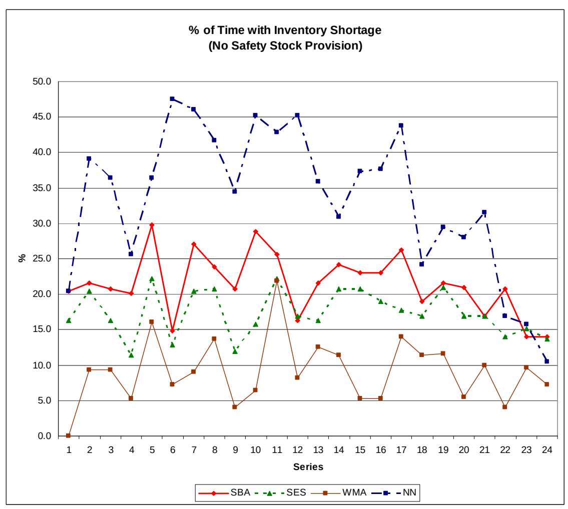


Figure 5: Percentage of Time Stocking Out with No Safety Stock Provision

When safety stock is set at $4 \cdot s_{tr}$, mean shortages as shown in Figure 7 have decreased significantly, although levels of on-hand inventory, as expected, have markedly increased. We see in Figure 8 that mean fill rates have substantially improved overall compared with those seen in Figure 6, even as mean fill rates when using NN continue to be generally lower than when WMA5, SES, and SBA are applied.

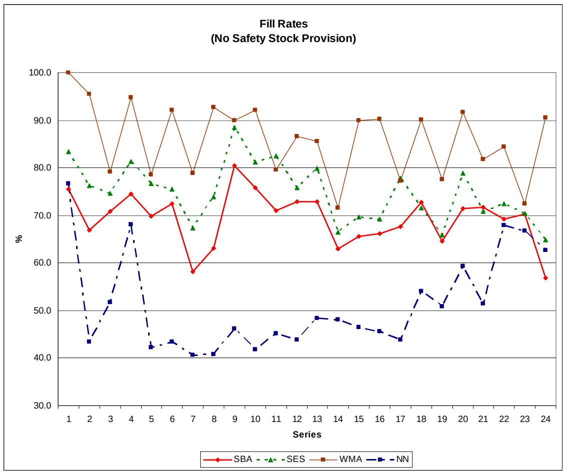


Figure 6: Average Fill Rates with No Safety Stock Provision

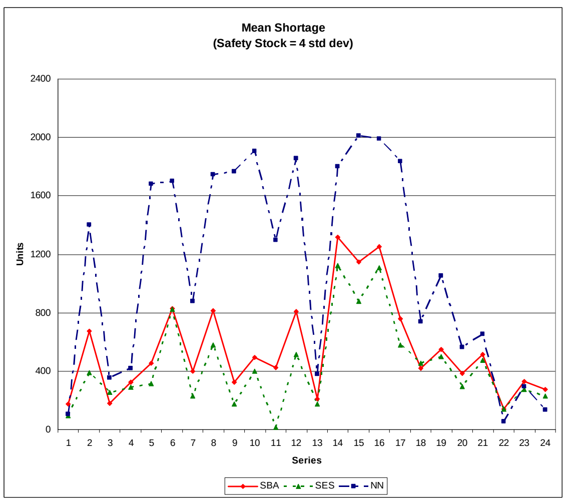


Figure 7: Mean Shortage with Safety Stock = $4 \cdot s_{tr}$

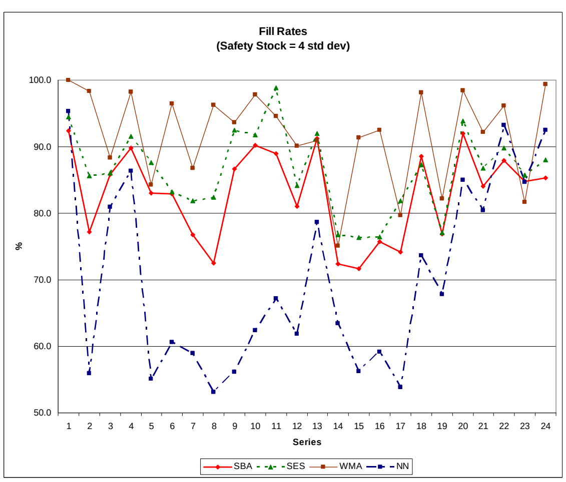


Figure 8: Average Fill Rates with Safety Stock = $4 \cdot s_{tr}$

We continue to see essentially the same mean fill rate comparisons as *k* is increased to 6, 8, 10, and 12. Mean fill rates when *k* = 8 are shown in Figure 9. We observe that all four methods under consideration lead to fill rates of 100% for series 22 and series 24.

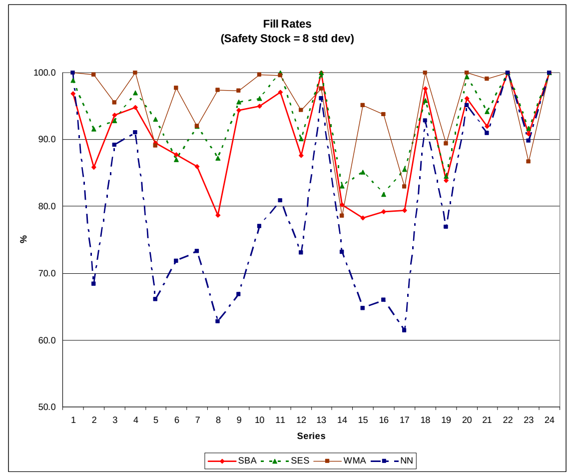


Figure 9: Average Fill Rates with Safety Stock = $8 \cdot s_{tr}$

Figure 10 shows the average on-hand inventory levels when *k* = 8. (Since fill rates arising from all methods are already at 100% for series 22 and series 24, these two series have been left out of Figure 10.) On the other hand, average backorder levels are shown in Figure 11. The lower mean fill rates with NN as forecasting method are clearly associated with generally lower average on-hand inventory levels but also generally greater mean shortages.

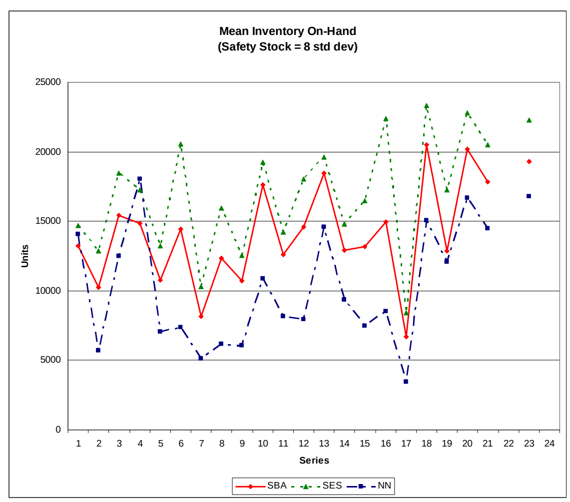


Figure 10: Mean On-Hand Inventory with Safety Stock = $8 \cdot s_{tr}$

Overall average fill rates across all 24 SKUs for each of the four methods under consideration are reported in Table 3 for the values of *k* tested.

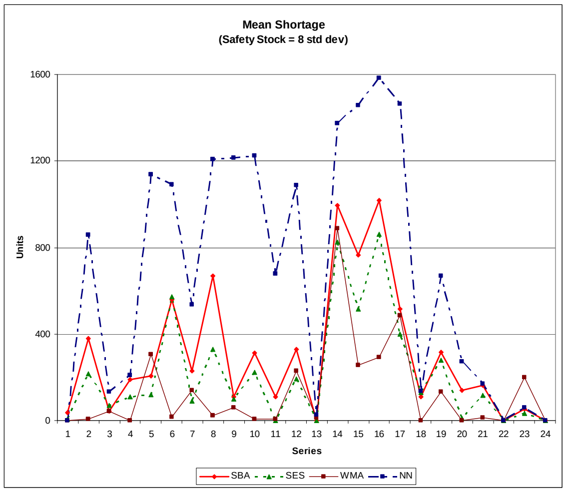


Figure 11: Mean Shortage with Safety Stock = $8 \cdot s_{tr}$

Table 3: Overall Average Fill Rates with Safety Stock = *k* Standard Deviations of Daily Demand

k	SBA	SES	WMA	NN
0	69.5	74.9	86.2	50.2
4	82.8	86.1	91.9	68.8
6	86.8	89.5	93.8	74.6
8	89.8	92.3	95.2	79.3
10	91.9	94.1	96.3	83.1
12	93.9	95.4	97.2	86.3

We have also conducted simulations with L = 3 days, for k = 0, 3, 5, 6, 7, 8, and 9. Similar comparisons of average on-hand inventory, backorders, and fill rates have arisen.

In view of the much higher levels of average on-hand inventory associated with demand forecasting using WMA5, we focus our attention on NN, SBA, and SES. At similarly specified safety stock levels, we observe much lower mean fill rates (i.e., inferior customer service levels) when NN – the "best" of the four methods based upon ratio-scaled traditional forecast accuracy measures – is applied in comparison with fill rates attained when using SES and SBA. In the same vein, NN yields relatively lower average on-hand inventory levels (i.e., lower inventory carrying costs) but higher mean shortages (i.e., higher backorder costs). While our dataset does not include specific cost information, a distributor of electronic components will be expected to pay significant attention to customer service levels and backorder costs.

Of additional interest is how SES and SBA compare in terms of stock control performance when demand is lumpy. Eaves and Kingsman (2004) and Syntetos and Boylan (2006) have found SBA to outperform several forecasting methods, SES included, when demand is intermittent though not lumpy. While SES is less sophisticated than SBA, the former yields generally higher average fill rates and lower average backorders than the latter. On the other hand, however, SES leads to somewhat higher average on-hand inventory levels than SBA.

5. CONCLUSIONS AND FURTHER WORK

In the current study, we find support for earlier assertions that demand forecasting performance with respect to standard accuracy measures may not translate into inventory systems efficiency. In particular, an NN model was found to outperform the SES and SBA methods in performance with respect to a number of scale-free traditional accuracy measures, but appears to be inferior when it comes to inventory control performance.

We intend to do further simulation work that will search, for each SKU in the dataset, for the value of k (and, hence, the safety stock component of the forecast-based order-up-to-level) that would meet a specified fill rate. Simulation studies of periodic review inventory systems generally involve searching for order-up-to-levels satisfying a target customer service level – e.g., a probability of not stocking out or a fill rate – often with a cost minimization objective (Solis and Schmidt 2009). For instance, Syntetos and Boylan (2006) evaluated performance of forecasting methods at target fill rates of 90% and 95%, while Teunter, Syntetos, and Babai (2010) considered target fill rates of 87%, 91%, 95%, and 99%. Boylan, Syntetos, and Karakostas (2008) initially set a fill rate of 95%, but later treated fill rate as a simulation parameter varying from 93% to 97%. Starting with comparable target fill rates, we will conduct simulation searches, with the resulting levels of on-hand inventory and backorders accordingly compared across the forecasting methods in terms of potential cost implications.

Based on the simulation searches outlined above, a more rational comparison between methods, especially between SES and SBA, should be possible.

APPENDIX

Table A.1: Optimized Weights for WMA5

	Optimized Weights on Lagged Demand				
Series	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5
1	0.434	0.348	0.036	0.055	0.127
2	0.113	0.086	0.174	0.344	0.282
3	0.151	0.162	0.205	0.133	0.349
4	1.017	-0.113	0.030	0.019	0.047
5	0.064	0.226	0.061	0.126	0.524
6	0.130	0.295	0.082	0.059	0.433
7	0.274	0.038	0.484	0.125	0.079
8	0.172	0.149	0.122	0.190	0.366
9	0.215	0.210	0.228	0.109	0.237
10	0.212	0.220	0.152	0.198	0.218
11	0.095	0.144	0.527	0.067	0.168
12	0.048	0.131	0.100	0.012	0.709
13	0.041	0.320	0.262	0.232	0.145
14	0.078	0.042	0.710	0.023	0.147
15	0.018	0.246	0.100	0.557	0.079
16	0.297	0.296	0.120	0.203	0.085
17	0.167	0.229	0.364	0.119	0.122
18	0.175	0.061	0.136	0.294	0.333
19	0.154	0.313	0.176	0.054	0.303
20	0.150	0.429	0.139	0.162	0.119
21	0.185	0.153	0.186	0.294	0.181
22	0.102	0.166	0.083	0.240	0.408
23	0.151	0.449	0.032	0.152	0.216
24	0.158	0.080	0.628	0.065	0.069

Table A.2: Minimum MAPEs of Two-alpha SBA and Two-alpha Croston's Methods on Training Sample

	Two-alpha SBA Method			Two-alpha Croston's Method		
Series	Minimum MAPE (%)	α_i	α_s	Minimum MAPE (%)	α_i	α_s
1	164.0	5%	5%	165.9	5%	5%
2	152.9	5%	5%	154.5	5%	5%
3	164.3	10%	5%	166.2	10%	5%
4	166.0	10%	5%	167.9	15%	5%
5	160.5	10%	5%	162.3	10%	5%
6	154.8	10%	5%	156.4	5%	5%
7	154.5	10%	5%	156.0	10%	5%
8	159.5	10%	5%	161.1	10%	5%
9	147.4	10%	5%	148.8	10%	5%
10	165.8	10%	5%	167.7	10%	5%
11	159.1	10%	5%	160.8	10%	5%
12	154.0	10%	5%	155.6	10%	5%
13	161.8	5%	5%	163.6	5%	5%
14	161.8	5%	5%	163.5	5%	5%
15	158.7	5%	5%	160.4	5%	5%
16	158.1	5%	5%	159.8	5%	5%
17	154.7	10%	5%	156.3	10%	5%
18	160.3	5%	5%	162.1	5%	5%
19	231.6	15%	20%	235.3	15%	20%
20	159.1	5%	5%	160.9	5%	5%
21	157.1	5%	5%	158.9	5%	5%
22	161.4	5%	5%	163.2	5%	5%
23	155.0	5%	5%	156.7	5%	5%
24	163.6	10%	5%	165.4	10%	5%

REFERENCES

Armstrong, J. S. and Collopy, F., 1992. Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, 8 (1), 69-80.

Bartezzaghi, E., Verganti, R., and Zotteri, G., 1999. A simulation framework for forecasting uncertain lumpy demand. *International Journal of Production Economics*, 59 (1-3), 499-510.

Bishop, C.M., 1995. *Neural Networks for Pattern Recognition*. Oxford, UK: Clarendon Press.

Boylan, J.E. and Syntetos, A.A., 2007. The accuracy of a Modified Croston procedure. *International Journal of Production Economics*, 107 (2), 511-517.

Boylan, J.E., Syntetos, A.A., and Karakostas, G.C., 2008. Classification for forecasting and stock control: a case study. *Journal of the Operational Research Society*, 59 (4), 473-481.

Carmo, J.L. and Rodrigues, A.J., 2004. Adaptive forecasting of irregular demand processes. *Engineering Applications of Artificial Intelligence*, 17 (2), 137-143.

Croston, J.D., 1972. Forecasting and stock control for intermittent demands. *Operational Research Quarterly*, 23 (3), 289-304.

Eaves, A.H.C., 2002. *Forecasting for the ordering and stock-holding of consumable spare parts*. Thesis (PhD). University of Lancaster.

Eaves, A.H.C. and Kingsman, B.G., 2004. Forecasting for the ordering and stock-holding of spare parts. *Journal of the Operational Research Society*, 55 (4), 431-437.

Elman, J.L. and Zipser, D., 1987. *Learning the hidden structure of speech*. Institute of Cognitive Science Report 8701. University of California, San Diego, USA.

Ghobbar, A.A. and Friend, C.H., 2002. Sources of intermittent demand for aircraft spare parts within airline operations. *Journal of Air Transport Management*, 8 (4), 221-231.

Ghobbar, A.A. and Friend, C.H., 2003. Evaluation of forecasting methods for intermittent parts demand in the field of aviation: a predictive model. *Computers & Operations Research*, 30 (14), 2097-2114.

Gilliland, M., 2002. Is forecasting a waste of time? *Supply Chain Management Review*, 6 (4), 16-23.

Gutierrez, R.S., Solis, A.O., and Mukhopadhyay, S., 2008. Lumpy demand forecasting using neural networks. *International Journal of Production Economics*, 111 (2), 409-420.

Hill, T., O'Connor, M., and Remus, W., 1996. Neural network models for time series forecasts. *Management Science*, 42 (7), 1082-1092.

Johnston, F.R. and Boylan, J.E., 1996. Forecasting for items with intermittent demand. *Journal of the Operational Research Society*, 47 (1), 113-121.

Lippmann, R.P., 1987. An introduction to computing with neural nets. *IEEE ASSP Magazine*, 4 (2), 4-22.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., and Winkler, R., 1982. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting*, 1 (2), 111-153.

Regattieri, A., Gamberi, M., Gamberini, R., and Manzini, R., 2005. Managing lumpy demand for aircraft spare parts. *Journal of Air Transport Management*, 11 (6), 426-431.

Rumelhart, D.E., Hinton, G.E., and Williams, R.J., 1988. Learning internal representations by error propagation. In: *Parallel Distributed Processing Explorations in the Microstructure of Cognition*. Cambridge, USA: MIT Press, 328-330.

Sani, B. and Kingsman, B.G., 1997. Selecting the best periodic inventory control and demand forecasting methods for low demand items. *Journal of the Operational Research Society*, 48 (7), 700-713.

Schultz, C.R., 1987. Forecasting and inventory control for sporadic demand under periodic review. *Journal of the Operational Research Society*, 38 (5), 453-458.

Silver, E.A., Pyke, D.F., and Peterson, R., 1998. *Inventory Management and Production Planning and Scheduling*. New York, USA: John Wiley & Sons.

Solis, A.O. and Schmidt, C.P., 2009. Stochastic leadtimes in a one-warehouse, N-retailer inventory system with the warehouse actually carrying stock. *International Journal of Simulation and Process Modelling*, 5 (4), 337-347.

Syntetos, A.A. and Boylan, J.E., 2001. On the bias of intermittent demand estimates. *International Journal of Production Economics*, 71 (1-3), 457-466.

Syntetos, A.A. and Boylan, J.E., 2005. The accuracy of intermittent demand estimates, *International Journal of Forecasting*, 21 (2), 303-314.

Syntetos, A.A., Boylan, J.E., and Croston, J.D., 2005. On the categorization of demand patterns. *Journal of the Operational Research Society*, 56 (5), 495-503.

Syntetos, A.A. and Boylan, J.E., 2006. On the stock control performance of intermittent demand estimators. *International Journal of Production Economics*, 103 (1), 36-47.

Syntetos, A.A., Babai, M.Z., Dallery, Y., and Teunter, R., 2009. Periodic control of intermittent demand items: theory and empirical analysis. *Journal of the Operational Research Society*, 60 (5), 611-618.

Syntetos, A.A., Nikolopoulos, K., and Boylan, J.E., 2010. Judging the judges through accuracy-implication metrics: The case of inventory forecasting. *International Journal of Forecasting*, 26 (1), 134-143.

Teunter, R.H., Syntetos, A.A., and Babai, M.Z., 2010. Determining order-up-to levels under periodic review for compound binomial (intermittent) demand. *European Journal of Operational Research*, 203 (3), 619-624.

Tiacci, L. and Saetta, S., 2009. An approach to evaluate the impact of interaction between demand forecasting method and stock control policy on the inventory system performances. *International Journal of Production Economics*, 11 (1), 63-71.

White, H., 1992. Learning and statistics. In: H. White, ed. *Artificial Neural Networks: Approximation and Learning Theory*. Oxford, UK: Blackwell, 79.

Willemain, T.R., Smart, C.N., Shockor, J.H., and DeSautels, P.A., 1994. Forecasting intermittent demand in manufacturing: a comparative evaluation of Croston's method. *International Journal of Forecasting*, 10 (4), 529-538.

Xiang, C., Ding, S.Q., and Lee, T.H., 2005. Geometrical interpretation and architecture selection of MLP. *IEEE Transactions on Neural Networks*, 16 (1), 84-96.

AUTHORS' BIOGRAPHIES

Adriano O. Solis is an Associate Professor of Logistics Management and Management Science at York University, Canada. After receiving BS, MS and MBA degrees from the University of the Philippines, he joined the Philippine operations of Philips Electronics where he became a Vice-President and Division Manager. He later received a PhD degree in Management Science from the University of Alabama. He was previously Associate Professor of Operations and Supply Chain Management at the University of Texas at El Paso. He has published in *European Journal of Operational Research, International Journal of Simulation and Process Modelling, Information Systems Management*, *International Journal of Production Economics*, *Computers & Operations Research*, and *Journal of the Operational Research Society*, among others.

Somnath Mukhopadhyay, an Associate Professor in the Information and Decision Sciences Department at the University of Texas at El Paso, has published in numerous well-respected journals like *Decision Sciences*, *INFORMS Journal on Computing*, *Communications of the AIS*, *IEEE Transactions on Neural Networks*, *Neural Networks*, *Neural Computation*, *International Journal of Production Economics*, and *Journal of World Business*. He received MS and PhD degrees in Management Science from Arizona State University. He was a visiting research assistant in the parallel distributed processing group of Stanford University. He has over 15 years of industry experience in building and implementing mathematical models.

Rafael S. Gutierrez is an Associate Professor in the Industrial, Manufacturing and Systems Engineering Department – of which he was formerly the Department Chair – at the University of Texas at El Paso. He received MS degrees from both the Instituto Tecnologico y de Estudios Superiores de Monterrey (ITESM), in Mexico, and the Georgia Institute of Technology, and a PhD degree in Industrial Engineering from the University of Arkansas. He previously taught at the Universidad Iberoamericana, the ITESM, and the Instituto Tecnologico de Ciudad Juarez. He has been a consultant on enterprise resource planning and materials management for, among others, Emerson Electric, Lucent Technologies, United Technologies, and Chrysler's Acustar Division. He has published in *International Journal of Production Economics*, *Military Medicine*, and *International Journal of Industrial Engineering*.

Monte Carlo Simulation: Remanufacturing or not –A Newsvendor perspective

Rafael Diaz[(a)], Mandar Tulpule[(b)], Francesco Longo[(c)], Antonio Cimino[(d)]

[(a),(b)] VMASC, Virginia Modeling, Analysis, & Simulation Center, Old Dominium University, VA, USA
[(c),(d)] MSC-LES, Mechanical Department, University of Calabria, Italy

[(a)] RDiaz@odu.edu, [(b)]MTulpule@odu.edu, [(c)] f.longo@unical.it [(d)] acimino@unical.it

ABSTRACT

The concept of random yields or uncertain supply is not new in inventory management. While this aspect has been extensively studied in the literature over several decades, the factors impacting uncertain supply have been largely been limited to quality defects, uncertain capacity and sourcing issues. The concept of random yields is very much applicable to the reverse logistics and remanufacturing activities that has received increased attention in recent years. The purpose of this paper is to demonstrate the modeling of the newsvendor problem with uncertain supply via spreadsheet based simulation as a managerial tool for optimizing the performance of a remanufacturing operation.

1. INTRODUCTION

Extensive amount of literature has been developed on the subject of Monte Carlo simulation and Newsvendor problem over the last several decades. The purpose of the paper is to demonstrate the use of spreadsheet simulation for remanufacturing operation from a managerial point of view, and hence we limit the discussion of the mathematical models in this area of study to a brief discussion. The second section of the literature review discusses the literature that has utilized spreadsheets as a means of simulation. Literature pertaining to the advantages of using spreadsheets for simulation and the various applications developed using spreadsheets would be the point of discussion in this section. In the last section we would review the literature addressing the issue of random yields in reverse logistics/remanufacturing scenario.

The classical newsvendor/ newsboy problem also known as the single period problem (SPP) aims at maximizing the expected profit for a single period under uncertain demand scenario. Khouja (1999) has presented a detailed review of literature pertaining to the SPP and classifies the extensions to this problem in eleven categories. One of the categories amongst these eleven is that of 'random yields'. The classical newsvendor problem assumes deterministic supply. The 'random yields' extension relaxes this assumption and considers stochastic supply. This type of situation is particularly applicable to electronic fabrication and assembly, chemical processes and in general processes having an unpredictable quality performance (Yano & Lee, 1995). The concept of random yield was initially introduced by Silver (1976) in which he proposes an extension to the EOQ (Economic order quantity) model by assuming uncertain yields. Yano & Lee (1995) present a comprehensive review of literature dedicated to inventory models assuming random yields. They report that modeling of costs, modeling of yield uncertainty and measures of performance are the three main issues while analyzing inventory systems with random yields. Silver (1976) provides a list of causes that may lead to uncertain supply which includes inventory errors, capacity issues, raw material constraints, quality issues (internal and supplier). The literature can generally be divided depending on the source of supply uncertainty. Uncertain supply due to quality issues have been discussed by Shih (1980), Gurnani et al. (2000), Inderfurth (2004), Maddah and Jaber (2008), Maddah et al. (2009). Capacity issues have been addressed by Erdem and Ozekici (2002), and Hariga and Haouari (1999). Ozekici and Parlar (1999), Keren (2009), Yang et al. (2007), and Rekik et al. (2007) consider sourcing and supply chain factors of random yield while defective inventory records leading to random yield have been discussed by Sahin et al. (2008), Fleisch and Tellkamp (2005), Hesse (2007).

A brief overview of each section is presented as follows. Section 2 and 3 describes some relevant work in spreadsheet simulation and reverse logistics and random yields. Section 4 and 5 presents the research question and decision environment respectively. Section 6 explains and applies the research approach using Monte Carlo simulation viewed as a newsvendor problem in the context of a remanufacturing decision. Section 7 analyzes the output data and provides a specific scenario analysis. Finally, the last section details the conclusions and briefly summarizes the contribution of the paper.

2. SPREADSHEET SIMULATION FOR INVENTORY PROBLEMS

Spreadsheets provide an easy and effective tool for building simulation models. Spreadsheet simulation is especially suited for static simulations such as the Monte-Carlo method. The ready availability and widespread familiarity make spreadsheets a useful platform for developing simulation and decision making tools that are simple and user friendly. Such tools can provide valuable assistance to the managers for decision-making and would help them improve the performance of their operations. Several papers discussing the use of spreadsheet simulation for facilitating the understanding of supply chain concepts amongst students can be found in recent literature. One of the first papers on this subject is Al-Faraj et al. (1991). They address the newsvendor problem and provide a spreadsheet based support system for controlling the inventory and service level. Pfeifer et al. (2001) make a case for using spreadsheet based simulation for teaching the classical newsvendor problem in management schools. They claim that spreadsheet based simulations are more realistic, and are a more suitable way of introducing students to uncertainty and more complicated scenarios as compared to other approached like decision trees and mathematical analysis. Walker (2000) presents a spreadsheet based decision support tool for managing the single period inventory problem. Sezen and Kitapci (2007) also present similar demonstration of spreadsheet based simulation for analyzing a supply chain inventory problem. Evans (2000) proposes the use of spreadsheets for teaching for simulation concepts. He uses the newsvendor problem for demonstrating the methodology of utilizing spreadsheets for constructing such simulation while noting several advantages of using them. Stahl (2005) also provides a spreadsheet based demonstration for the newsvendor problem based on Monte-Carlo simulation and reiterates the usefulness of such application in teaching as done by other authors. Seila (2005) also discusses the usefulness of spreadsheets for constructing simulations. He also provides a demonstration using the inventory model while also putting forth the limitations of the spreadsheets when dealing with complex data structures and algorithms, their slow speed and limited data storage.

3. Reverse logistics and random yields

The number of research papers considering random yield in reverse logistics/remanufacturing activities is comparatively low. Thierry et al. (1995) identify the information on the magnitude and uncertain of the returned products as one of the strategic issues in product recovery management. Guide (2000) has identified, uncertain timing and quantity of returned products as one of the seven characteristics that impact the reverse logistics/remanufacturing production planning and control activities. Ferrer (2003) evaluates the options that a manager of a remanufacturing unit has when faced with random yields and high lead time supplier and evaluates various scenarios comparing the value of information on random yields and having a highly responsive supplier. In more recent works, Bakal and Akcali, (2006) evaluates the effects of random yields on the profitability of an automotive remanufacturing facility where the firm has the power to influence both the supply and demand of the remanufactured products through pricing. Diaz (2010) used random varieties to exemplify the application of Monte Carlo simulation in a shorter version of problem while educating in supply chains and reverse logistics concepts. Longo and Mirabelli (2008), De Sensi et al. (2008), Curcio and Longo (2009), Cimino et al. (2010) used random variates and simulation both in normal supplì chains and in the case of reverse logistics for studying inventory problems.

In the preceding sections, we have reviewed the literature pertaining to random yield in general and to reverse logistics/remanufacturing in particular. We have also reviewed the literature highlighting the importance and usefulness of spreadsheets based simulation. This review emphasizes the overall lack of literature dedicated to developing managerial tools for management of reverse logistic/remanufacturing operations and also points to the spreadsheet based simulation as a suitable platform for developing such a tool. In the next section, we would discuss a research associated to these issues and propose a model to assist the decision-making process in this context.

4. RESEARCH QUESTION

In general, firms that are faced with critical decisions in relation to undertaking significant additional production loads tend to be risk averse. The inherent uncertainty related to resources, demand, and supply, makes the decision difficult. However, most firms are pressured to increase efficiencies while keeping or increasing responsiveness. Thus, the enterprise is expected to find and engage in activities that maintain or increase the difference between revenues and costs. Remanufacturing is an alternative activity that reduces this gap. What is uncertain are the various impacts on profits stemming from decision related to remanufacturing. Thus, the goal is to model and simulate the impacts on profits for the likely near-future electric energy portfolios within Southeastern Virginia

5. DECISION ENVIRONMENT

The objective of this case study is to present a common decision-making situation that can be modeled and simulated using Monte Carlo. This paper considers a hypothetical but realistically represented firm that manufactures a large number of electronic components. As most firms, there are subsets of components that are highly demanded and very profitable. The products analyzed in this paper are assumed to be highly sold in the winter season. The cost of producing each item is $8.15 while sold at $10.00. The volume of sales is 10,000,000 items per season. Returns varies between 3%-6%. Traditionally, items that have been returned to the firm are sold to third party companies that scrap them, and according to market circumstances, these companies dissemble them to be used as raw components. They receive $0.25 apiece. However, the demand for their product has increased exponentially. Expansion of current capacity is on its way. Marketing research indicates that new segments are willing to purchase like-new remanufactured products. Remanufactured items can be sold at $7.85 per item. The forecasted demand has discrete probability distribution according to table 2. The associated costs of retrieving and collecting are $0.50 each Inspecting and dissembling $2.30 per item 50% of this product can be remanufactured. Remanufactured costs are estimated to be $4.25 per item. After the season, unsold remanufactured items can be sold to other forms of secondary markets for 2.50 each. While you don't have to make the decision, you have been asked to analyze the current situation and make a final recommendation on whether or not to remanufacture this item. Based on the probabilities distribution of remanufactured products you have to suggest how many of the next's period items to collect and remanufacture. Because the profits have been extensively studied in the Tidewater region, you know that they are normally distributed.

Table 1: Return Rates

Demand	*Probability*
150	0.3
200	0.2
350	0.3
250	0.15
300	0.05

Table 2: Probabilistic Demand

Supply	*Probability*
400	0.2
350	0.3
600	0.15
300	0.2
500	0.15

6. RESEARCH APPROACH

This research proposes a Modeling and Simulation (M&S) approach to investigate decisions concerning the remanufacturing process. The four basic steps that gird our approach are as follows: 1) represent input data and generate random variates, 2) simulate the supply and demand 3) determine Measures of Performance, and 4) analyze output data.

Monte Carlo simulation produces a probability that an outcome will occur. As a result, multiple sets of trials are required. Since the frequency of possible outcomes can be controlled, this frequency is always known. In this context, we need to mimic how the supply of returned products as well as the demand for them will be. We generate random numbers for the demand and the supply in separated streams. Each stream represents the supply and the demand respectively. A uniform distribution is used to perform the drawing of these random numbers.

Let ξ represent the stochastic supply, C the collection costs, I the inspection costs, $RmPerc\%$ portion of remanufacturable products, RmC the remanufacturing costs, R the revenues, RP the remanufacturing price, Rf the refunds, and P profits generated by the sales of remanufactured components.

The collection and inspection cost, TCC, are determined by:

$$TCC = \xi(\mathrm{C} + \mathrm{I}) \tag{1}$$

The total remanufacturing costs, $TRmC$, is given by:

$$TRmC = \xi * RmC * RmPerc\% \tag{2}$$

The total revenues, TR from remanufactured items sales is given by the Remanufacturing price times the minimum between the supply and demand.

$$TR = RP * \min(\xi, \delta) \quad (3)$$

Refund for unsold items:

$$TRf = Rf * \max(\xi - \delta, 0) \quad (4)$$

Total Profits, P:

$$P = TR + TRf - TCC \quad (5)$$

After calculating the mean, it is usual to accompany this with a confidence interval. It provides the accuracy of the estimate. From statistics: estimated mean (+ /- a multiple of the standard error, so:

$$Confidence\ Interval = \overline{X} \pm (Multiple \times Std.\ Error\ \overline{X})$$

$$Std.\ Error\ \overline{X} = \frac{Std.\ Deviation}{\sqrt{number\ observations}} = \frac{s}{\sqrt{n}}$$

According to normal distribution Multiple for confidence level 95% = 1.96 (Central limit theory). Thus, for 95% confidence:

$$Confidence\ Interval = \overline{X} \pm (1.96 \times \frac{s}{\sqrt{n}}) \quad (6)$$

7. RESULTS

In this exercise, we analyze three possible scenarios that include:

Probability of making any profit P(Profit>0)

$$P(\text{Profit} > 0) = \frac{\text{Count Number of Profits} > 0}{n} \quad (7)$$

Probability of any lost P(Profit<0)

$$P(\text{Profit} < 0) = \frac{\text{Count Number of Profits} < 0}{n} \quad \text{or} = 1 - P(\text{Profit} > 0) \quad (8)$$

Probability of making more profits than the most conservative strategy P(Profit>125)

$$P(\text{Profit} > 125) = \frac{\text{Count Number of Profits} > 125}{n} \quad (9)$$

The estimated average of *current situation* can be calculated by the cost of good sold times the volume, which results in $815. The total revenues is given the current price times the current volume, so we have $1,000. As a result, the difference between revenues and posts provide us with profits, which is estimated to be $125.

The estimated average profit resultant from the Monte Carlo simulation is $278.87. The average profits considering current situation (without remanufacturing) is $125. Since the difference is positive 153, there is no doubt that engaging in remanufacturing activities improves the probabilities of making additional profits. The estimated Confidence Interval (95%) is between 258.11 and 299.63.

Additional information will assist the decision-making process from the risk aversion perspective. We used the three scenario described above to guide our suggestion.

- Probabilities of making any profit = 67.10%
- Probabilities of lost = 32.90%
- Probabilities of more profits than current situation= 54.45%

Given this outcome, the critical question to answer is deciding what to recommend. Firm's risk aversion is intimately related to answering this question.

CONCLUSIONS

Analyzing uncertainty in supply and demand is not new. Research related to these issues has been studied widely. However, most studies focus on quality defects, uncertain capacity, and sourcing issues. Random yields definitions can be related to the reverse logistics and remanufacturing analyses. In this paper, we modeled a common reverse logistic problem using a newsvendor formulation. In the analysis, we used a spreadsheet model to represent not only the stochastic demand, but also the uncertain supply. We used Monte Carlo simulation to determine the effects of several operational constraints on several measures of performance including costs, revenues, and profits. We used a theoretical --but realistic model-- to illustrate such procedures. We applied scenario analysis to study the effects on making profits and obtaining a profit threshold.

We conclude that Monte Carlo Simulation is a well-suited approach to represent probabilistic demand and supply. We determine that Monte Carlo yield insights into various effects on common operational measures of performance such as costs, revenues, and profits. The approach is used in the newsvendor context that assists decision-makers in single period decisions.

REFERENCES

Al-Faraj, T. N., Al-Zayer, J. A., & Alidi, A. S. (1991). A PC-based Spreadsheet Support System for the Newsboy Inventory Control Problem. *International Journal of Operations & Production Management , 11* (10), 58-63.

Bakal, I. S., and Akcali, E. (2006). Effects of Random Yield in Remanufacturing with Price-Sensitive Supply and Demand. *Production and Operations Management , 15* (3), 407-420.

Cimino A., Longo F., Mirabelli G., (2010). Products return flows in inventory management: a general framework based on a supply chain simulator. Proceedings of the *International Society for Inventory Research* (ISIR), 23-25 August, Budapest, Hungary.

Curcio D, Longo F., (2009). Inventory and Internal Logistics Management as Critical Factors Affecting the Supply Chain Performances. International Journal of Simulation & Process Modelling, vol. 5(4); p. 278-288.

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, vol. 46(24); p. 6997-7016.

Diaz, R. (2010). *Educating College Students in M&S Engineering: The Reverse Logistic Project. Report # 10-008.* Virginia: ODU-VMASC.

Erdem, A. S., and Ozekici, S. (2002). Inventory models with random yield in a random environment. *International Journal of Production Economics , 78*, 239-235.

Evans, J. R. (2000). Spreadsheets as a Tool for Teaching Simulation. *INFORMS Transactions on Education , 1* (1), 27-37.

Ferrer, G. (2003). Yield information and supplier responsiveness in remanufacturing operations. *European Journal of Operational Research , 149*, 540-556.

Fleisch, E., and Tellkamp, C. (2005). Inventory inaccuracy and supply chain performance:a simulation study of a retail supply chain. *International Journal of Production Econmics , 95*, 373-385.

Guide, V. R. (2000). Production planning and control for remanufacturing: industry practice and research needs. *Journal of Operations Management , 18*, 467-483.

Gurnani, H., Akella, R., & Lehoczky, J. (2000). Supply management in assembly systems with random yield and random demand. *IIE Transactions* , 701-714.

Hariga, M., and Haouari, M. (1999). An EOQ lot sizing model with random supplier capacity. *International Journal of Production Economics , 58*, 39-47.

Hesse, H. (2007). Inventory Record Inaccuracy, Double Marginalization, and RFID Adoption. *Production and Operations Management , 16* (5), 542-553.

Inderfurth, K. (2004). Analytical Solution for a Single-period Production Inventory Problem with Uniformity Distributed Yield and Demand. *Central European Journal of Operations Research , 12* (2), 117-127.

Keren, B. (2009). The single-period inventory problem: Extension to random yield from the perspective of the supply chain. *Omega , 37*, 801-810.

Khouja, M. (1999). The single-period (news-vendor) problem: literature review and suggestions for future research. *Omega , 27*, 537-553.

Longo, F. and Mirabelli G., (2008). An advanced supply chain management tool based on modeling and simulation. *Computers & Industrial Engineering*, 54(3), 570-588.

Maddah, B., and Jaber, M. Y. (2008). Economic order quantity for items with imperfect quality: Revisited. *International Journal of Production Economics , 112*, 808-815.

Maddah, B., Salameh, M. K., & Karame, G. M. (2009). Lot sizing with random yield and different qualities. *Applied Mathematics Modelling , 33*, 1997-2009.

Ozekici, S., and Parlar, M. (1999). Inventory models with unreliable suppliers in a random environment. *Annals of Operations Research , 91*, 126-136.

Pfeifer, P. E., Bodily, S. E., Carraway, R. L., Clyman, D. R., & Frey, S. C. (2001). Preparing Our Students to Be Newsvendors. *Interfaces , 31* (6), 112-122.

Rekik, Y., Sahin, E., and Dallery, Y. (2007). A comprehensive analysis of the Newsvendor model with unreliable supply. *OR Spectrum , 29*, 207-233.

Sahin, E., Buzacott, J., and Dallery, Y. (2008). Analysis of a newsvendor which has errors in inventory data records. *European Journal of Operational Research , 188*, 370-389.

Seila, A. F. (2005). Spreadsheet Simulation. *Proceedings of the 2005 Winter Simulation Conference.*

Sezen, B., and Kitapci, H. (2007). Spreadsheet simulation for the supply chain inventory problem. *Production Planning & Control , 18* (1), 9-15.

Shih, W. (1980). Optimal inventory policies when stockouts result from defective products.

International Journal of Production Research , *18* (6), 677-686.
Silver, E. A. (1976). Establishing the order quantity when the amount received is uncertain. *INFOR* , *14* (1), 32-39.
Stahl, I. (2005). Using Discrete Event Simulation in the Teaching of Decision Analysis. *Proceedings of the 2005 Winter Simulation Conference.*
Thierry, M., Salomon, M., Van Nunen, J., & Van Wassenhove, L. (1995). Strategic Issues in Product Recovery Management. *California Management Review* , *37* (2), 114-135.
Walker, J. (2000). Decision support for the single-period inventory problem. *Industrial Management and Data Systems* , *100* (2), 61-66.
Yang, S., Yang, J., and Abdel-Malek, L. (2007). Sourcing with random yields and stochastic demand:A newsvendor approach. *Computers & Operations Research* , *34*, 3682-3690.
Yano, C. A., and Lee, H. L. (1995). Lot Sizing with Random Yields: A Review. *Operations Research* , *43* (2), 311-334.

AUTHORS BIOGRAPHIES

Rafael Diaz graduated from the Old Dominion University with a Ph.D. in Modeling and Simulation in 2007, and became a Research Assistant Professor of Modeling and Simulation at Old Dominion University's Virginia Modeling, Analysis, and Simulation Center (VMASC). He holds an M.B.A degree in financial analysis and information technology from Old Dominion University and a B.S. in Industrial Engineering from Jose Maria Vargas University, Venezuela. His research interests include operations research, operations management, production and logistic systems, reverse logistics, dependence modeling for stochastic simulation, and simulation-based optimization methods. He worked for six years as a process engineer and management consultant prior to his academic career.

Mandar Tulpule is currently pursuing a Ph.D. in Modeling and Simulation at the Old Dominion University's Virginia Modeling, Analysis, and Simulation Center (VMASC). He holds a M.E degree in Industrial and Systems Engineering from the North Carolina State University, Raleigh and a B.E in Mechanical Engineering from Pune University, India. His key research interest includes modeling & simulation, operations management, supply chain and logistics. He has two years experience as a manufacturing and supply chain engineer prior to his academic career.

Francesco Longo received his Ph.D. in Mechanical Engineering from University of Calabria in January 2006. He is currently Assistant Professor at the Mechanical Department of University of Calabria and Director of the Modelling & Simulation Center – Laboratory of Enterprise Solutions (MSC-LES). He has published more than 100 papers on international journals and conferences. His research interests include Modeling & Simulation tools for training procedures in complex environment, supply chain management and security. He is Associate Editor of the "Simulation: Transaction of the society for Modeling & Simulation International". For the same journal he is Guest Editor of the special issue on Advances of Modeling & Simulation in Supply Chain and Industry. He is Guest Editor of the "International Journal of Simulation and Process Modelling", special issue on Industry and Supply Chain: Technical, Economic and Environmental Sustainability. He is Editor in Chief of the SCS M&S Newsletter and he works as reviewer for different international journals. His e-mail address is: f.longo@unical.it and his Web-page can be found at http://www.ingegneria.unical.it/impiantiindustriali/index_file/Longo.htm

Antonio Cimino 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. He has published more than 20 papers on international journals and conferences. His research activities concern the integration of ergonomic standards, work measurement techniques, artificial intelligence techniques and Modeling & Simulation tools for the effective workplace design. His e-mail address is: acimino@unical.it and his Web-page can be found at http://www.ingegneria.unical.it/impiantiindustriali/index_file/Cimino.htm

PRODUCTION SYSTEMS DESIGN AND MANAGEMENT: A CASE STUDY ON A HAZELNUTS INDUSTRIAL PLANT

Duilio Curcio[(a)], Francesco Longo[(b)], Giovanni Mirabelli[(c)], Piero Giordano[(d)]

[(a) (b) (c) (d)]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

[(a)] dcurcio@unical.it, [(b)]f.longo@unical.it, [(c)]g.mirabelli@unical.it [(d)]p.giordano@msc-les.org

ABSTRACT
The focus of this paper is an hazelnuts industrial plant located in Calabria (Italy). The objective is to implement a support tool (a simulator) to be used for carrying out specific analyses in order to test system performance under different operative scenarios improving and/or optimizing, if required, system design. After the modeling phase, the simulation model has been verified and validated. Four different performance parameters are introduced to evaluate system behavior in correspondence of different operative scenarios.

Keywords: industrial plant, Modeling & Simulation, performance analysis

1. INTRODUCTION

During the last years several research works in the area of Modeling and Simulation (*M&S*) applied to production systems and industrial plants design and management have been proposed (Callahan *et al.* 2006).

The M&S approach generally does not provide exact or optimal solutions to problems but allows the users to analyze the behavior of complex systems, to perform what-if analysis and correctly choose among alternative scenarios (Karacal 1998; Banks 1998). In fact, oppositely to analytical approaches, the main advantage of M&S when studying and analyzing manufacturing and logistic systems is the possibility to take into consideration multiple aspects without introducing restrictive assumptions. Other advantages of M&S include (Banks 1998):

- understanding why certain phenomena occur in real systems;
- diagnosing problems considering all the interactions which take place in a given moment;
- identifying constraints, e.g. performing bottleneck analysis, it is possible to discover the causes of delays;
- building consensus by presenting design changes and their impact on the real system;
- specifying requirements during the system design.

A state of art overview highlights a great number of research works in the field of M&S for production systems and industrial plants design and management, see Berry (1972), Nunnikhoven and Emmons (1977), Stenger (1996), Mullarkey *et al.* (2000), Longo *et al.* (2005). According to Banks (1998), simulation plays an important role above all for its main property to provide what-if analysis and to evaluate all the benefits and issues related to the environment where it is applied.

As a consequence simulation models are decision support tools adopted by company managers to solve problems. In fact, a simulation model is able to reproduce the evolution of the system taking into consideration several operative scenarios. Simulation models are classified in function of decisions they support. Strategic decisions typically concern production systems and industrial plants design and resources allocation in the medium/long period. Tactical decisions are related to planning and control of production systems and industrial plants in the medium period (weeks or months). Finally, operative decisions concern production systems and industrial plants management in the short period.

The main objective of this paper is to present a simulation model used as decision support tool for investigating the behavior/performance of an industrial plant devoted to produce hazelnuts. Simulation Model development, verification and validation and preliminary analysis are presented. The paper is organized as follows: Section 2 reports a description of the hazelnuts industrial plant; section 3 presents the simulation model as well as verification and validation results while section 4 describes the preliminary analysis and simulation results. Finally, conclusions summarise critical issues and main results of the study.

2. THE HAZELNUTS INDUSTRIAL PLANT

As before mentioned, the production system considered in this research work is located in Calabria, south part of Italy, and manufactures hazelnuts for satisfying the demand of the Pizzo Handmade Ice Cream Consortium, see Cimino *et al.* (2009).

The industrial plant has a rectangular shape with a surface of about 2000 m^2. Figure 1 shows the industrial plant layout (red arrows show the material flow through the different work stations).

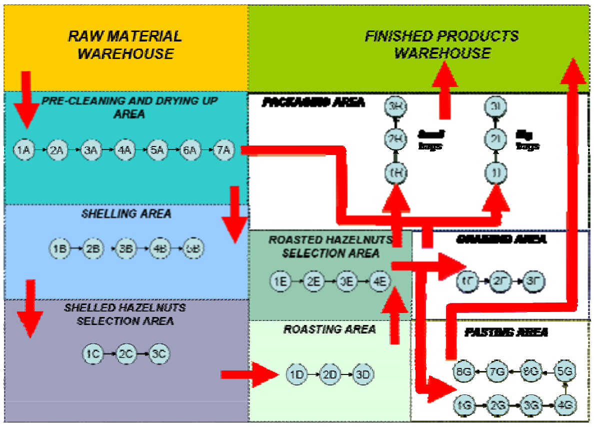


Figure 1: The Layout of the Manufacturing System

According to Figure 1, the plant layout is subdivided in 8 different areas/departments each one including different workstations carrying out the following main operations:

- pre-cleaning;
- drying;
- calibration;
- shelling;
- selection;
- roasting;
- graining;
- pasting;
- packaging (large and small bags).

Figure 2 shows the flow chart of the production process including all the main operations and highlighting the amount of product at the end of each operation.

3. THE SIMULATION MODEL

Based on authors experience (simulation is the most effective tool for designing and analyzing manufacturing systems, industrial plants and supply chain as well (Bruzzone and Longo, 2010; Castilla and Longo, 2010; Cimino et al., 2009; Longo and Mirabelli, 2009; Longo and Mirabelli, 2008). In fact, one of the most important advantages of simulation is to explore and experiment possibilities for evaluating system behavior under internal/external changes.

As a consequence, for a complete scenarios analysis based on a well defined experimental design (i.e. full or fractional factorial experimental design), a specific feature of the simulation model is flexibility. Consider as example Bocca *et al.* (2008); the authors implement a simulation model of a real warehouse highlighting the importance of building flexible simulation models for carrying out experimental analysis. Cimino et *al.* (2009), Longo and Mirabelli (2008) use flexible simulation model to analyze the performance of real manufacturing systems and supply chains by monitoring multiple performance measures under multiple system configurations and constraints. In the next section the implementation of the simulation model is briefly described.

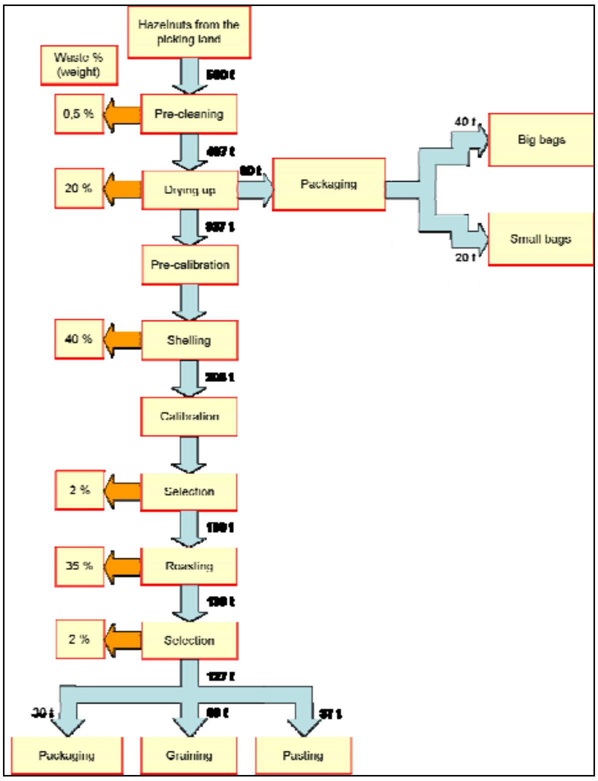


Figure 2: The production process flow chart

3.1. The production system processes modeling

The simulation model presented in this research work reproduces all the most important processes and operations of the hazelnuts industrial plant. The software tool adopted for the simulation model implementation is the commercial package Anylogic™ by *XJ Technologies*.

In particular, for reproducing all the logics and rules used within the industrial plant and for increasing model flexibility, different classes are implemented by using software libraries objects and ad-hoc Java routines. The simulation model is in two parts: the flow chart (or structure diagram) and the animation.

The flow chart displayed in Figure 3 recreates system structure and contains software libraries objects opportunely connected and integrated in order to reproduce with high accuracy the flow of entities (raw material, semi-finished or finished products and workers) through the model.

More in detail, entities defined in the simulation model can be classified into *static* and *dynamic* entities.

Static entities (or resources) belong to specific areas of the model supporting dynamic entities that pass through. From the other side, dynamic entities represent the objects flowing through different classes of the simulation model (workstations of the real manufacturing system). As a consequence, in the simulation model implemented static entities are represented by workers while hazelnuts are the dynamic entities.

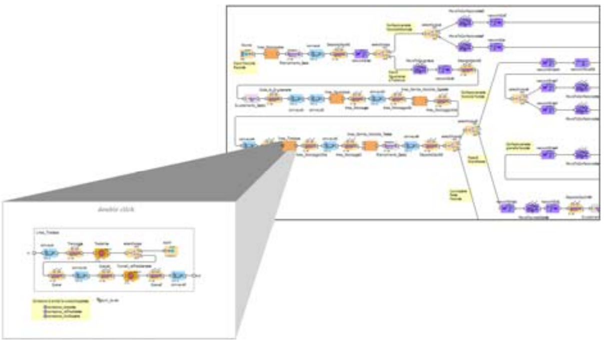


Figure 3: The Simulation Model Structure Diagram

Figure 4 shows the simulation model animation which faithfully reproduces the hazelnuts flow in the real system.

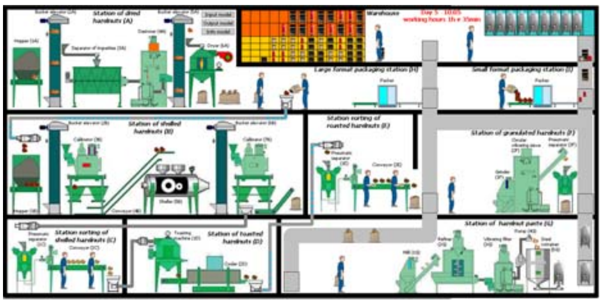


Figure 4: The Simulation Model Animation

3.2. The Graphic User Interface

The main variables of the simulation model are completely parameterized in order to reproduce different operative scenarios. To this end the authors developed a dedicated Graphic User Interface (*GUI*) with a twofold functionality:

- to increase the simulation model flexibility changing its input parameters both at the beginning of the simulation run and at runtime (by using sliding bars, buttons and check boxes) observing the effect on the system behaviour (*Input Section*);
- to provide the user with all simulation outputs for evaluating and monitoring system performances (*Output Section*).

The *Input Section* reported in Figure 5 is subdivided in five different subsections:

- the *Industrial Plant parameters* section in which, for each department, the productive capacity of machines and intermediate buffers capacity can be modified;
- the *Consumption of raw material* section which contains the parameters related to the quantity of hazelnuts to be processed and their arrivals frequency;
- the *Workers* section in which the number of workers to be allocated in each department can be easily selected;
- the *Work shifts* section in which the user can decide for each production line/department the work shifts (up to three work shifts per day);
- the *Products mix* section in which the production mix can be defined.

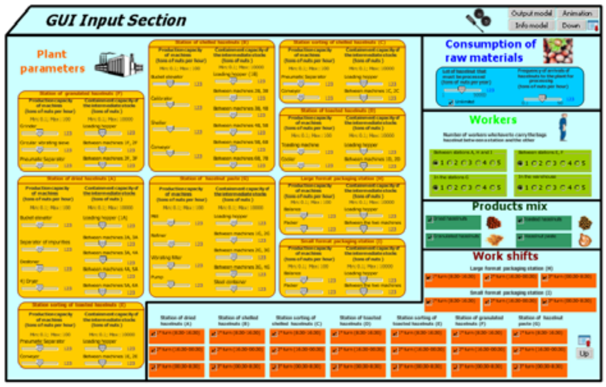


Figure 5: The GUI Input Section

The *Output Section* provides the user with the simulation outputs to evaluate and monitor the industrial plant performances. According to Figure 6 the output section is subdivided in three different subsections:

- the *Plant production* section in which the quantity of dried, roasted, grained hazelnuts and hazelnuts paste is displayed;
- the *Packages* section in which the number of packages for each product is reported;
- the *Plant performance* section in which the performance of the whole industrial plant is monitored. Furthermore, for each department, output data related to machines average utilization level and buffers saturation level can be collected.

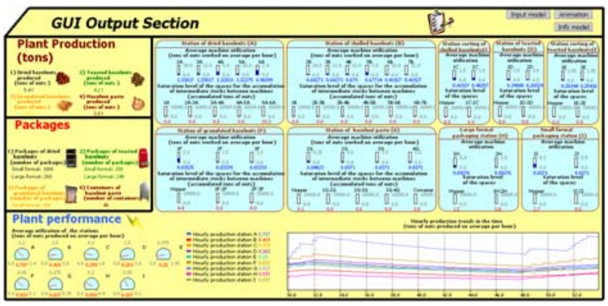


Figure 6: The GUI Output Section

4. MODEL VERIFICATION AND VALIDATION

Verification is the process of determining that a model implementation accurately represents developer conceptual description and specifications (Balci 1998).

The simulation model verification has been made using the debugging technique. As explained in Dunn (1987), debugging is an iterative process that aims at finding and eliminating all the bugs due to model translation. The model is opportunely modified and tested (once again) for ensuring errors elimination as well as for detecting new errors. All the methods (routines written in Java) have been iteratively debugged line by line, detecting and correcting all the errors.

4.1. The Validation

Validation is the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended use of the model (Balci 1998). Data used for simulation model validation regard an historical period of 5 years, from January 2005 to December 2009.

In order to evaluate simulation data accuracy, four different statistical indexes are introduced: the Root Mean Squared Error (*RMSE*), the Mean Absolute Error (*MAE*), the Modeling Efficiency (*EF*) and the Coefficient of Residual Mass (*CRM*).

In particular, the RMSE and MAE indexes are calculated according to Fox (1981):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}} \tag{1}$$

$$MAE = \frac{\sum_{i=1}^{n}| P_i - O_i |}{n} \tag{2}$$

in which P_i represents values estimated by the model and O_i are values observed on the real system. MAE is less sensitive to extreme values than RMSE. The lower are these indexes, the higher is the model accuracy.

The other two indexes, the EF and CRM, are calculated using the following formulas (Loague and Green 1991):

$$EF = \frac{\sum_{i=1}^{n}(O_i - \overline{O})^2 - \sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2} \tag{3}$$

$$CRM = \frac{\sum_{i=1}^{n} O_i - \sum_{i=1}^{n} P_i}{\sum_{i=1}^{n} O_i} \tag{4}$$

in which $\bar{O}$ is the average value of observations on the real system. The optimal value for EF is 1; values greater than 0 indicate that model estimated values are better than the average of the observations while negative values confirm that the average of observations is a better estimator of model accuracy. The optimal value for CRM is 0; positive values indicate that model underestimates measured data while negative values indicate the opposite.

In order to assure the goodness of simulation model statistic results each simulation run has been replicated 5 times so P_i are the average values of each run.

Let us consider results of the validation on dried hazelnuts annual production.

Table 1: Validation on Dried Hazelnuts Annual Production

year	P_i *(t/year)*	O_i *(t/year)*	*RMSE*	*MAE*	*EF*	*CRM*
1	66	63,04	29,71	17,72	0,99	0,03
2	415,4	420,29				
3	220	210,14				
4	700,4	765,60				
5	512,6	506,88				

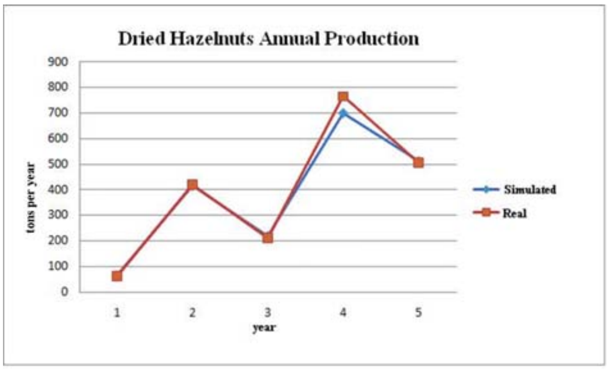


Figure 7: Dried Hazelnuts Annual Production (real and simulated)

According to Table 1, RMSE, MAE, EF, CRM values are good estimators of simulation data accuracy. Moreover, Figure 7 shows real and simulated curves of dried hazelnuts annual production (with different industrial plant setting and production mix every year): these curves are nearly similar so the simulation model is an accurate representation of the real system. Figures 8–9–10 report validation results for roasted, grained hazelnuts and hazelnuts paste (again each year the production mix is different in order to test simulator capability in different operative scenarios).

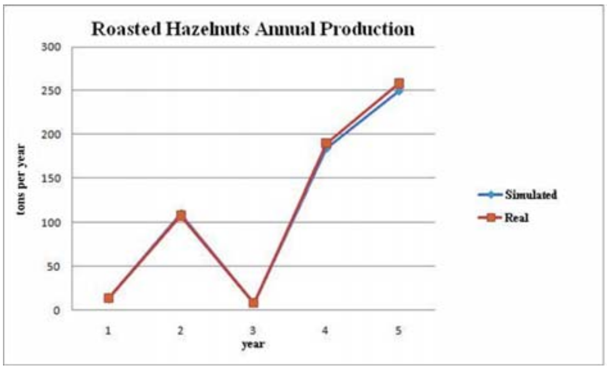


Figure 8: Roasted Hazelnuts Annual Production (real and simulated)

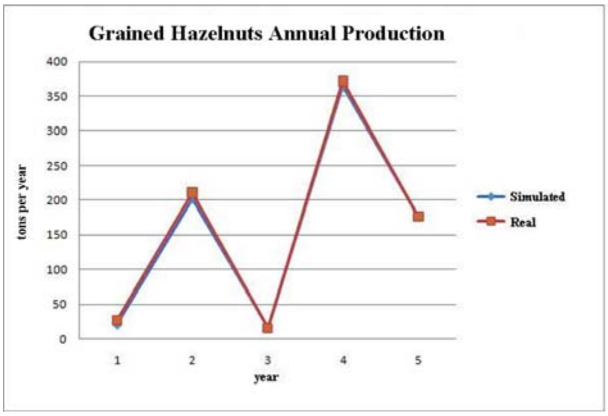


Figure 9: Grained Hazelnuts Annual Production (real and simulated)

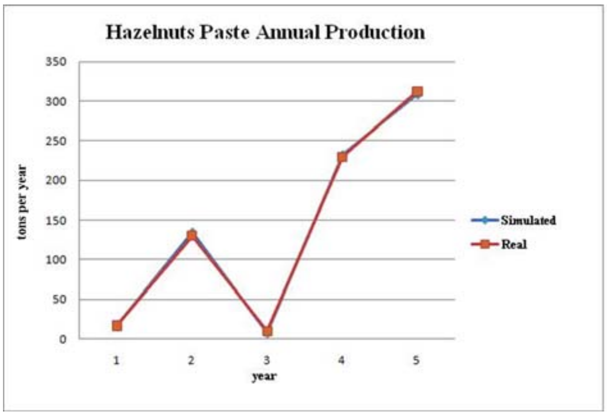


Figure 10: Hazelnuts Paste Annual Production (real and simulated)

5. DESIGN OF EXPERIMENTS AND SIMULATION RESULTS ANALYSIS

As before mentioned, the objective of this research work is to implement a simulation model to be used for carrying out specific analyses in order to test system performance under different operative scenarios improving and/or optimizing, if required, its design.

More in detail, the authors analyze system performance through four different performance parameters and by changing the pre-cleaning, roasting and roasted hazelnuts selection departments productive capacity keeping constant all the remaining parameters/variables. As reported in Table 2, each productive capacity is expressed as percentage of the actual value.

Table 2: Factors and Levels

Factors	L1	L2	L3
Pre-cleaning productive capacity	90%	100%	110%
Roasting productive capacity	90%	100%	110%
Roasted Hazelnuts Selection productive capacity	90%	100%	110%

The four different performance parameters introduced are:

- P_1 related to machines average utilization level (UL_i), see Equation 5;
- P_2 evaluated as the ratio between the intermediate stocks of hazelnuts in tons (*WH*) and the tons of hazelnuts to be processed (*WIP*) as reported in Equation 6;
- P_3 calculated as the ratio between tons of raw hazelnuts (*IN*) and tons of dried, roasted, grained hazelnuts and hazelnuts paste produced (*OUT*) as showed in Equation 7;
- P_4 which is a global system performance estimator, see Equation 8.

$$P_1 = \frac{\sum_{i=1}^{n} UL_i}{n} \tag{5}$$

$$P_2 = \frac{WH}{WIP} \tag{6}$$

$$P_3 = \frac{IN}{OUT} \tag{7}$$

$$P_4 = \frac{P_1 + (1 - P_2) + P_3}{3} \tag{8}$$

Simulation results, for each factors levels combination, are reported in Tables 3–4–5. In particular, the following scenarios have been analyzed:

- comparison of the 90%, 100% and 110% scenarios in terms of pre-cleaning productive capacity;
- comparison of the 90%, 100% and 110% scenarios in terms of roasting productive capacity;
- comparison of the 90%, 100% and 110% scenarios in terms of roasted hazelnuts selection productive capacity.

For each scenario the four different performance parameters have been monitored. Table 3 reports the simulation results under different pre-cleaning productive capacity.

Table 3: Simulation results under different pre-cleaning productive capacity

Scenarios	P_1	P_2	P_3	P_4
90% Pre-cleaning productive capacity	0,539	0,997	0,020	0,187
100% Pre-cleaning productive capacity	0,516	0,995	0,207	0,243
110% Pre-cleaning productive capacity	0,926	0,373	0,980	0,844

Considering the P_1 and P_2 parameters, the first and the second scenarios shows a similar behavior while the third scenario provides a better behavior for these parameters and for the global system performance. In fact, the increase of pre-cleaning productive capacity means the increase of the machines utilization level for this production line and, as a consequence, the addition of raw hazelnuts in input.

Table 4 reports the simulation results under different roasting productive capacity. Also in this case, the P_1 and P_2 parameters have a similar value in the first and second scenarios while the third scenario provides the worst behavior for the P_2 parameter.

Table 4: Simulation results under different roasting productive capacity

Scenarios	P_1	P_2	P_3	P_4
90% Roasting productive capacity	0,377	0,798	0,016	0,199
100% Roasting productive capacity	0,413	0,796	0,165	0,261
110% Roasting productive capacity	0,833	0,336	0,882	0,793

Table 5 shows the simulation results for the roasted hazelnuts selection productive capacity.

Table 5: Simulation results under different roasted hazelnuts selection productive capacity

Scenarios	P_1	P_2	P_3	P_4
90% Roasted Hazelnuts Selection productive capacity	0,647	0,897	0,022	0,257
100% Roasted Hazelnuts Selection productive capacity	0,677	0,896	0,165	0,316
110% Roasted Hazelnuts Selection productive capacity	0,787	0,186	0,735	0,779

Such scenario investigates how system performance changes under different roasted hazelnuts selection productive capacity. In this case the global system performance increase passing from 90% to 100% roasted hazelnuts selection productive capacity is about 30% while the best results in terms of global system performance is related to 110% roasted hazelnuts selection productive capacity.

6. CONCLUSIONS

A simulation model of a hazelnuts industrial Plant, its implementation, verification and validation are presented. Preliminary analysis to investigate system behavior under different factors levels combinations are carried out. In particular, four different performance measures are introduced in order to evaluate system performances under different operative scenarios. Changes in factors levels highlights the tendency of the system to over-react with major changes in some of the performance measures therefore stressing the importance to use the simulator to tune the system correctly to improve system efficiency.

REFERENCES

Balci, O., 1998. Verification, Validation and Testing. In: J. Banks, *Handbook of Simulation*, John Wiley & Sons, Inc., 335–393.

Banks, J., 1998. *Handbook of Simulation*. John Wiley & Sons, Inc.

Berry, W.L., 1972. Priority scheduling and inventory control in a job shop lot manufacturing systems. *AIIE Transactions*, 4, 267–276.

Bocca, E., Curcio, D., Longo, F., Tremori, A., 2008. Warehouse and internal logistics management based on modeling & simulation. *Proceedings of the International Workshop on Modeling & Applied Simulation 2008*, pp. 41–48. September 17–19, Campora S.Giovanni, Cosenza, Italy.

Bruzzone, A.G., Curcio, D., Mirabelli, G., Papoff, E., 2009. Warehouse management: inventory control policies comparison. *Proceedings of the International Workshop on Modeling & Applied Simulation 2009*, pp. 148–154. September 23–25, Puerto de La Cruz, Tenerife, Spain.

Bruzzone A.G, Longo F., 2010. An advanced System for supporting the decision process within large scale retail stores. SIMULATION, ISSN: 0037-5497, doi: 10.1177/0037549709348801.

Callahan, R.N., Hubbard, K.M., Bacoski, N.M., 2006. The use of simulation modelling and factorial analysis as a method for process flow improvement. *Advanced Manufacturing Technology*, 29, 202–208.

Castilla I., Longo F., 2010. Modelling and Simulation Methodologies, Techniques and Applications: a State of the Art Overview. International Journal of Simulation & Process Modelling, vol. 6(1); p. 1-6, ISSN: 1740-2123.

Cimino A., Longo F., Mirabelli G., 2009. A Multi-Measures Based Methodology for the Ergonomic Effective Design of Manufacturing System Workstations. INTERNATIONAL JOURNAL OF INDUSTRIAL ERGONOMICS, vol. 39; p. 447-455, ISSN: 0169-8141, doi: 10.1016/j.ergon.2008.12.004

Cimino, A., Curcio, D., Longo, F., Mirabelli, G., 2009. *The hazelnuts production in Calabria*. University of Calabria.

Cimino, A., Curcio, D., Mirabelli, G., Papoff, E., 2008. Warehouse inventory management based on fill rate analysis. *Proceedings of the International Workshop on Modeling & Applied Simulation 2008*, pp. 23–30. September 17–19, Campora S.Giovanni, Cosenza, Italy.

Curcio D, Longo F., 2009. Inventory and Internal Logistics Management as Critical Factors Affecting the Supply Chain Performances. *International Journal of Simulation & Process Modelling,* vol. 5(4); p. 278-288, ISSN: 1740-2123

Dunn, R.H., 1987. The quest for software reliability. In: *Handbook of Software Quality Assurance*, Van Nostrand Reynold, 342–384.

Eben-Chaime, M., Pliskin, N., Sosna, D., 2004. An integrated architecture for simulation. *Computer and Industrial Engineering*, 46, 159–170.

Fox, D.G., 1981. Judging air quality model performance. *Bulletin of the American Meteorological Society*, 62, 599–609.

Karacal, S.C., 1998. A novel approach to simulation modeling. *Computers & Industrial Engineering*, 34 (3), 573–587.

Loague, K., Green, R.E., 1991. Statistical and graphical methods for evaluating solute transport models: overview and application. *Contaminant Hydrology*, 7 (1-2), 51–73.

Longo, F., Mirabelli, G., Papoff, E., 2005. Material flow analysis and plant lay-out optimization of a manufacturing system. *International Journal of Computing*, 1 (5), 107–116.

Longo F., Mirabelli G., 2009. Effective Design of an Assembly Line using Modeling & Simulation. *Journal of Simulation,* vol. 3; p. 50-60, ISSN: 1747-7778, doi: 10.1057/JOS.2008.18.

Longo F., Mirabelli G., 2008. An Advanced Supply Chain Management Tool Based on Modeling & Simulation. COMPUTERS & INDUSTRIAL ENGINEERING, vol. 54/3; p. 570-588, ISSN: 0360-8352, doi: 10.1016/j.cie.2007.09.008.

Mullarkey, P., Gavirneni, S., Morrice, D.J., 2000. Dynamic output analysis for simulations of manufacturing environments. *Proceedings of the 2000 Winter Simulation Conference*, pp. 1290–1296. December 10–13, Orlando, Florida (USA).

Nunnikhoven, T.S., Emmons, H., 1977. Scheduling on parallel machines to minimize two criteria related to job tardiness. *AIIE Transactions*, 3, 288–296.

Stenger, A.J., 1996. Reducing inventories in a multi-echelon manufacturing firm: a case study. *Production Economics*, 45, 239–249.

AUTHORS BIOGRAPHIES

DUILIO CURCIO was born in Vibo Valentia (Italy), on December the 15th, 1981. He took the degree in Mechanical Engineering from University of Calabria in 2006 and received the PhD in Industrial Engineering at the same University in February 2010. His research activities include Modeling & Simulation and Inventory Management theory for production systems and Supply Chain design and management. Now he is collaborating with MSC-LES at the Industrial Engineering Section of the University of Calabria to research projects for supporting Research and Development in SMEs.

PIERO GIORDANO was born in Cosenza (Italy), on August the 14th, 1982. He took the bachelor's degree in Management Engineering from University of Calabria in 2007. Now he is completing his graduate studies in Management Engineering at the same University.

FRANCESCO LONGO was born in Crotone (Italy), on February the 08th, 1979. He took the degree in Mechanical Engineering from University of Calabria (2002). He received the PhD in Industrial Engineering (2005). He is currently researcher at the Mechanical Department (Industrial Engineering Section) of University of Calabria and scientific responsible of the Modeling & Simulation Center – Laboratory of Enterprise Solutions (MSC-LES) in the same department. His research interests regard modeling & simulation of manufacturing systems and supply chain, DOE, ANOVA.

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.

COORDINATION OF ACTIVITIES: APPLICATION OF SOME CONCEPTS AND FORMALIZATIONS TO AGRICULTURAL SYSTEMS SIMULATION

Z. Afoutni [a, b], R. Martin-Clouaire [c], R. Courdier [b] and F. Guerrin [a, c]

[a] CIRAD, Environmental Risks of Recycling Research Unit, F-97408 Saint-Denis, Reunion, France
[b] University of La Réunion, Computer Science & Mathematics Lab., F-97490 Ste-Clotilde, Reunion, France
[c] INRA, Biometrics and Artificial Intelligence Research Unit, F-31326 Castanet-Tolosan, France

[a, b] zoubida.afoutni@cirad.fr, [c] rmc@toulouse.inra.fr, [b] remy.courdier@univ-reunion.fr, [a, c] guerrin@cirad.fr

ABSTRACT
Coordination is defined as the management of dependencies between activities in order to reach an objective. These dependencies generally concern resource sharing and the compliance with temporal (simultaneity, precedence) and spatial constraints. This management is made according to two principal modes: explicit and implicit. The explicit mode is based on using 'protocols' (e.g. procedures, plans) explicitly describing how the agents must perform their actions to guarantee the good functioning of the system they are committed to. The implicit mode, characterized by the absence of protocol, is mainly based on using artefacts, implicitly fostering the behaviour of the agents through their interaction with their physical environment (concepts of 'stigmergy' and 'affordance'). In this paper we try to synthesize theories and relevant concepts necessary to represent coordination. Our goal is to propose, at last, a modelling framework to simulate the coordination of human activities in complex agricultural production systems.

Keywords: activity coordination, planned action, situated action, action representation, resource allocation, agricultural system modelling.

1. INTRODUCTION

Agricultural production systems (APS) are made of interacting components among which human agents performing interdependent activities. These activities need resources of material, financial or human natures to be accomplished. They may be characterized by their temporal dimension (i.e. start and end dates, duration) along with their spatial dimension (i.e. they occur at determined locations). The agents interact through the activities they perform and their effects on the environment, altogether contributing to the attainment of some desirable system's goal, as defined by the system's designer or manager. One of the questions posed is how to manage these interactions? This is crucial to be able to propose which is our aim: a modelling framework enabling one to represent and simulate interacting farming activities at operations level (Guerrin, 2009). Many authors (Malone and Crowston, 1994; Schmidt and Simone, 1996; Whang, 1995) have dealt with the issue of 'coordination' in various domains, namely, computer-supported cooperative work systems and supply chain management. This article is a tentative synthesis of some existing theories and concepts about coordination. Before all, we need make more precise two terms which meanings are too often confused using the definitions given by Clancey (2002):

- Task: "a specification of work … to be performed";
- Activity: "how people actually work within the constraints of their environment".

Hence, to denote farming practices, i.e. what is actually done by the farmers, we use 'activity', conceived as a complex set of coordinated elementary actions.

This article is organized as follows. In Section 1 we define the concept of coordination with various types of dependence between them. Sections 2 and 3 are devoted to two coordination modes: explicit and implicit. In Section 4, we introduce two formal tools enabling activity coordination in APS's to be represented, i.e. Allen's temporal logic (Allen, 1984) and the modelling framework of action by Guerrin (2009).

2. WHAT IS COORDINATION?

Malone and Crowstone (1994) have defined coordination as the management of interdependencies between activities performed by one or more agents, necessary to attain a goal that can be common or not. When several agents share the same objective these authors speak of coordinating cooperative activities. They distinguish between two main types of dependences: (i) resource sharing and (ii) temporal dependences (simultaneity and precedence). To these two we propose to add (iii) spatial dependence.

2.1. Resource sharing

Because resources in all kinds of production systems are limited they may be required by several actions at the same time or exhausted by previous ones. Consequently, resource allocation is necessary to avoid conflicting situations or to foster some preferred actions against others.

2.2. Temporal dependencies

Some actions must be performed simultaneously as driving a tractor while spreading manure (simultaneity). Others, conversely, should not: for a same crop in a same field, ploughing must take place before sowing, and sowing before harvest (precedence). These temporal dependencies among actions can be represented and simulated using the formal tools described in Section 5: Allen's temporal relations (Allen, 1984) and Guerrin's framework (2009).

2.3. Spatial dependencies

To our knowledge, this aspect is scarcely dealt with in the literature whereas an action should also be characterized by the place it is executed. A production system is, very generally, composed of productive units located at different places; e.g. a farm or a set of farms with several fields and livestock enterprises scattered over a territory. Obviously, necessary resources for action must be disposed at the right place at the right time. Thus, an agent likely two perform two actions at two different places should schedule one before the other. A precedence constraint must hence be added as well as a third intermediary action, that of agent transportation (with possibly other necessary resources) from the first location to the second. Assume a farmer must deliver a product to two buyers at a given place and due date. It is hence necessary, not only to coordinate those three people in time (i.e. synchronize them) but also in space in order they meet at the right place and date. The relative locations of the productive units are also important to be accounted for as they can strongly determine the agents' actions. For example, a farmer having made something on a field can perform in the sequel another action on a neighbouring field to spare time and transport.

Once the temporal and spatial dependencies among activities are determined, the manager has to find a way to coordinate them in both dimensions. This can be made according to two modes, explicit and implicit.

3. EXPLICIT COORDINATION

In explicit coordination, technical facilities are implemented to clarify how the agents should execute their activities. These are 'artefacts' (e.g. document, blackboard) jointly used with 'protocols' (e.g. rules, procedures, plans) prescribing the ways of acting (Schmidt and Simone, 1996). Although also a means of coordination in day-to-day life, conventions (arbitrary habits) are not considered here as they seem less relevant for APS. Artefacts are used to share information among agents as material supports to coordination protocols. This mode of coordination can take two modalities: centralized and distributed. The latter can take two perspectives: the team, which members pursue a common objective, and the market, in which, by letting each agent pursue his own objective, Adam Smith's 'invisible hand' makes the system converge to equilibrium (Whang, 1995).

3.1. Coordination protocols

3.1.1. Rule

A rule is a statement prescribing a determined behaviour as an injunction, a prohibition, or even a simple recommendation (Batifoulier, 2001). It is generally accompanied by an explicit threat of sanction and, so, must be justified to enabling the application of a penalty in the case of non-compliance. Observing a rule is accomplished through a judgment made contextually by the agents. Hence it needs a common representation of the situations at hand.

3.1.2. Procedure

Two issues inherently linked to interpretation appear to execute a rule (Kechidi, 2005):

- How assess the situation to decide if it matches with the rule premises?
- How select the rule to be triggered when several are candidates in a given situation?

Triggering a rule needs in fact to reduce its subjective interpretability giving it a stronger prescriptive feature, that is specifying precisely which behaviour is required, preferred or prohibited in determined contexts. When such a rule exists it is a 'procedure' (Kechidi, 2005).

3.1.3. Plan

Planning is an emblematic sub-domain of Artificial Intelligence which aims, as one of the theories of action, at answering the question "What should be done?" That are: Which actions are to be performed? In which order? In its more classical sense, a plan is a sequence of actions capable to drive a system from its current state to a final desired state called a goal. In executing the plan, an action is triggered as soon as its conditions are met. A plan can encompass alternative conditional paths to cope with external events.

3.2. Modalities of explicit coordination

3.2.1. Centralized coordination

The production system is here managed by a unique coordinator endowed with roles of observation, information collection and decision-making (Li and Wang, 2007). The information relevant to it is about the dependencies among activities, the system states and external observed processes (e.g. market or climate evolutions).

In this case, the protocol is often an action plan specifying the sequence of actions to be performed over time and the resources needed. To design this plan, the first step is to identify the precedence and simultaneity constraints among activities and those sharing the same resources (Malone et Crowstone, 1994). The second step consists in determining resource allocation rules. For example, as some activities are critical and must be performed within specified time-windows, they may be assigned higher priorities to get the resources they need in time; other activities, owing to be executed in parallel should be given their resources at the same time to avoid delays. The coordinator may be obliged to revise the plan in cases unexpected situations appear.

3.2.2. Distributed coordination

This modality of explicit coordination is characterized by the absence of a central coordinator: management is thus shared by all agents. Two perspectives may be distinguished: the team and market perspectives.

3.2.2.1. Team perspective

Each agent has limited information on the system and must coordinates its own activities by communicating with other agents to achieve their common goal. This can be made through plan exchanges, according to the "Partial Global Planning" approach (PGP) described by Ferber (1995) and involving three types of plans:

- Local plans for managing each agent's own activities.
- 'Node-plans', synthesizing the sole relevant information in local plans to be exchanged with others.
- PGP's, gathering all the information relevant for each agent about its own and others' activities.

Consubstantial to this perspective is the notion of 'cooperation', which "usually implies shared goals among different actors" (Malone and Crowstone, 1994).

3.2.2.2. Market perspective

Contrary to the team perspective, which members share a common goal, each market agent pursues its own goal, the coordination with other agents emerging naturally from the functioning of the whole. In some cases, coordination may be based on contracts among stakeholders (Whang, 1995). This system prevails in supply/demand APS such as a set of farms collectively managing their wastes on a territory scale (see application to livestock waste management in Courdier et al., 2002). This perspective generally coexists with others: a production system can be coordinated internally according to a centralized mode and externally by the market with many other firms.

4. IMPLICIT COORDINATION

Another coordination mode, called implicit or reactive, also exists based on concepts of 'stigmergy' and 'affordance' allowing actions to be coordinated without specifying protocols.

4.1. Stigmergy

This way of coordination stems from the research by Pierre-Paul Grassé on ants colonies (Susi and Ziemke, 2001). The general principle of stigmergy is as follows: every ant wanders randomly in its environment searching for food. As soon as a food source is found it goes back directly to its nest, dropping on its way back pheromone traces so that other ants may found them and follow the path until the food heap. These new ants, doing the same, reinforce the path gradually. Ants thus use their environment to communicate by the means of pheromone droppings let on their way. Using modifications of the environment to influence other agents behaviour is stigmergy (Susi and Ziemke, 2001).

We can try to generalize this concept to human activity when the result of an agent's action influences the behaviour of other agents. For example, consider two neighbouring farmers that use to help each other. One has crop fields and the other livestock. The fact the first one has completed the harvest of some crops may be a signal for the second bring manure on these fields. Observing the heaps left on their edges, the first farmer may be fostered to spread this manure within the next few days. Stigmergy is obviously an implicit means of coordination as it allows an indirect communication between agents based on the persistence of effects of past activities in the environment to determine activities in the future.

4.2. Affordances

Another concept, called affordance, can be used as coordination means. It originally emerged from the works made by Gibson (1979) on human vision in the field of Ecological Psychology, whose goal is to explain how an individual adapts to its environment. An 'affordance' is the perception of possibilities of action that are "offered" by objects in the agent's environment. It allows an immediate adaptation of the individual perceiving it in the form of a response action. In a sense, the artefacts used in the theory of stigmergy could be considered as affordances fostering agents to commit to some specific action. With affordances and stigmergy, the activities of agents are not determined by protocols but by an evolving space of possibilities in which they navigate and choose, at any time, the action to commit to. Being confronted permanently to concurrent solicitations from the environment poses, nevertheless, the issue of how individuals select the one which will make them act. This has been a criticism addressed to Gibson's by authors like Reed (1996). In the case of APS this dilemma is solved by the farmer's experience, knowledge and memory that will make him/her focus on some signs rather than others: although many affordances can be generated by a tractor (that can allow various works: ploughing, sowing, transporting,…) he/she will select the one corresponding to its current priority (e.g. if he has already prepared the soil, he may sow).

5. FORMAL TOOLS OF REPRESENTATION

5.1. Allen's temporal logic

Allen's formalism (1984) is based on 7 binary relations (and their inverses, omitted here for simplicity) between any pair of temporal intervals (T_i,T_j):

- DURING(T_i,T_j): T_i is fully contained within T_j;
- STARTS(T_i,T_j): T_i shares the same start date as T_j, but ends before T_j;
- FINISHES(T_i,T_j): T_i shares the same end date as T_j, but begins after T_j;
- BEFORE(T_i,T_j): T_i lies before T_j with no overlap;
- OVERLAPS(T_i,T_j): T_i starting before T_j overlaps it;
- MEETS(T_i,T_j): T_i ends exactly when T_j starts;
- EQUAL(T_i,T_j): T_i and T_j are superimposed.

These relations are mutually disjoints (if one holds for two intervals, no other holds) and complete (given any two intervals always one relation holds). This formalism is useful to represent and manage the essential temporal dependencies among actions: simultaneity (DURING, STARTS, FINISHES, EQUAL) and precedence (BEFORE, MEETS). It is used in the modelling framework of action proposed by Guerrin (2009) to simulate human activities in APS.

5.2. Dynamic simulation of action at operations level

In Guerrin's (2009) framework activities are considered as complex coordinated set of actions. Every action is represented as a dynamic process determined by conditions stemming from observed processes of various kinds (including other actions). It is characterized by a start date and an end date or duration. Hence, actions are actually represented in the same way as temporal intervals, making the use of Allen's primitives natural. In the sequel we present some aspects of the mathematical formalization that will be used to deal with coordination representation.

5.2.1. Representation of action

Every action A is represented by a binary function of time:

$$S_A(t) = \begin{cases} 1 & \text{if } \; C_A(t) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Where $C_A(t)$ is a logical proposition evaluated true or false whether its value is respectively 1 or 0. Therefore, an action A defines a sequence of temporal intervals during which its value is 1 or 0.

5.2.2. Temporal bounds of action

The start and end date (t^-, t^+) and the duration ($\tau_A(t)$) of an action are also functions of time, determined according to a condition $P^-_A(t)$ (resp. $P^+_A(t)$):

$$t^{\pm}_A(t) = \begin{cases} t & \text{if } \; P^{\pm}_A(t) \\ t^{\pm}_A(max(0, t-\tau_S)) & \text{otherwise} \end{cases} \quad (2)$$

Where τ_S is the simulation time-step, $P^-_A(t)$ (resp. $P^+_A(t)$) is a logical proposition function of time specified according to any process $X(t)$ on which events relevant to trigger or stop an action are possibly detected. It may be a clock, a schedule, a biophysical process, or even another action.

Note that, as a minimal condition $C_A(t) \equiv (t^-_A > t^+_A)$: A holds as long as $P^-_A(t)$ is true and stops as soon as an event occurs to stop it, i.e. when $P^+_A(t)$ becomes true and an end date is set.

5.2.3. Coordination of actions

5.2.3.1. Specification of complex activities

As an illustration, consider two cultural schedules of two market garden crops, carrot and potato, each being cultivated by two farmers on two different plots. Tables 1 and 2 show these schedules for each crop, the work time and the equipment necessary to each operation.

Table 1: Cultural schedule of carrot

Operations	Period	W time	Equipment
Soil disinfection (DC)	Oct.-Apr.	5 days	Sprayer
Tillage (TC)	Mar.-Jun.	9 days	Plough
Sowing (SC)	Apr.-Jul.	18 days	Seeder
Hoeing (HC)	Jun.-Sept.	5 days	Hoe
Harvest (AC)	Jul.-Nov.	19 days	Carrot harvester

Tableau 2: Cultural schedule of potato

Operations	Period	W time	Equipment
Tillage (TP)	Apr.-May	10 days	Plough
Planting (PP)	Apr.-May	6 days	Potato planter
Hoeing-Ridging (HP)	May-Aug.	3 days	Hoe-Ridger
Harvest (AP)	Aug.-Nov.	5 days	Potato harvester

Here we consider only the constraints linked to material resources. Often farmers rent together with neighbours heavy expansive equipments to save costs and, so, must set a common schedule of utilization. Here we assume two farmers having each a tractor and equipments specific to their own crop are sharing the same plough.

The holding condition for each cultural operation (= action), according to Eq. 1, is true for a crop whenever the current time is within its feasibility period and necessary resources (equipment, labour) are available. Hence in this example the start date $t_A^-(t)$ of each action must verify the following condition:

$$t_P^- \le t_A^-(t) < t_P^+ - \tau_A \qquad (3)$$

Where, with values given in Tables 1 and 2, τ_A is the duration of action A and t_P^- (resp. t_P^+) is the opening (resp. closing) date of each feasibility period, that is the earliest start date (resp. the latest end date) of A.

In each schedule are found precedence constraints among operations. E.g., for carrot, soil disinfection must precede tillage which, in turn, must precede sowing. Hoeing must be done while plants are growing, i.e. between sowing and harvest. Delays between two consecutive operations should be adapted (e.g. it is preferable not to let a bare soil too long). Operations may also different priority: here we take potato with greater priority than carrot. Figure 1 displays a solution for combining these two cultural as a Gantt diagram.

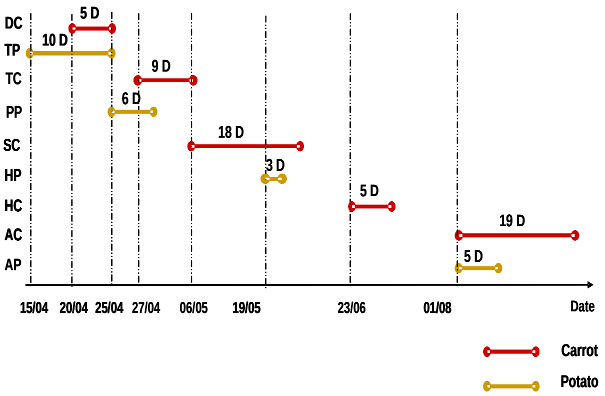


Figure 1: Combination of cultural schedules for carrot and potato for 2 neighbouring farmers sharing a plough.

The resulting schedule can be represented with Allen's relations (cf. notations in Tables 1 and 2) starting from operation DC (soil disinfection in carrot):

- DC
- TP: FINISHES(DC,TP)
- TC: BEFORE(TP,TC),
- PP: MEETS(TP,PP) & OVERLAPS(PP,TC)
- SC: BEFORE(PP,SC) & MEETS(TC,SC)
- HP: DURING(HP,SC)
- HC: BEFORE(HP,HC)
- AC: BEFORE(HC,AC)
- AP: STARTS (AP, AC).

Guerrin (2009) has shown how to simulate such specification of activities.

5.2.3.2. Resource allocation

The issue of resource allocation is posed whenever several actions require the same resources at the same time or when they are exhausted due to a previous action. This may be solved by allocating priorities to actions according to their critical nature in the system. This may lead to cancel or delay some non-priority actions or to execute concurrently actions with same priorities. The priority degrees are given as constants according to an arbitrary numerical scale or as relevant dynamic variables taken in the system (see Guerrin, 2009 for details).

6. DISCUSSION AND CONCLUSIONS

In this article we tackled the issue of coordination based on existing theories namely by identifying two modes: explicit and implicit.

The first mode is based on defining ex-ante the actions to be performed as protocols to be used together with communication artefacts. Be they a rule, procedure or plan, protocols are a way of specifying actions coordination, namely by enabling all agents to anticipate the behaviour of the others. However, protocols (and particularly plans), as necessary means for action, have been strongly criticized by many authors (Selznick, 1948; Suchman, 1987; Clancey, 2002) considering they cannot completely define action in the real practice. For this, it is necessary to take into account the inherent 'situated' dimension of action (Suchman, 1987). According to this theory, every action stems mainly from the dynamical interaction of agents with their environment. Hence, the notion of 'protocol' is theoretically inconsistent with the necessary improvisation an agent must implement to adapt to its changing context. If protocols do not allow one to determine completely and coordinate actions, what are their role? Could there be other means of coordination? According to Schmidt and Simone (1996), all kinds of protocols can play two different roles:

- "Weak": a guide as a "codified set of functional requirements which provides a general heuristic framework".
- "Strong": a script offering "a pre-computation of interdependencies among activities (options, sequential constraints, temporal constraints, etc.) which, for each step, provides instructions to actors of possible or required next steps".

The protocol role, be it weak or strong, depends on agents capacity in determining in advance the dependencies among actions. To determine these dependencies one must anticipate the actions to be performed, which is only partially the case in APS's.
Therefore, another coordination mode, implicit or reactive, must be considered, based on the concepts of 'stigmergy' or 'affordance', both enabling a "protocol-free" coordination of agents. In that respect, this mode seems also appropriate to APS. Stigmergy and affordances allow agents to coordinate implicitly, in

both temporal and spatial dimensions by adapting dynamically to their environment. Coordination is in that case made in a distributed way: each agent, endowed with its own perceptive and interpretative abilities, reacts individually and the global coherence emerges, eventually, of the whole. However, the question is posed whether these concepts can also be useful to solve conflicts on resource sharing which are generally dealt with by establishing protocols.

Fundamentally, we came to the conclusion that dealing with the issue of coordination needs to take into account both the temporal and the spatial dimensions of activities. If 'synchronization' can be taken as a synonymous for 'coordination in the temporal dimension', we did not find yet any equivalent concerning the spatial dimension. There exist indeed formalizations that can be used to take into account the temporal dimension of action. This is the case for Allen's work (1984) that cannot be ignored, and most of formalizations quoted in the excellent state-of-the-art on temporal reasoning made by Chittaro et Montanari (2000). However, it must be emphasized that all of these formalizations deal with time in an essentially static mode: they rather allow one to reason about courses of action already made or planned than about action while it is being made. For this, beyond our own works in simulation modelling of APS where human activities are explicitly represented (Guerrin, 2009; Martin-Clouaire and Rellier, 2009) little formalizations exist. In fact activity is often ignored in production system simulation. However, the spatial dimension is lacking in our approaches and we found very little literature on this topic, except about Schelling's focal point (Morel, 2004). This concept, stemming from Schelling's own practical experience tries to tackle the following issue: how two individuals knowing each other lost in a foreign city may found them without communicating? The answer is based on the common knowledge of the participants that can allow, both anticipating the other's solution, eventually, make them converge to a common place. This notion seems to be interesting to explore, however it does not correspond well to the problematic we defined in §2.3 where locations are generally known and agents must navigate among them to act according to various global and local constraints altogether with the time course...

These aspects will be dealt with in our future works, our ambition being, eventually, to propose a modelling framework of human activities applicable to APS. We think to explore, particularly, implicit coordination through the use of artefacts following the stigmergy and affordance theories, as steps to approximate the nice idea of "situated action" promoted by Suchman (1987).

REFERENCES

Allen, J., 1984. Towards a general theory of action and time. *Artificial Intelligence*, 23, 123–154.

Batifoulier, P., 2001. *Théorie des conventions*. Paris: Economica.

Chittaro, L., Montanari, A., 2000. Temporal representation and reasoning in Artificial Intelligence: Issues and approaches. *Annals of Mathematics and Artificial Intelligence*, 28, 47–106.

Clancey, W., 2002. Simulating activities: Relating motives, deliberation, and attentive coordination. *Cognitive Systems Research*, 3, 471–499.

Courdier, R., Guerrin, F., Andriamasinoro, F.H., Paillat, J.-M., 2002. Agent-based simulation of complex systems: Application to collective management of animal wastes. *Journal of Artificial Societies and Social Simulation*, 5 (3), 30–56.

Ferber, J., 1995. Coordination d'actions. In: J. Ferber, *Les systèmes multi-agents, vers une intelligence collective*, Paris : InterEditions, 422–492.

Gibson, J., 1995. *The ecological approach to visual perception*. Boston, MA, USA: Houghton Mifflin.

Guerrin, F., 2009. Dynamic simulation of action at operations level. *Autonomous agents and multi-agents systems*, 18 (1), 156–185.

Li, X., Wang, Q., 2007. Coordination mechanisms of supply chain systems. *European Journal of Operational Research*, 179, 1-16.

Malone, T.W., Crowston, K., 1994. The interdisciplinary study of coordination. *ACM Computing Surveys*, 26 (1), 87–119.

Martin-Clouaire, R., Rellier, J.-P., 2009. Modelling and simulating work practices in agriculture. *Int. J. Metadata, Semantics and Ontologies*, 4, 42–53.

Morel, C., 2004. Variation sur la négociation tacite et le point focal de Thomas Schelling. *Négociations*, 1(1), 15-26.

Kechidi, M., 2005. Une taxinomie des figures collectives de coordination de l'action : de la convention à la routine. *Cahier du GRES*, 2005-07, Univ. Bordeaux 4 & Toulouse 1, 16 p.

Reed, E., 1996. Encountering the world: Toward an Ecological Psychology. New-York: Oxford University Press.

Selznick, P., 1948. Foundations of the Theory of Organization. *American Sociological Review*, 13, 25–35.

Schmidt, K., Simone, C., 1996. Coordination mechanisms: Toward a conceptual foundation of CSCW systems design. *Computer supportive cooperative work: The Journal of Collaborative Computing*, 5, 155–200.

Suchman, L., 1987. Plans and situated actions. The problem of human-machine communication. Cambridge: Cambridge University Press.

Susi, T., Ziemke, T., 2001. Social cognition, artefacts, and stigmergy: A comparative analysis of theoretical frameworks for the understanding of artefact-mediated collaborative activity. *Journal of Cognitive Systems Research*, 2, 273–290.

Whang, S., 1995. Coordination in operations: A taxonomy. *Journal of Operations Management*, 1, 413–422.

STOCHASTIC VENDOR SELECTION PROBLEM

Baohua Wang[(a)], Shiwei He[(b)], Sohail S. Chaudhry[(c)]

[(a)]IBM China Research Lab, Beijing, China 100094
[(b)]Department of Transportation Management Engineering, College of Traffic and Transportation, Beijing Jiaotong University, Beijing, China 100044
[(c)]Department of Management and Operations/International Business, Villanova School of Business, Villanova University, Villanova, PA USA 19085

[(a)] baohuaw@cn.ibm.com, [(b)] shwhe@bjtu.edu.cn, [(c)] sohail.chaudhry@villanova.edu

ABSTRACT

We study the vendor selection problem in which capacity, quality level, service level, and lead time associated with each vendor are considered to be stochastic. The problem is modeled as a stochastic dependent-chance programming model. As stochastic programming models are difficult to solve by traditional methods, a hybrid adaptive genetic algorithm, which embeds the neural network and stochastic simulation, was designed and implemented. To further improve the performance of the algorithm, the adaptive genetic algorithm was adjusted by varying the crossover probability and mutation rate according to the stage of evolution and fitness of the population. The solution procedure was tested on several randomly generated problems with varying parameters. Our extensive computational experience on these problems indicates that the hybrid adaptive genetic algorithm has strong adaptability on the tested problems as the algorithm converged more rapidly than the simple genetic algorithm.

Keywords: Vendor selection problem, stochastic dependent-chance programming, genetic algorithm

1. INTRODUCTION

For most manufacturing firms, the purchasing of raw materials and component parts from outside vendors constitutes a major expense. In a vendor selection problem (aka the supplier selection problem in literature), the purchasing manager must choose order quantities to place among vendors in a multi-sourcing network to satisfy the demand of the firms under a variety of conditions. Generally, the selected vendors need to be evaluated on more than a single criterion. Dickson (1966) studied the vendor selection problem and reported that there are 23 factors that are important to purchasing managers when selecting vendors whereas Dempsey (1978) identified 18 criteria. Among these criteria, it was found that price, delivery, quality, and capacity were ranked at the top of the list of purchasing managers when they selected vendors. A similar conclusion was reported by Weber et al. (1991) using a review of 74 related papers for the vendor selection problem.

The focus of this paper is to analyze the vendor selection problem under stochastic environment. We propose a stochastic dependent-chance programming model which aims to maximize the probability that the demand of the firm can be satisfied while minimizing the expected cost. Since stochastic programming model is hard to solve by traditional methods, a hybrid intelligent algorithm, which integrates stochastic simulation, neural network into genetic algorithm, is designed to solve the problem.

The paper is organized as follows. In the next section, a literature review is presented. In the following section, we present some basic concepts related to dependent-chance programming based on Liu (1997) and the stochastic dependent-chance programming model for vendor selection problem. This is followed by design of the hybrid genetic algorithm for the model. To further improve the performance of the algorithm, an adaptive genetic algorithm is also presented in this section. In the penultimate section, we present the application of the hybrid genetic algorithm to a series of randomly generated problem instances. The conclusions and future research direction are presented in the final section.

2. LITERATURE REVIEW

During the past 50 years or more, many different methods have been proposed to solve a variety of different the vendor selection problems. These methods can be grouped into whether the technique focuses on qualitative or quantitative factors that are relevant in the vendor selection problem. However, in the recent years, researchers have developed solution approaches that based are based on two or more of these methodologies.

Wind and Robinson (1968), Mazurak et al. (1985), Cooper (1977) and others have used a weighted linear method of multiple criteria for this problem. Timmerman (1986) and Gregory (1986) linked this approach to a matrix representation of data and Narasimhan (1983) employed the analytical hierarchical process to generate weights for such models. Recently, Micheli (2008) investigated supplier selection problem as a way to mitigate the overall supply risk. A risk efficiency-based supplier selection approach was developed for critical supplies that allowed a decision maker to consider the procurement-related “risk” and “investment” for mitigation/exploitation interventions. Kirytopoulos et al. (2008) analyzed the supply chain processes within the pharmaceutical industry in Greece in which analytic network process based method for the selection was used to solve the problem. Saen (2008) addressed the supplier ranking in a volume discount environment and introduced an innovative approach which was based on the super-efficiency analysis.

Besides qualitative methods, there is an abundance of published research that utilizes quantitative methods, which may also be integrated with qualitative methods to solve this problem. The earliest papers that utilized quantitative methods to solve the vendor selection can be traced back to Stanley et al. (1954) and Gainen (1955) in which linear programming was used for awarding contracts to contractors at the Department of Defense. Bender et al. (1985) proposed a mixed integer programming (MIP) model for vendor selection problem. This approach was used at IBM to select vendors and their order quantities with the objective to minimize purchasing, inventory, and transportation costs; however, the specific mathematical formulation was not presented. Pan (1989) developed a single item linear programming model to allocate order quantities among suppliers in which the objective was to minimize aggregate price that was restricted by the constraints on quality, service level, and lead-time. Sharma et al. (1989) suggested a goal programming formulation that considered price, quality and lead-time goals with demand and budget constraints.

Chaudhry et al. (1993), Degraeve et al. (2000), De Boer et al. (2001), and Aissaoui et al. (2007) also have provided a well-structured literature survey on the application of different techniques to the vendor selection problem. More recently, Chen and Huang (2007) related product characteristics to supply chain strategy and adopted supply chain operations reference (SCOR) performance metrics as the decision criteria. A scheme integrated analytic hierarchy process (AHP) with bi-negotiation agents based on the multi-criteria decision-making approach and software agent techniqueis then developed to take into account both qualitative and quantitative attributes in supplier selection. Ting and Cho (2008) developed a two-step decision-making procedure utilizing analytic hierarchy process and multi-objective linear programming in which analytic hierarchy process was used for select candidate supplier and multi-objective linear programming was used to allocate the quantities among the selected suppliers

However, in solving practical vendor selection and purchasing plans, businesses are faced with some uncertain factors. For example, the quantity supplied by the vendors, the quality level, and the service level of the vendors sometimes can be considered to be random variables with known distribution function. Kasilingam and Lee (1996) considered the stochastic nature of demand and propose a mixed-integer programming model to select vendors and determine the order quantities. Shiromaru et al. (2000) treated coal purchase planning in a real electric power plant and applied a fuzzy satisfying method to deal with the vagueness of the goals. Kumar et al. (2005) presented a fuzzy multi-objective integer programming model and discussed the corresponding crisp equivalence for optimization.

In a recent work, Rezaei and Davoodi (2006) formulated a fuzzy mixed integer programming model of a multi-period inventory lot sizing problem with supplier selection. Amid et al. (2006) firstly developed a fuzzy multiobjective model in which different weights can be considered for various objectives. This fuzzy model enabled the purchasing managers not only to consider the imprecision of information but also take into consideration the limitations of buyer and supplier into account in order to calculate the order quantity assigned to each supplier. Liao and Rittscher (2006) considered the demand quantities and timing uncertainties into consideration and proposed a multi-objective supplier selection model. A genetic algorithm was utilized to handle this model. Sevkli et al. (2007) proposed an analytical hierarchy process weighted fuzzy linear programming model for supplier selection and compared this new model with the classical analytic hierarchy process. Amid et al. (2007) developed a fuzzy multi-objective model for the supplier selection problem under price breaks and presented a weighted additive method to generate an optimal solution in the fuzzy environment. Chan et al. (2008) discussed the fuzzy based analytic hierarchy process to efficiently tackle both quantitative and qualitative decision factors involved in selection of global supplier in current business scenario. Wu and Olson (2008) considered three types of vendor selection methodologies in supply chains with risk, which are chance constrained programming, data envelopment analysis, and multi-objective programming models. The Monte-Carlo simulation was applied to these three methodologies. He et al. (2008) developed a class of special stochastic chance-constrained programming models and presented a genetic algorithm for vendor selection problem under stochastic environment.

3. STOCHASTIC DEPENDENT-CHANCE PROGRAMMING MODEL FOR VENDOR SELECTION PROBLEM

In practice, the decision-maker may want to maximize the chance functions of some events (i.e., the probabilities of satisfying the events) under stochastic environment. In order to model this type of stochastic decision system, Liu (1999) provided a new type of stochastic programming, called dependent-chance programming. Dependent-chance programming involves maximizing chance functions of events in an uncertain environment.

According to Liu (1999), an uncertain environment signifies the following stochastic constraint, $g_j(x,\xi) \le 0$, $j = 1,2,3\cdots\cdots p$, where x is a decision vector, and ξ is a stochastic vector. Also, let an event be represented by $h_k(x,\xi) \le 0$, $k = 1,2,3\cdots\cdots q$, whose chance function is defined as the probability measure of the event, $f(x) = \Pr\{h_k(x,\xi) \le 0, k = 1,2,3\cdots\cdots q\}$, subject to the uncertain environment as defined above.

In deterministic model, expected value model, and chance-constrained programming, the feasible set is essentially assumed to be deterministic after the real problem is modeled. That is, an optimal solution is given regardless of whether it can be performed in practice. However, the given solution may be impossible to perform if the realization of uncertain parameter is unfavorable. Thus dependent-chance programming theory never assumes that the feasible set is deterministic. In fact, it is constructed in an uncertain environment.

Formally, a typical dependent-chance programming model can be represented as maximizing the chance function of an event subject to an uncertain environment in the following way:

$$\max\ \Pr\left\{h_k(x,\xi) \le 0, k = 1,2,3\cdots\cdots q\right\}$$
$$s.t.$$
$$g_j(x,\xi) \le 0, j = 1,2,3\cdots\cdots p,$$

where x is an n-dimensional decision vector, ξ is a random vector of parameters, the system $h_k(x,\xi) \le 0$, $k = 1,2,3\cdots\cdots q$, represents an event, and the constraints $g_j(x,\xi) \le 0$, $j = 1,2,3\cdots\cdots p$ are an uncertain environment.

3.1. Notation and mathematical model

Let the decision variable x_i represent the percentage of the quantity to be ordered from vendor i. In addition, let the parameters be defined as:

D Total demand of the item;

Ω Set of vendors competing for selection, $\Omega = \{1,2,3\cdots\cdots N\}$, $i \in \Omega$;

c_i Unit cost of purchasing plus transportation from vendor i;

d Unit cost due to receiving poor quality items;

e Unit cost due to receiving late delivered items;

ξ_i Upper limit of the quantity available for vendor i, random variable;

λ_i Percentage of good items supplied by vendor i, random variable;

η_i Percentage of items receiving good after service offered by vendor i, random variable;

τ_i Percentage of the late delivered items by the vendor i, random variable;

S Minimum allowable aggregate quantity of items receiving good after service (Required service level);

L Maximum allowable aggregate quantity of late delivered items (Required lead-time level);

W Minimum allowable aggregate quantity of good items (Required quality level);

B Budget constraint.

In our model we assume that quantity discounts are not allowed. There is only one item to be considered. However, multi-item vendor selection problem can be simplified into several single-item vendor selection problems. The maximum number of vendors which can be selected is not restricted. Finally, all the random variables are independent.

3.2. Dependent-chance programming model

Given the definitions, assumptions and notations above, the vendor selection problem can be formulated as the following dependent-chance integer goal programming model.

1) Constraints

$$D \cdot x_i \le \xi_i \text{ for all } i \tag{1}$$

Constraint (1) puts restrictions due to the maximum capacity of the vendors.

$$\sum_i (D \cdot x_i) \cdot \lambda_i \ge W \tag{2}$$

Constraint (2) means that the required quality level should be achieved.

$$\sum_i (D \cdot x_i)\tau_i \le L \tag{3}$$

Constraint (3) means that the required lead-time level should be achieved.

$$\sum_i (D \cdot x_i)\eta_i \ge S \tag{4}$$

Constraint (4) means that the required service level should be achieved.

$$\sum_i (D \cdot x_i)\cdot c_i + (D \cdot x_i)\cdot(1-\lambda_i)\cdot d + (D \cdot x_i)\cdot e \cdot \tau_i \le B \tag{5}$$

Constraint (5) puts restrictions on the budget.

$$x_i \ge 0 \text{ for all } i \tag{6}$$

Constraint (6) ensures the non-negativity of the solution.

2) Objective Function

$$\max\ \Pr\left\{\sum_i x_i = 1\right\} \tag{7}$$

$$\min\ E\left(\sum_i (D \cdot x_i)\cdot c_i + (D \cdot x_i)\cdot d \cdot (1-\lambda_i) + (D \cdot x_i)\cdot e \cdot \tau_i\right) \tag{8}$$

The objectives are to maximize the probability that the demand can be satisfied and minimize the total expected cost under stochastic environment, which is characterized by constraint (1) – (6). The probability can also be considered to the reliability, or the risk, of the purchasing plan. The reason why we take the total expected cost into consideration is that with the same probability there can be more than a single purchasing plan and the decision-maker would like to know which one is the best. We define μ as the weight coefficient. We also assume the priority of objective (7) is higher than objective (8). So μ should be a sufficiently large positive number. The two objectives can be integrated into one objective by the following equation:

$$\max\ \mu \cdot \Pr\left\{\sum_i x_i = 1\right\} - E\left(\sum_i (D\cdot x_i)\cdot c_i + (D\cdot x_i)\cdot d\cdot(1-\lambda_i) + (D\cdot x_i)\cdot e\cdot\tau_i\right) \quad (9)$$

When some management targets are given, the objective function may minimize the deviations, positive, negative, or both, with a certain priority structure set by the decision-maker. In this paper, if we let α denote the chance of meeting the demand given by the decision-maker, then the objective function of the dependent-chance goal programming model can be formulated as follow, where d^- and d_2^+ is to be minimized:

$$\min\ \left\{d_1^-, d_2^+\right\}$$

$$\Pr\left\{\sum_i x_i = 1\right\} + d_1^- - d_2^+ = \alpha \quad (10)$$

$$E\left(\sum_i (D\cdot x_i)\cdot c_i + (D\cdot x_i)\cdot d\cdot(1-\lambda_i) + (D\cdot x_i)\cdot e\cdot\tau_i\right) + d_2^- - d_2^+ = B \quad (11)$$

4. HYBRID GENETIC ALGORITHM

Generally speaking, stochastic programming models are difficult to solve by traditional methods. It has been shown that a good way to solve these difficult problems is to design hybrid intelligent algorithms (Liu and Iwamura 1997, Liu 1997, Liu 2000). In this section, we integrate the neural network, stochastic simulation, and genetic algorithm to produce a hybrid intelligent algorithm for solving stochastic dependent-chance programming models of vendor selection problem, which is formulated by equations (1) – (7).

According to Liu (1997)'s study, equation (7) is equivalent to the uncertain function as follow:

$$f_1(x) = \Pr\left\{\begin{array}{l} \sum_i x_i = 1 \\ D\cdot x_i \le \xi_i \text{ for all } i \\ \sum_i (D\cdot x_i)\cdot\lambda_i \ge W \\ \sum_i (D\cdot x_i)\cdot\tau_i \le L \\ \sum_i (D\cdot x_i)\cdot\eta_i \ge S \\ \sum_i (D\cdot x_i)\cdot c_i + (D\cdot x_i)\cdot d\cdot(1-\lambda_i) + (D\cdot x_i)\cdot e\cdot\tau_i \le B \\ x_i \ge 0 \text{ for all } i \end{array}\right\}.$$

Equation (8) is equivalent to the uncertain function as follow:

$$f_2(x) = E\left[\sum_i (D\cdot x_i)\cdot c_i + (D\cdot x_i)\cdot d\cdot(1-\lambda_i) + (D\cdot x_i)\cdot e\cdot\tau_i\right].$$

Given a certain $\mathbf{x} = \left\{(x_1, x_2 \cdots x_N) \middle| \sum_i x_i = 1, x_i \ge 0 \text{ for all } i\right\}$, the value of $f_1(x)$ and $f_2(x)$ may be estimated by the following stochastic simulation.

4.1. Algorithm (Stochastic Simulation)

Step 1. Set $N' = 0$, $Cost = 0$.

Step 2. Generate ξ_i, λ_i, τ_i, η_i according to the their distribution function.

Step 3. $Cost = Cost + E[\bullet]$. If constraints (2) – (5) can be satisfied, then $N'++$.

Step 4. Repeat the second to fourth steps for N times, where N is a sufficiently large number.

Step 5. $f_1(x) = N'/N$, $f_2(x) = Cost/N$.

Although stochastic simulations are able to compute the chance functions, we need relatively simple functions to approximate the uncertain functions because the stochastic simulations are a time-consuming process. In order to speed up the solution process, a neural network is employed to approximate the chance functions since the neural network has the ability to approximate the uncertain functions by using the training data, it can compensate for the error of training data (all input-output data obtained by stochastic simulation are clearly not precise), and has the high speed of operation after they are trained. Hence, the hybrid genetic algorithm is presented next.

4.2. Algorithm (Hybrid Genetic Algorithm)

Step 1. Generate training input-output data for the chance function $f_1(x)$ and $f_2(x)$ by stochastic simulation (Algorithm 4.1).

Step 2. Train a neural network to approximate the chance function according to the generated training data.

Step 3. In the paper, we use the floating vector to represent a solution in which each chromosome vector is coded as a vector of floating numbers, of the same length as the solution vector. Let $V=(x_1,x_2\cdots x_N)$ be the chromosome representing the solution $x=(x_1,x_2\cdots x_N)$. We assume all the vendors have the same priority. Then the chromosomes should be initialized by the following manner.

Step 3.1 Define $Total=0$;

Step 3.2 Choose a vendor i randomly. The quantity purchased from vendor i, x_i, is initialized by generating a random number q in $(0,E(\xi_i))$.

Step 3.3 If $x_i=0$, $x_i=x_i+q$. Otherwise, if $x_i+q>E(\xi_i)$, $x_i=E(\xi_i)$;

Step 3.4 $Total=Total+x_i$. If $Total>D$, $x_i=x_i-(Total-D)$; Otherwise, if $Total<D$, go to Step 3.2.

Repeat the algorithm above pop_size times, we can obtain pop_size chromosomes.

Step 4. Compute the fitness of all chromosomes $V_k, k=1,2\cdots pop\_size$ by the trained neural network to rearrange them from best to worse according to their objective function values.

Step 5. Select the chromosomes by spinning the roulette wheel.

Step 6. Renew the chromosomes $V_k, k=1,2\cdots pop\_size$ by crossover operation.

We define a parameter P_c of a genetic system as the probability of crossover. This probability gives us the expected number $P_c\cdot pop\_size$ of chromosomes undergoing the crossover operation. In order to determine the parents for crossover operation, we can generate a random real number r from the interval $[0,1]$. If $r<P_c$, the chromosome V_i is selected as a parent.

We denote the selected parents by $V_1^{'},V_2^{'},V_3^{'},\ldots$ and divide them into the following pairs: $((V_1^{'},V_2^{'}),(V_3^{'},V_4^{'}),(V_5^{'},V_6^{'})\cdots\cdots$. Let us illustrate the crossover operator on each pair by using the pair $(V_1^{'},V_2^{'})$. Initially, a random number c is generated from the open interval $(0,1)$. Then, the crossover operator on $V_1^{'}$ and $V_2^{'}$ will produce two children X and Y as follows:

$$X=c\cdot V_1^{'}+(1-c)\cdot V_2^{'},\ Y=(1-c)\cdot V_1^{'}+c\cdot V_2^{'}$$

We must check the feasibility of each child before accepting it, and only replace the parents with the feasible children.

Step 7. Update the chromosomes $V_k, k=1,2\cdots pop\_size$ by mutation operation.

We define a parameter P_m of a genetic system as the probability of mutation. This probability gives us the expected number of $P_m\cdot pop\_size$ of chromosomes undergoing the mutation operations. The mutation operation will be carried out as the following manner, which is similar to Gaussian Mutation.

For each selected parent, denoted by $V=(x_1,x_2,...,x_N)$, we randomly generated N real positive numbers, $r_1,r_2\cdots r_N$, with the distribution $U(0,x_i)$.

Step 7.1 Randomly choose x_i for mutation.

Step 7.2 Randomly generated a number ω from $(0,1)$.

If $\omega>0.5$, $x_i=x_i+r_i$; else, $x_i=x_i-r_i$;

Step 7.3 If $\sum_i x_i>1$, adjust x_i to make sure $\sum_i x_i=1$. Then the mutation operation is over; else, go to step 7.1.

Step 8. Repeat the third to sixth steps for a given number of cycles.

Step 9. Report the best chromosome as the optimal solution.

Based on the algorithm above, an adaptive genetic algorithm (AGA), in which the probability of the crossover and mutation operation will be adjusted according to the stage of evolution and fitness of the population, is used to improve the performance. In AGA, we define f_{avg} as the average fitness of the population, f as the fitness of the chromosome, $P_{c-\max}$ as the maximum probability of crossover, $P_{c-\min}$ as the minimum probability of crossover, $P_{m-\max}$ as the maximum probability of mutation, $P_{m-\min}$ as the minimum probability of mutation, $MaxGen$ as the maximum generation of the algorithm, gen is the current generation of the algorithm.

So in every generation, the probability of crossover and mutation can be obtained by following equations:

$$P_c = \begin{cases} P_{c-\max} - \left(\dfrac{P_{c-\max} - P_{c-\min}}{MaxGen} \right) \bullet gen, f > f_{avg} \\ P_{c-\max}, f \le f_{avg} \end{cases}$$

(12)

$$P_m = \begin{cases} P_{m-\max} - \left(\dfrac{P_{m-\max} - P_{m-\min}}{MaxGen} \right) \bullet gen, f > f_{avg} \\ P_{m-\max}, f \le f_{avg} \end{cases}$$

(13)

5. COMPUTATIONAL EXPERIENCE AND DISCUSSION

In this section, we apply the hybrid genetic algorithm to a series of instances of the stochastic dependent-chance programming model for the allocation of order quantity among vendors. All computational analysis was performed on an AMD Turion 1.7 GHz notepad and the algorithm code is implemented in C++.

Please note that due to space limitations of 6 pages for a regular paper, the results and discussion as well as additional references can be requested from the corresponding author, sohail.chaudhry@villanova.edu.

REFERENCES

Aissaoui, N., Haouari, M. and Hassini, E., 2007. Supplier selection and order lot sizing modeling: a review. *Computers and Operations Research*, 34(12), 3516-3540.

Amid A., Ghodsypour, S.H. and O'Brien C., 2006. Fuzzy multiobjective linear model for supplier selection in a supply chain. *International Journal of Production Economics*, 104(2), 394-407.

Amid A., Ghodsypour, S.H. and O'Brien C., 2007. A weighted additive fuzzy multiobjective model for the supplier selection problem under price breaks in a supply Chain. *International Journal of Production Economics*, In press, doi:10.1016/j.ijpe.2007.02.040.

Bender, P.S., Brown, R.W., Issac, M.H. and Shapiro, J.F., 1985. Improving purchasing productivity at IBM with a normative decision support system. *Interfaces*, 15(3), 106–115.

Charnes, A. and Cooper, W.W., 1959. Chance-constrained programming. *Management Science*, 6(1), 73-79.

Chan, F.T.S., Kumar, N., Tiwari, M.K., Lau, H.C.W. and Choy, K.L., 2008. Global supplier selection: a fuzzy-AHP approach. *International Journal of Production Research*, 46(14), 3825-2857.

Chaudhry, S.S., Forst, F.G. and Zydiak, J.L., 1993. Vendor selection with price breaks. *European Journal of Operational Research*, 70(1), 52-66.

Chen Y.M. and Huang, P.N., 2007. Bi-negotiation integrated AHP in supplier selection. *International Journal of Operations & Production Management*, 27(11), 1254 - 1274.

Crama, Y., R. Pascual J. and Torres, A., 2004. Optimal procurement decisions in the presence of total quantity discounts and alternative product recipes. *European Journal of Operational Research*, 159(2), 364–378.

Cooper, S.D., 1977. A total system for measuring of performance. *Journal of Purchasing and Materials Management*, 13(3), 22–26.

De Boer, L. and van der Wegen, L.L.M., 2003. Practice and promise of formal supplier selection: a study of four empirical cases. *Journal of Purchasing & Supply Management*, 9(3), 109-118.

Degraeve, Z., Labro, E. and Roodhooft, F., 2000. An Evaluation of Vendor Selection Models from a Total Cost of Ownership Perspective. *European Journal of Operational Research*, 125(1), 34-58.

Dempsey, W.A., 1978. Vendor selection and the buying process. *Industrial Marketing Management*, 7(4), 257–267.

Dickson, G.W., 1966. An analysis of vendor selection systems and decisions. *Journal of Purchasing*, 2(1), 5-17.

Gainen, L., 1955. Linear programming in bid evaluations. In: Proceedings of the 2nd Symposium of Linear Programming, Washington DC, 29-38.

Gregory, R.E., 1986, Source selection: a matrix approach. *Journal of Purchasing and Materials Management*, 22(2), 24–29.

Sevkli, M., Koh, S.C.L., Zaim, S., Demirbag, M. and Tatoglu, E., 2008. Hybrid analytical hierarchy process model for supplier selection. *Industrial Management & Data Systems*, 108(1), 122-142.

He, S., Chaudhry, S.S., Zhonglin, L. and Wang, B., 2008. Stochastic vendor selection problem: chance-constrained model and genetic algorithms. *Annals of Operation Research*, In press, doi:10.1007/s10479-008-0367-5

Kasilingam, R.G. and Lee, C.P., 1996. Selection of vendors - a mixed-integer programming approach. *Computers & Industrial Engineering*, 31(1-2), 347-350.

Kumar, M., Vrat, P. and Shankar, R., 2006. A fuzzy programming approach for vendor selection problem in a supply chain. *International Journal of Production Economics*, 101(2), 273-285.

Liao, Z. and Rittscher, J., 2007. A multi-objective supplier selection model under stochastic demand conditions. *International Journal of Production Economics*, 105(1), 150-159.

Kirytopoulos, K., Leopoulos, V. and Voulgaridou, D., 2008. Supplier selection in pharmaceutical industry: an analytic network process approach. *Benchmarking: An International Journal*, 15(4), 494-516.

Liu, B., 1997. Dependent-chance programming: a class of stochastic optimization. *Computers & Mathematics with Applications*, 34(12), 89-104.

A DECISION SUPPORT TOOL BASED ON ANP AND FMEA TO DETERMINE CAUSE FAILURES

Fabio De Felice[(a)], Antonella Petrillo[(b)]

[(a)]D.i.M.S.A.T - Department of Mechanism, Structures and Environment - University of Cassino, Italy
[(b)] D.i.M.S.A.T -Department of Mechanism, Structures and Environment - University of Cassino, Italy

[(a)]defelice@unicas.it, [(b)]a.petrillo@unicas.it

ABSTRACT
Failure mode and effect analysis (FMEA) is one of the formal techniques for evaluation of failure mode in mechanical and electronic equipments. The FMEA may be a very helpful tool for identifying weak points in the design stage of a product/process. The aim of this paper is to develop a *new maintenance decision strategy* by integrating the criticality of various factors related to failure and repair of a component/subsystem as an alternative to traditional approach of Failure Mode and Effect Analysis (FMEA). The methodology presented is based on *Analytical Network Process* (ANP), a multi-criteria decision making technique.
We propose a *decision support tool analysis* for determining maintenance priority action in which the typical FMEA parameters are modeled. The approach has been validated in a real case study concerning the European Train Control System - E.T.C.S.

Keywords: FMEA, Maintenance, ANP, Risk Analysis

1. INTRODUCTION
The identification and choice of a suitable risk assessment model has been considered as a crucial issue for decades. So far, models used in the practice were developed for different applications and adapted for health and safety at work (Hazards and Operability Study – HAZOP, Failure Methods and Critical Analysis- FMECA, Fault tree analysis, Events tree, etc.).

In our work we focused attention on the FMEA technique. FMEA has been widely standardized, as MIL-STD-1629A, MIL-HDBK-217 in the USA and as BS 5760 in the UK. Industrial users have reported significant benefits from these design tools. Successful users have achieved a 15–45% improvement in quality, and reduction in cost and time to market (Huang *et al.* 2000).

This technique is a well known assessment tool used to identify the components of an equipment most likely to cause failures, and to enhance the reliability of a system through the development of the appropriate corrective actions (Hung *et al.*, 1999).

FMEA is important for directing maintenance tasks and identifying more efficient operational methods and for allocating the recommended actions at those points with higher damage potentials.

The main problem faced in the utilization of this technique is the necessity to help management to consider different parameters simultaneously. Thus, it is useful to adopt multicriteria decision making techniques.

From this point of view in a recent article, Kjellen et al. (2009) pointed out the importance of risk of accidents as a criterion in decision making. Amongst many factors, maintenance practice will also affect the occurrence of accidents.

Multi criteria decision making approach gained momentum in the field of maintenance strategy selection (de Almedia and Bohoris, 1995; Triantaphyllou et al., 1997; Labib et al., 1998) suggested the use of AHP/ANP for maintenance strategy selection considering cost, reparability, reliability, and also used AHP/ANP for selecting the maintenance strategy for an Italian oil refinery based on four important criteria, namely cost, damages, applicability and added value (Pillay *et al*., 2003; Sachdeva, 2008).

Definitely, the aim of our work is to propose an *ANP decision support tool to evaluate systems reliability performance and to select the best maintenance strategy.*

2. ANP APPROACH
ANP (Saaty, 2001) is a comprehensive decision-making technique that captures the outcome of the dependence and feedback within and between the clusters of elements.

The main reason for choosing the ANP as our methodology is due to its suitability in offering solutions in a complex multicriteria decision making process (De Felice *et al*., 2009).

The ANP model consists of the control hierarchies, clusters, elements, interrelationship between elements and interrelationship between clusters. The modeling process can be divided into different phases for the ease of understanding which are described as follows:

PHASE 1: *Pairwise comparison and relative weight estimation.* Pairwise comparisons of the elements in each level are conducted with respect to their relative importance towards their control criterion based on the principle of AHP. Saaty (1980) suggested a scale of 1-9 when comparing two components (see Table 1).

Table 1: 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
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 result of the comparison is the so-called dominance coefficient aij that represents the relative importance of the component on row (i) over the component on column (j), i.e., aij=wi / wj. The pairwise comparisons can be represented in the form of a matrix (Saaty, 2007). The score of 1 represents equal importance of two components and 9 represents extreme importance of the component i over the component j.

$$A=\begin{vmatrix} w_1/w_1 & w_1/w_2 & \dots\dots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \dots\dots & w_2/w_n \\ \dots\dots & & & \dots\dots \\ w_n/w_1 & w_n/w_2 & \dots\dots & w_n/w_n \end{vmatrix} = \begin{vmatrix} a_{11}\dots\dots a_{1n} \\ a_{21}\dots\dots a_{2n} \\ \dots\dots\dots\dots \\ a_{n1}\dots\dots a_{nn} \end{vmatrix}$$

PHASE 2: *Priority vector.* After all pairwise comparison is completed the priority weight vector (w) is computed as the unique solution of $Aw = \lambda_{max}w$, where λ_{max} is the largest eigenvalue of matrix A.

PHASE 3: *Consistency index estimation.* The consistency index (CI) of the derived weights could then be calculated by: CI = (λmax−n) n−1. In general, if CI is less than 0.10, satisfaction of judgments may be derived (Saaty, 2005).

3. THE RESEARCH METHODOLOGY

The FMEA design and implementation requires a careful knowledge of the system (Anthony *et al.*,1998). A combination of techniques is therefore needed to perform system level availability modeling of complex heterogeneous control systems, considering both structural and behavioral studies (Puente et al., 2002; Bowles, 2003). For this reason we proposed a new decision multicriteria methodology. Below methodological steps are illustrated (Figure 1).

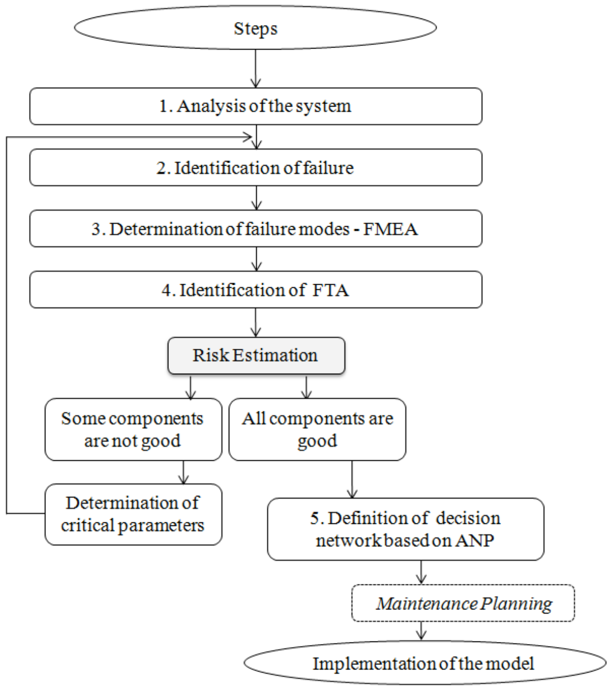


Figure 1: Methodological steps

The aim of our formal model is:

- to check the completeness of the specification of the ETCS;
- to use it for a systematic derivation of test cases;
- to evaluate at an early stage the specification of the European stardized interfaces of ETCS according to the national railway environment.

Here below are the methodological steps proposed:

STEP 1 - Analysis of the system. This preliminary activity represents a focal analysis as it supplies information about the organization of processes and procedures of each level.

STEP 2 - Identification of failure. This activity mainly consists of the identify failure that could characterize the system.

STEP 3 - Identification of failure modes – FMEA. This activity mainly consists of the identify failure modes: the aim is to integrate information coming from operational fields with global level effects.

STEP 4 - Identification of FTA. A fault tree structure is proposed to analyze undesired events with different levels of operation quality.

STEP 5 - Definition of decision network based on ANP. In this phase the main objective was to assess design and/or operational procedure alternatives which could improve reliability and maintainability of the whole system. The multi-criteria analysis allowed to obtain a numerical assessment on the representative of the various components, usually characterized by a qualitative measurement.

4. CASE STUDY: THE E.T.C.S.

As case study we analyzed the European Train Control System. The ETCS is a signalling, control and train protection system designed to replace the many incompatible safety systems currently used by European railways, especially on high-speed lines (see Figure 2).

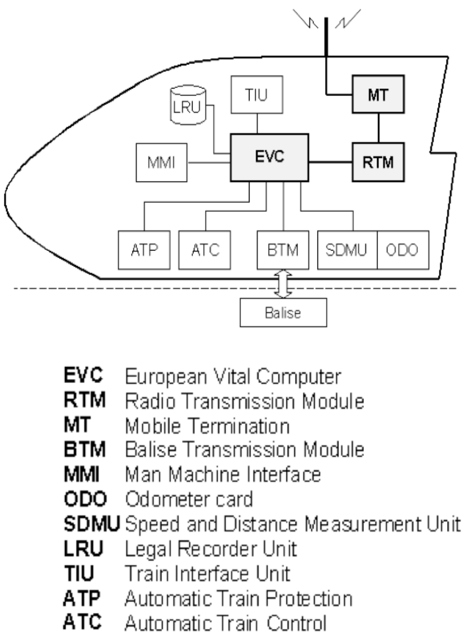


Figure 2: European Train Control System - E.T.C.S. framework

E.C.T.S. requests the observance of high safety and reliability standards (BS 5760, MIL-HDBK-217, MIL-STD1629A) thus is a proper case study for our work. Here below we analyzed methodological steps.

STEP 1 - Analysis of the system. The European Train Control System is a European project aiming at the cross-border operation of trains without obstacles, i.e. the free movement of train operators along various infrastructures while maintaining the necessary level of safety, thus creating real interoperability in the area of control command and signalling. In this context we note that computer systems used in critical control applications are rapidly growing in complexity, featuring a very high number of requirements together with large, distributed and heterogeneous architectures, both at the hardware and software levels. Traditional functional testing techniques based, for example, only on Fault Tree Analysis reveal inadequate for the verification of modern control systems, for their increased complexity and criticality properties (Frosig, P., 1995).
Here below (Table 2) we show a description of major components of the system.

Table 2: Components of system and Failure Rate

DESCRIPTION	FAILURE RATE
RADIO BLOCK CENTRE	
Safety Centre	*3,07E-06*
Functional Keyboard	*1,23E-05*
Interface Operator – Alarm – Remote control	*4,25E-06*
Interface TLC-LD & GSM-R	*4,00E-06*
Power RBC	*5,22E-06*
Encoder LEU	
Interconnection with relay electric system	*1,43E-07*
Distribution-Power	*3,99E-07*
LEU (Encoder)	*1,36E-06*
Front End Diagnostic	*2,37E-06*
Interconnections Balise & LEU	*2,07E-07*
Filter Module	*1,27E-07*
Splitter Module - FED	*9,20E-09*
Splitter Module - BUS I/O	*9,20E-09*
GPS Module	*3,63E-06*
Boa Balise	*4,50E-07*

STEP 2 - Identification of failure. In modelling the European Train Control System, different modeling aspects have been integrated:

- Components;
- Scenarios;
- Functions

are shown on different model levels.
When modelling the component view, the focus is on communication and interaction of different subsystems. A general representation of these nets shows the subsystems and their interfaces, every subsystem being detailed on additional levels. The scenario-based view is the modelling of operational procedures. Its main

elements are the interaction between on-board and trackside equipment and the sequence of events required to maintain operation. The functions are represented at lower model levels. The functions are specifically associated with the objects of the process aspect and represent the activities or the response to interaction requirements following from the scenarios.
In Table 2 is a definition of failure rate for major components.

STEP 3 - Identification of failure modes- FMEA. For FMEA analysis we adopted MIL-STD-1629 standard. We evaluated failures considering:

- The consequences of failure are the worst conceivable (conservative assumption);
- Failure is never contemporary to another (analyzing a fault at a time);
- The devices are in normal operation.

In appendix (Table 3) we show an example of "*Identification of failure modes*".

STEP 4 - Identification of FTA. For modeling fault we used the technique FTA (Fault Tree Analysis), deemed appropriate to highlight the dependencies between logical and functional components of the subsystem that can lead to abnormality determination of exercise (Top Event) and to quantify the probability of occurrence. FTA analysis assumes that the subsystem at the beginning of the mission is fully efficient, that every component is in good working condition and that all redundancies planned are active. The mission time was assumed to be 24 hours, or equal to the time of daily use of the subsystem. Here below (Figure 3) is an example of FTA.

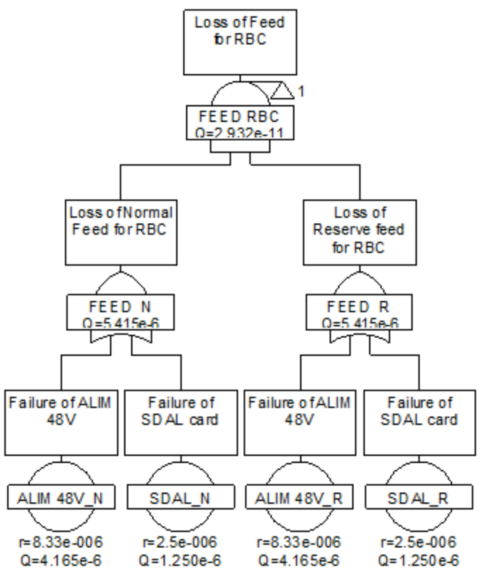


Figure 3: FTA example for "loss of feed for RBC"

STEP 5 - Definition of decision network based on ANP. In this phase we proposed an ANP model to develop a maintenance plan to reduce the unacceptable risk to an acceptable level. We distinguished between "*preventive maintainability*" and "*corrective maintainability*".

The definition of the network scheme proposed involved expert people in maintenance problems. We established a "FMEA team" including Mechanical Engineerings, Quality Experts, the Maintenance Experts Once the network structure of the maintenance decision making problem is defined (see Figure 4 in Appendix), the priorities of the scheme are calculated using pairwise comparisons.

Definitely, the network was characterized by the following clusters: Radio Block Center, Encoder, Boa Balise, Failure Causes, Failure Effects, Failure Modes, Alternatives that include Preventive Maintainability and Corrective Maintainability.

Here below (Figure 5 and Figure 6) we illustrate some results:

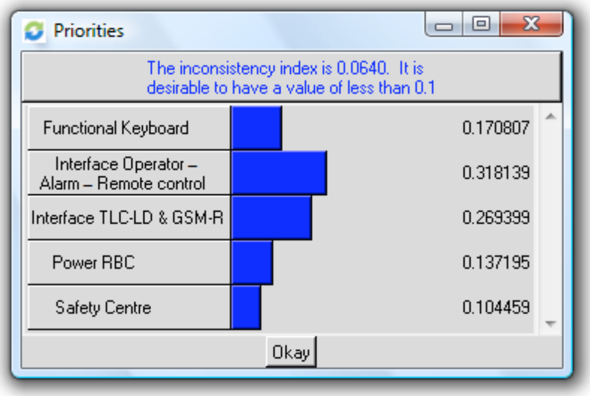


Figure 5: Priority vector for "Corrective Maintainability" node in "RADIO BLOCK CENTER" cluster

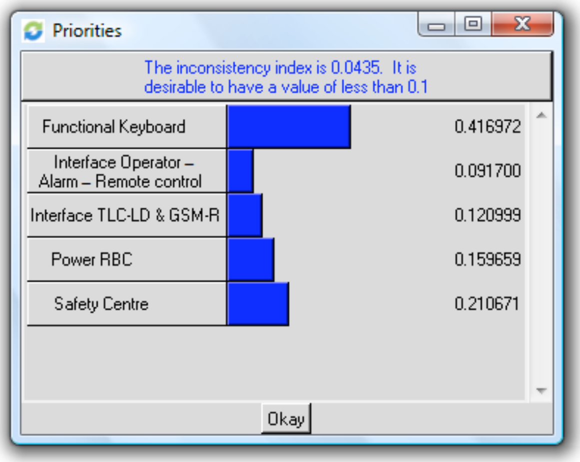


Figure 6: Priority vector for "Preventive Maintainability" node in "RADIO BLOCK CENTER" cluster

Once defined all comparison judgments we obtain a ranking of global priority for "Preventive Maintainability" and for "Corrective Maintainability" (Table 4).

Table 4: Global Priority

Component	Global Priority for Preventive Maintainability	Global Priority for Corrective Maintainability
Safety Centre	0,23	0.15
Functional Keyboard	0,25	0.23
Interface Operator – Alarm – Remote control	0.30	0.27
Interface TLC-LD & GSM-R	0.35	0.29
Power RBC	0.15	0.24
Interconnection with relay electric system	0.16	0.19
Distribution-Power	0.18	0.28
LEU (Encoder)	0.21	0.13
Front End Diagnostic	0.19	0.14
Interconnections Balise & LEU	0.27	0.22
Filter Module	0.24	0.21
Splitter Module - FED	0.13	0.17
Splitter Module - BUS I/O	0.26	0.24
GPS Module	0.14	0.20
Boa	0.38	0,29

Established priorities of actions is necessary to implement the maintenance programs. In appendix (Table 5 and Table 6) we propose as an example a preventive and corrective program (we show only a partial implementation).

5. CONCLUSION

The unexpected failures, the down time associated with such failures, the loss of production and, the higher maintenance costs are major problems in any process.

It is necessary to develop a specific methodology to assess the reliability of systems/components. In fact, the failure of a system is rarely the result of a single cause, but rather the result of a combination of a series of interacting events.

As a result, risk-based maintenance must not be perceived as a static exercise to be performed only once. It is a dynamic process, which must be continuously updated as additional information becomes available.

For this reason in our work, we have proposed a methodology to develop an optimum risk-based maintenance strategy. The paper presents a new methodology for designing maintenance programs based on Analytic Network Process.

This approach:

1. allows to integrate classical methodology based on the FMEA technique with a multicriteria decision making approach;
2. ensures to take in consideration different parameters simultaneous. This will contribute to guarantee the reliability and availability of the whole system.

Further developments could be oriented in applying the approach in several fields (automotive, aeronautic, etc.).

REFERENCES

Anthony, N.R., Govindarajan, G., 1998. *Management Control. Systems*. Irwin McGraw-Hill, New York, NY.

Bowles, J. B., 2003). An assessment of RPN prioritization in a failure modes effects and criticality analysis. *Prec. Ann. Rel. Maint. Symp., IEEE*, New York, p. 380.

BS 5760, Part 5, *Guide to failure modes, effects and criticality analysis* (FMEA and FMECA).

de Almedia, A.T., Bohoris, G.A., 1995. Decision theory in maintenance decision making. *Journal of Quality in Maintenance Engineering* 1, 39–45.

De Felice, F., Petrillo, A., 2009. Proposal of a new model for the optimization of the organizational process in industrial company through the application of the Analytic Network Process. *Proceedings of the International Conference on Modelling & Applied Simulation*, Porto de la Cruz (Tenerife), 23-25 September.

Huang, G.Q., Shi, J., Mak, K.L., 2000. Failure Mode and Effect Analysis (FMEA) Over the WWW. *International Journal Adavanced Manufacturing Technoly*, 16 603–608.

Hung, G.Q., Nie, M. and Mark, K.L.,1999. Web-based failure mode and effect analysis, *Computers and Industrial Engineering*, Vol. 37, pp. 177–180.

Frosig, P., 1995. ETCS: Requirements to Seamless Cross-Border Operation. *International Railway Journal and Rapid Transit* Review, 9 September, New York.

Kjellen, U., Motet, G., Hale, A., 2009. Resolving multiple criteria decision making involving risk of accident loss. *Safety Science* 47, 795–797.

MIL-HDBK-217, *Reliability Prediction of Electronic Equipment*, U.S. Department of Defense.

MIL-STD1629A, *Procedures for performing a failure mode, effects, and criticality analysis,* US Department of Defense.

Labib, A.W., O'Connor, R.F., Williams, G.B., 1998. An effective maintenance system using the analytic hierarchy process. *Integrated Manufacturing Systems* 9, 87- 98.

Pillay, A., and Wang, J., 2003. Modified Failure Mode and Effects Analysis using Approximate Reasoning. *Rel. Eng. Sys. Saf*., 79, 69.

Puente et al., 2002. A decision support system for applying failure mode and effects analysis. *International Journal of Quality & Reliability Management*, 19 (2) 137-150.

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., 2005. *Theory and Applications of the Analytic network Process: decision making with benefits, opportunities, costs, and risks*. RWS Publications, (Pittsburg ,USA).

Saaty T. L., Peniwati K., 2007. *Group Decision Making: Drawing out and Reconciling Differences*. RWS Publications, (Pittsburg , USA).

Sachdeva, A., Kumar, D., Kumar, P., 2008. A methodology to determine maintenance criticality using AHP. *International Journal of Productivity and Quality Management*, Vol. 3, Number 4 / 2008, 396 – 412.

Triantaphyllou, E., Kovalerchuk, B., Mann Jr., L., Knapp, G.M., 1997. Determining most important criteria in maintenance. *Journal of Quality in Maintenance Engineering*, Vol. 3 Iss: 1, pp.16 - 28

AUTHORS BIOGRAPHY

Fabio De Felice, Professor 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. Holder of the course planning and management of industrial plants.
The scientific activity developed through studies and researches on problems concerning industrial plant engineering. Such activity ranges over all fields from improvement of quality in productive processes to the simulation of industrial plants, from support multicriteria techniques to decisions (Analytic Hierarchy Process, Analytic Network Process), to RAMS Analysis and Human Reliability Analysis.
The main courses, both curricular and specialistic for territorial businesses in which he is involved are: Safety of Industrial Plants, Industrial Production Management, Industrial Simulation, Human Reliability Analysis.
General Secretary of the Analytic Hierarchy Process – A.H.P. Academy - International Association for the promotion of multi-criteria decision making methods.
Antonella Petrillo, degree in Mechanical Engineering, now PhD at the Faculty of Engineering of University of Cassino where she conducts research activities on Multi-criteria decision analysis (MCDA) at the Department of Mechanism, Structures and Environment.

APPENDIX

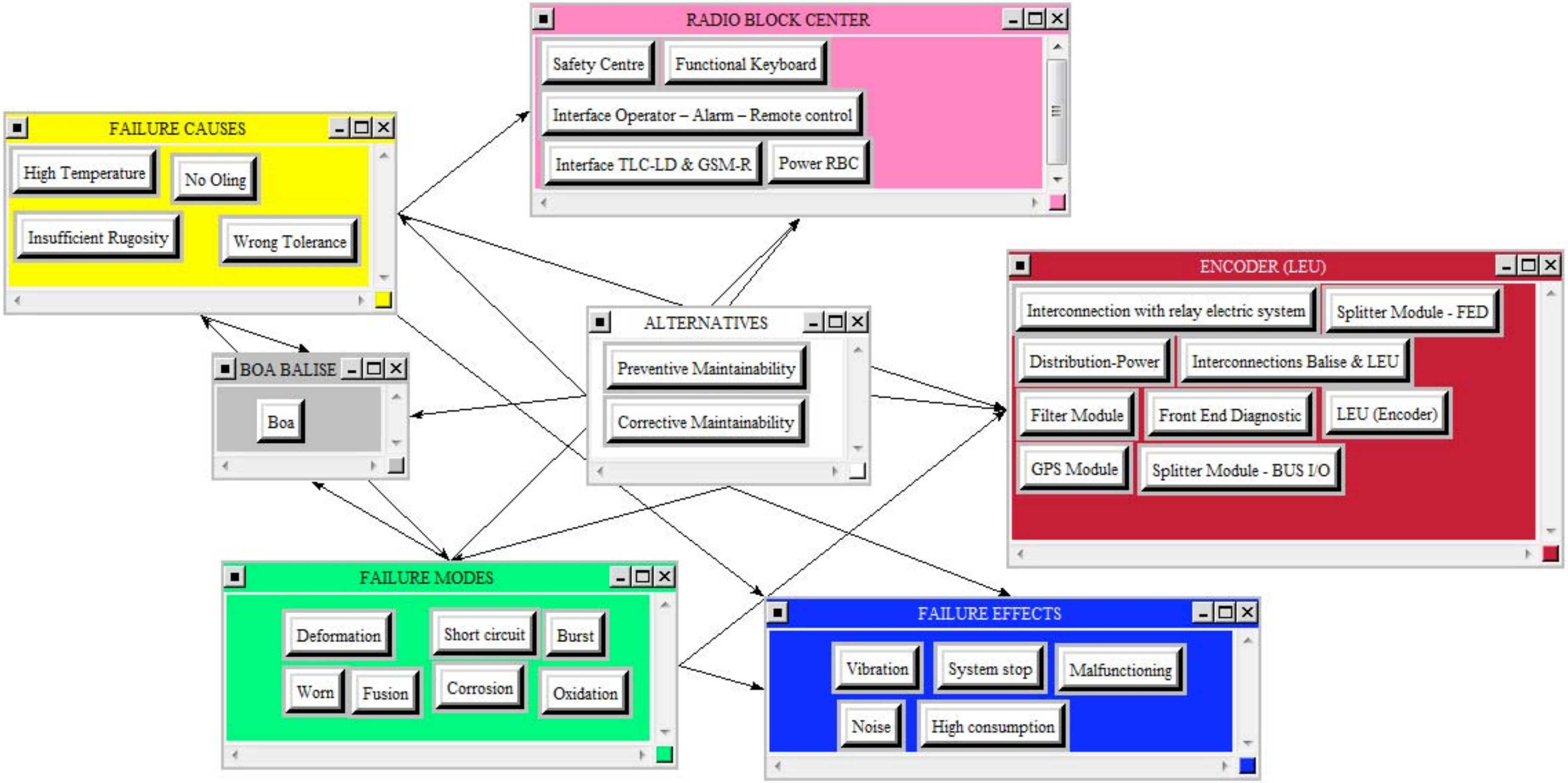


Figure 4: ANP Model

Component: *Safety Centre*				FMEA							
Subsystem: **RBC**											
N.	Part N.	Symbol	Function	Mode and cause of failure	Functioning	Failure Effects			Method of Detection	Compensation measures	Criticality Level.
						Local	Main Component	Final			
1	/	XyZ	Making the internal connections between sections of processing and interface guidelines and distribute the supply voltages.	Mechanical failure	Continuous	Lack connections and / or power between sections of processing and interface guidelines	ECTS Central Out of order	Prevent the train running under the jurisdiction of RBC	Diagnostic Test	Immediate replacement equipment of failure	A failure or disorder that prevents or influences multi-gear trains.

Table 3: FMEA Example

Preventive Maintainability

Component: **BOA Balise**
Subsystem: **BOA**

Step	Action	Frequency [months]	Tools	Materials		Number of operators		Duration of action [min]	Total hour/operator
				Identification	Quantity [No.]	No.	Skill		
1	Visual inspection of BOA Monitor the presence of debris within the limits allowed. Control correct positioning of Boa. Control of the efficiency of the electric connections	12	Toolbox			1	Necessary a specific knowledge of the system. Operators must be able to make medium repairs.	7	0,12
2	Proper tightening of control cable to LEU cabinet	12	Toolbox			1	Not necessary a specific knowledge of the system. Operators must be able to make easy repairs.	7	0,12
3	Verify the functionality is not interested in normal operation (eg switching to telegram by default).	12	Toolbox			1	Not necessary a specific knowledge of the system. Operators must be able to make easy repairs.	6	0,10

Table 5: Preventive Maintainability Example

Corrective maintainability

Component: **BOA**
Subsystem: **BOA**

Symbol	Failure Mode	Mapping of failure	Standard procedure	Tool	Number of operators		Duration of action	Observations
					Specialization	Quantity No.	[h]	
BOA	Out of order	Self-revealing	1) Location component failure 2) Isolation component 3) Removal 4) Replace failed component 5) Realignment 6) Functional Tests	Toolbox	Medium	1	0,333	

Table 6: Corrective Maintainability Example

HARMONIZED AND REVERSIBLE DEVELOPMENT FRAMEWORK FOR HLA BASED INTEROPERABLE APPLICATION

Z. Tu(a), G. Zacharewicz(b) D. Chen(c) and Z. Zhang(d)

(a)(b)(c)(d) IMS-LAPS / GRAI, Université de Bordeaux – CNRS, 351, Cours de la Libération, 33405, TALENCE cedex, France

(a)zhiying.tu@ims-bordeaux.fr, (b) gregory.zacharewicz@ims-bordeaux.fr, (c) david.chen@ims-bordeaux.fr, (d) Zhenchuan.Zhang.Simon@gmail.com

ABSTRACT
This paper aims at improving the re-implementation of existing information systems when they are called to be involved in a system of systems, i.e. a federation of enterprise information systems that interoperate. The idea is reusing the local experiences coming from the development of the original information system with the process of Model Discovery and Ontological approach. We draw the strong points of MDA and HLA, former in transforming concepts and models from the conceptual level to the implementation and latter in implementation of distributed systems, then we propose a MDA/HLA framework to implement distributed enterprise models from the conceptual level of federated enterprise interoperability approach. In addition, we propose a model reversal methodology to help re-implement the legacy information system, in order to achieve the interoperability with other systems. We also draw the strong points from web services in order to improve the performance of HLA in distribute simulation.

Keywords: Interoperability, HLA, Model reversal, HLA evolved

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 ability 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 the semantics necessary to assure interoperability.

This paper will contribute on rapid and intelligent development of Distributed Enterprise Information Systems by proposing a harmonized and reversible development framework for HLA based interoperable application. In fact, A multiple agent/HLA enterprise interoperability methodology and an unified reversible life cycle for interoperable enterprise distributed information systems have been proposed in (Zacharewicz 2008) & (Tu 2010). They mention that several enterprises, who participate in a cooperative project, need to exchange various information any time through a platform based on HLA. However, most of the enterprises have their own legacy information systems and the question is how to use HLA to connect them conveniently, rapidly, flexibly and perfectly solve the interoperability issues. This paper will use the harmonization of MDA (Model Driven Architecture) and HLA to form a rapid and flexible development life cycle. In order to discover the enterprises' knowledge quickly, Model Reverse Engineering will be used to reverse the legacy information system of enterprises backwards to models, and concurrently to solve the interoperability problems without rebuilding the information systems or integrate them (Jouault 2009). HLA and Web Services will be mixed in order to improve HLA's performance, which enables a closer integration of simulation based systems with traditional operational systems in a systems-of-systems approach. Web Services will help HLA to achieve Distributed Simulation Interoperability with external Systems in world wide web (Möller 2005).

The paper is structured as follows. We start out with a description of background, explain the problems we want to solve (section 2). Then we give a survey of related work in the area of HLA, MDI, Web services and Model Reverse Engineering (section 3). Then, we describe our framework in general and related scenario (section 4). After that, we explain our methodology separately in three parts, harmonization of MDA and HLA FEDEP, model reversal and HLA Evolved Web services (section 5).

2. PROBLEM STATEMENT

The enterprise interoperability framework mentions that there are three barriers for enterprise interoperability, conceptual barrier, technological barrier, organizational barrier (Ullberg 2007). And those barriers are the problems we want to solve. We propose a synthetic development life cycle to facilitate the development of the dynamic, secure and synchronized interoperable HLA based platform for enterprises with heterogeneous legacy IT system.

The three crosses in figure 1 represent three barriers in enterprise interoperability framework. For example, *conceptual barrier*, different IT systems use different terms to describe “car”; *technological barrier*, different IT systems use different data format for message transmission; *organizational barrier*, diverse organization structures exist in different enterprises. In addition, there is another problem. When an enterprise intents to participate in a existing cooperative project and connect to other heterogeneous IS, how can we establish this connection efficiently, correctly, and with low cost? In this paper, we are going to answer this question.

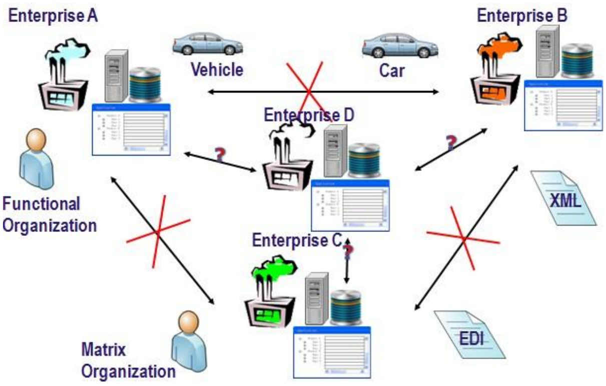


Figure 1 Problems to solve

3. TECHNICAL RECALL

3.1. Model-Driven Interoperability framework

The approach "Model Driven Interoperability" (MDI) consists in considering interoperability problems from enterprise models level instead of only at the coding step.

These works were realized in the Task Group 2 (TG2) of INTEROP-NoE state at defining an approach inspired from OMG MDA (Bourey 2007). The goal is to tackle the interoperability problem at each abstraction level defined in MDA and to use models transformations techniques to link vertically the different levels of abstraction or horizontally to ensure each models of the level interoperability. The main goal of MDI, based on model transformation, is to allow a complete follow from expressing requirements to coding of a solution and also a greater flexibility thanks to the automation of these transformations.

In the context of TG2, experimentations have been realized and in particular the feasibility study to transform GRAI Methodology (Chen 1997) (Doumeingts 2001) Models to UML between CIM and PIM levels (Bourey 2007). These works are additional works realized in the context of ATHENA to define UML profiles to take into account Service Oriented Architectures at PIM level (Gorka 2007). These results have consolidated by results presented by (TOUZI 2007) that have proposed an interoperability transformations method from BPMN to UML in the context of services oriented architecture.

Nevertheless, the soundness of the methodology has been demonstrated, but no full industrial scale validation have been yet realized. A project ISTA3 has especially for goal to demonstrate these concepts in an industrial real world significant application. The different methodological propositions will be tested and refined by focusing on model interoperability and their interoperability. It will consist in particular to improve the flexibility of the MDI transformation process in the way of obtaining dynamic interoperability in the context of federated approach..

3.2. High Level Architecture

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.

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 2000).

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.

The FEDEP (Federation Development and Execution Process) describes a high-level framework for the development and execution of HLA federation. FEDEP uses the seven-step process to guide the spiral development of the simulation system through phases of requirements, conceptual modeling, design, software development, integration, and execution (IEEE 2003).

3.3. MDA & Model Reverse Engineering

MDA is well-known for promoting the use of models and their transformations to consider and implement different systems. After MDA became an important change in software development practice, there is another research soon launched in OMG, later called Architecture Driven Modernization (ADM) (OMG 2008).

The basic idea proposed in the MDA approach is to translate from an abstract platform-independent model (PIM) expressed in UML into a more concrete platform-specific model (PSM). When the PSM is obtained, then we still need to generate code (OMG 2003). In the opposite aspect of view, ADM is trying to track the models by backwards to discover models from legacy systems. However, sometimes one would like to discover more specific models from a legacy system. This is why the ADM group has defined several meta-models to this purpose, the best known being the knowledge discovery meta-model (KDM) and abstract syntax tree meta-model (ASTM) (OMG 2008).

KDM is a meta-model for representing existing software system and the associations, relationships among the function models in the system, and also describe the operation environments. It can insure the interoperability among the existing systems, make the data exchange among different vendors' tools easier.

3.4. Web Services

Web Services is a well known technique in distribute interoperability simulation domain. It has achieve a great success in business domain, which stems from the good characteristics of the technology itself, and widely recognized by enterprises and business organizations and effective support for the open source community (Richardson 2007).

The data exchange of Web services is based on open standards such as HTTP and XML, is not associated with any particular vendor, operating system and programming language, which makes Web services platform with a good vendor neutrality, coarse-grained business functions can also be packaged for the Web service, and it can be discovered by potential consumers.

At the same time, business-to-Web service is widely recognized. Microsoft, IBM and Sun and other leading manufacturers as well as Apache and other open source organizations support it, and World Wide Web and other standardization organizations active participation in the Web services technology provides great maturity and popularity of the organization advantage.

4. OVERVIEW OF FRAMEWORK

4.1. Framework Description

Harmonized and Reversible development framework for HLA based interoperable application, is the general description of this framework. In this title, there are four keywords, harmonized, reversible, HLA, and interoperable.

Harmonized means this framework is a synthetic framework, which consists several techniques. As figure 1 shows, in this framework, we create a new five steps development life cycle which aligns MDA and HLA FEDEP. MDA is easy to use and easy to understand, and tightly bounding with Unified Modeling Language , Meta-Object Facility (MOF), and Common Warehouse Meta-model (CWM). It is the best choice to overcome the interoperability barriers, which mentioned in MDI framework (Elvesæter 2007). HLA FEDEP is the standard of development and execution of HLA federation, and it is similar with the waterfall development but with look-back test phrase. The motivation of alignment of MDA & FEDEP

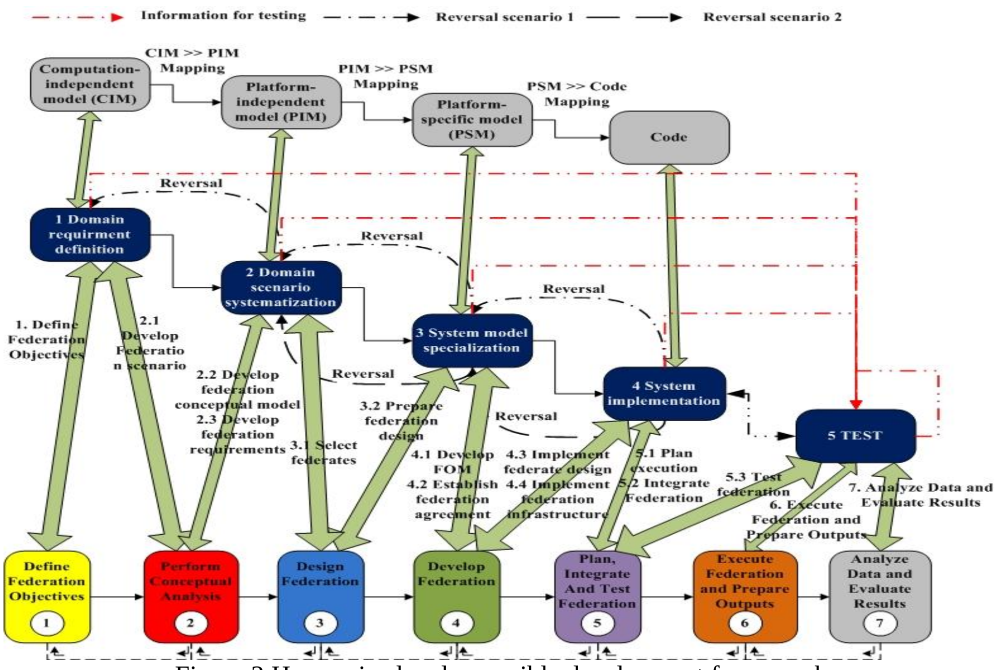


Figure 2 Harmonized and reversible development framework

is they have lots of similar steps and tasks and easy to reconstruct and align. In addition, this framework also use web services to improve the performance of HLA, which will explain in section 5.3.

Reversible means that this framework uses model reverse engineering technique to discover the model from the legacy system. To involve model reverse engineering technique aims to avoid to rebuild the legacy system for different cooperation. Our objective is to accelerate the development and reduce the cost. As figure 2 illustrates, there are two arrows, which have opposite direction to five step development life cycle. These two arrows represent two different scenarios of model reversal in this framework. We will explain this reversal in 5.2.

HLA means that this framework dedicates to the development of HLA based application. In our project, we choose an open source RTI, PORTICO. The reason of choosing that, is not only because it is free, but also we want to initiate an open framework to receive more comments and contributions from people who are interested in this idea.

Interoperable means that this framework provides a solution for achieving enterprise interoperability. As mentioned in section 2, we are going to overcome the enterprise interoperability barriers and help to realize the short-lived ontology approach, which mentioned by (Zacharewicz 2009).

4.2. Scenario

A schema of the related scenario is shown in figure 3. We assume that before enterprises start to launch a cooperative project, they all have their own system. Thus, our goal is to achieve the interoperability among those legacy systems. The steps of our approach as follows.

Step 1: model reverse engineering will be used to discover the models from the legacy system based on enterprises' requirement and our interest. And then, rewind these models into MDA models based on the alignment of MDA and HLA FEDEP, and solve interoperability problem of each level of MDA models according to the principle of MDI framework.

Step 2: a test of the final models we got by model reverse engineering is carried out. After that, we are going to transform correct models from CIM to code, and generate a Federate Interface, which can plug into the HLA platform and transmit the information with other companies' information systems via RTI.

Step 3: if there are other enterprises want to join this cooperative project, they also need to follow the step 1 and step 2, to rewind their legacy system into MDA models, and remove the incompatible parts, then generate the Federate Interface, finally, synchronize with other systems.

5. SPECIFICATION OF FRAMEWORK

5.1. Harmonization of MDA and HLA FEDEP

As known, HLA has lot of advantages, such as generalized development process: Federation Development and Execution Process (FEDEP), synchronization standard: Runtime Infrastructure Specification, Data Standards and etc, however, in order to keep these advantages, HLA needs to benefit from the commercial developments. As MDA is popular, widely recognized and it is an align able development life cycle with HLA FEDEP, and MDA can facilitate the construction of simulators and provide the standardized meta models to this integration (Andreas 2002) (Shawn 2003) (Trbovich 2005), MDA will be the best choice for HLA to benefit from. Besides that, from interoperability aspect of view, because most of the enterprises prefer to build their information system by using MDA, MDA would be the best choice for overcoming the interoperability barriers (Ullberg 2007). Model-Driven Interoperability (MDI) Framework provides a foundation, consisting of a set of reference models, for how to apply Model Driven Development (MDD) in software engineering disciplines in order to support the business interoperability needs of an enterprise (Elvesæter 2007).

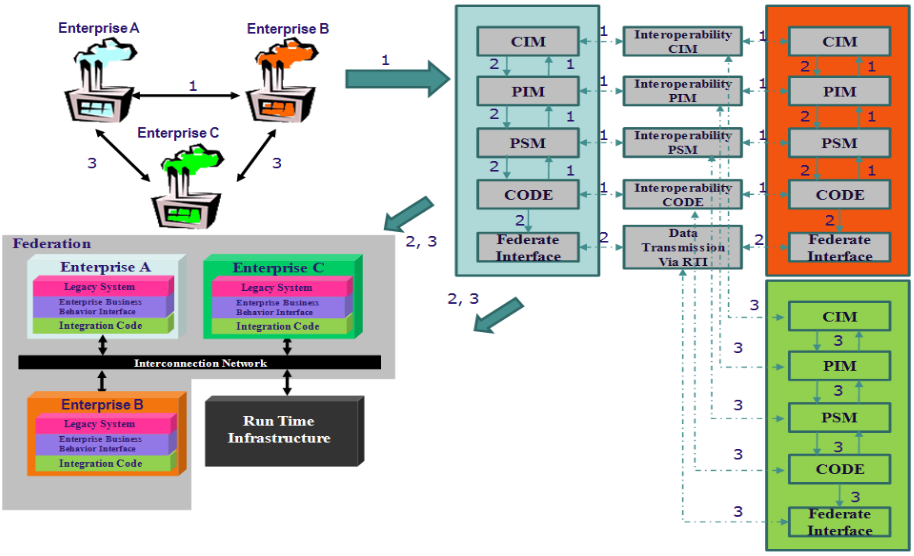


Figure 3 Scenario description

5.1.1. Specification of the harmonization

In this section, we propose a new development lifecycle to reconstruct HLA FEDEP and MDA, and generate a new five steps development framework (as shown in figure 4). This new methodology aims to adopt the strong points from both HLA FEDEP and MDA while overcoming their weak points, then, to achieve proper component reuse and rapid development.

Phase 1: *Domain requirement definition*, whose main task is to collect clear and enough requirements from customer in order to define the objective of the system, to describe the environment of the system, the scenario of the system. At the same time, all these definition and description need to be reasonable, understandable for all the stakeholder. CIM of MDA has more similar task with Define Federation Objectives, Develop Federation scenario together in HLA FEDEP. As the result, we align them in this phase, to convert the user requirement, which is more textual based, into more visual model, such as UML use case to derive the federation requirement, etc.

Phase 2: *Domain scenario systematization*, whose main task is to refine the domain scenario and business process defined in the first phase, to identify and describe the entities involved in the scenario and business process. And then, to define the relationships among entities and behaviors, events for each entity, etc. This phase integrates PIM in MDA, which describes the operation of system but doesn't address the detail platform information yet, as well as steps of Perform Conceptual Analysis, Develop Federation Requirements and Select Federates in HLA FEDEP, which also define and select general participators of the federation, then describe their relationship, behaviors and event in general.

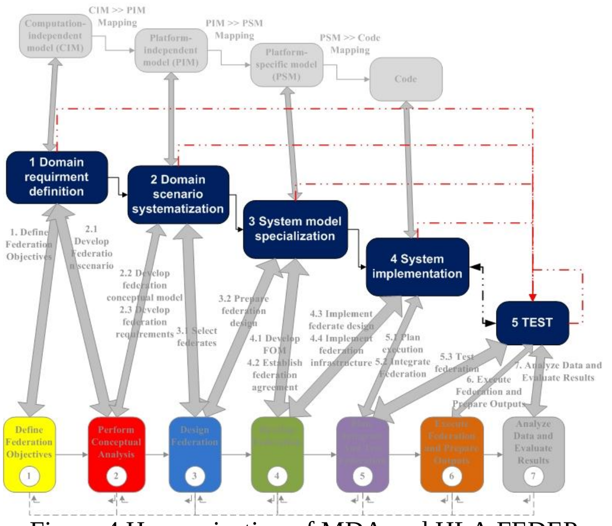


Figure 4 Harmonization of MDA and HLA FEDEP

Phase 3: *System model specialization*. In this phase, according to the technique chosen and platform selected, the system needs to be refined, for instance, to refine federation and federate structure, to allocate functions and attributes, etc. Detailed design will carry out at this time. This phase integrates the following parts in MDA and FEDEP. PSM in MDA, which is in the form of software and hardware manuals or even in an architect's head, will be based on detailed platform models, for example, models expressed in UML and OCL, or UML, and stored in a MOF compliant repository. The Prepare federation design, Prepare plan, Develop FOM, and Establish federation agreement in FEDEP will produce federate responsibilities, federation architecture, supporting tools, integration plan, VV&A plan, FOM, FED/FDD and time management, date management, distribution agreements, etc.

Phase 4: *System Implementation*, whose task is to transfer the specific system model into code, to create the executable federation and runable federate. At this level, MDA has various transformation techniques from model to code. In the FEDEP, Implement Federate designs will provide modified and/or new federates and supporting database. Implement Federation Infrastructure will provide implemented federation infrastructure and modified RTI initialization data. Plan Execution and Integrate Federation will provide execution environment description and integrated federation.

Phase 5: *Test*. Throughout the previous steps of the MDA and HLA FEDEP alignment process, testing is essential to ensure fidelity of the models. Testing phase includes the Test Federation, Execute Federation and Prepare Outputs, and Analyze Data and Evaluate Results in HLA FEDEP. Meanwhile, it will also refer to the outputs from the previous steps, such as the original user requirement in the first step, and federation test criteria from second phase.

5.1.2. Harmonized federate structure

A schema of the expected result can be seen in figure 5. After this harmonization, we can consider a federate as a *converter*. A federate has two parts, one is *Adapter* and another is *Plug-in*. *Adapter* includes *Enterprise Business Behavior Interface*, which connects to the enterprise business process related to specific strategies and algorithms of different enterprises. *Plug-in* includes *Integration code*, which manages the interactions between the enterprise business behavior interface and the RTI, providing an RTI independent API to the enterprise business behavior interface, and a simulation independent API to the RTI services.

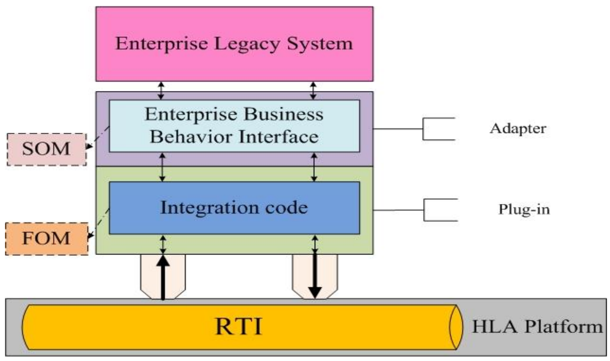


Figure 5 Harmonized federate structure

The objective of these abstractions is to ensure that the enterprise business behavior remains decoupled from RTI services. After the harmonization, all the federates will have the same Integration code but different Enterprise Business Behavior Interface, and any simulation related services required by the enterprise business behavior interface are accessed via the integration code, rather than through direct interaction with the RTI. As the result, integration code is common components for all the federate of the existing coordinators and also the reusable components for the future coordinators. And enterprise business behavior interface will adapt to different legacy systems of different enterprises, it will be implemented based on specific strategies and algorithms of different enterprises, it will accomplish the cipher mission.

5.2. Model Reversal

5.2.1. Specification of Model Reversal

This section describes a brand-new process of model reverse engineering with different scenarios constraints (see figure 6). The reverse process will re-characterize the legacy system in order to capitalize on the information and functions of the existing system, and make it easy for reusing in a new HLA compliant system. This methodology will assist to HLA FEDEP / MDA alignment mentioned in previous section, to fully achieve rapid development of federation and/or federate based on the legacy IT systems.

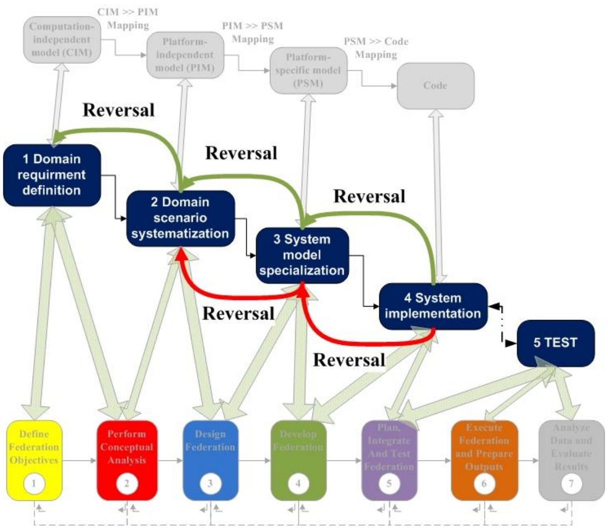


Figure 6 Model Reversal Framework

There are two different arrows, which represent two different scenarios of the model reverse process, illustrated in figure 6.

1. First, when an enterprise intents to start exchanging information in a new cooperative project with other enterprises. In that case, the HLA federation has not been created yet, so we propose to reverse the code of the legacy information systems to the first definition phase (domain requirement definition). Then from top to down, we generate the model for each phase, finally we produce a federation and federate rapid development template.
2. Second, if an enterprise intents to participate in a existing cooperative project and exchange data with other heterogeneous IS. Thus, we assume an HLA federation has already been created. Here, according to the HLA FEDEP, federate starts to be considered from the second step (Perform Conceptual Analysis) as the reversal scenario 2 shows in fig. 3.2. Therefore it is not necessary to reverse to the first phase, the reversal can stop at the second phase (Domain scenario systematization). One will only reuse the model of the existing federation to create the model for the federate related to the legacy system of the new participator. Finally, the model of the existing federation and the new federate model are used to generate the code template for the new federate for rapid development.

5.2.2. Expected result of Model Reversal

A schema of the expected result can be seen in figure 7. This illustration is based on the approach of MoDisco (for Model Discovery).

In MoDisco principle (Jouault 2009), A model (Mi) in the modeling world is a representation of a system in the real world and the nature of the model (Mi) is defined by its meta-model (MMi). it means that model Mi conforms to its meta-model MMi, and every step is guided by a meta-model. The very first step of a model discovery process is always to define the meta-model corresponding to the models you want to discover. Then, the second step is about creating one or many discoverers, which is illustrated in the middle of figure 7. These discoverers extract necessary information from the system in order to build a model conforming to the previously defined meta-model. The way to create these discoverers is often manual but can also be semi-automatic.

In addition, in order to adapt MoDisco principle to our approach, we add *constrains* before the *discoverers*. As we know, the legacy system consists lots of subsystems, which are always based on various kind of platforms and techniques, thus it is big and complex. If we are going to reverse the whole legacy system, the project will be extremely hug and complicated, and it departs from our motivation. As the result, constrains aims to specify the target source, which means that we must first define the bound of model reversal before start to reverse. And the boundary must be defined based on our interest and each enterprise's confidential information. This boundary specification will be recorded as a configuration file which can be read by *discoverers*.

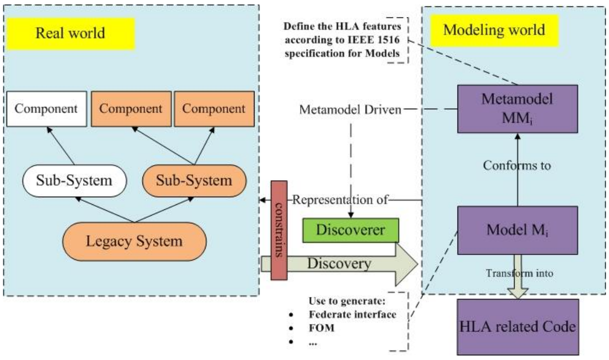


Figure 7 The expected result of model reversal

According to the ongoing research, none of the software tools can fully reverse the legacy system from code to model. Some of the tools can rewind the code to static model without the dynamic one, and some of them can only discover the data model from database. In our approach, we choose MoDisco, which is an Eclipse GMT component for model-driven reverse engineering. The common infrastructure of MoDisco is inspired by the KDM, which provides a comprehensive high-level view of application behavior, structure, and data and corresponds to a coarse-grained representation of a given legacy portfolio. MoDisco is a powerful tools even though it needs to be further improved. It provides different kinds of DSLs (associated to a meta-model) to describe the facet you want to discover from a given legacy. Furthermore MoDisco provides a toolbox that facilitates the building of model discoverers from the definition of the target meta-models. MoDisco can generate static model from the code of program, and can also generate dynamic model from the execution of program. However, the output of MoDisco is not adaptive to our project. Now, the output is a XML based file and with a tree structure illustration on eclipse, it is not easy for the people without strong IT background to understand it. So, our objective is based on MoDisco, visualize its output and add constrains in order to adapt to our project.

5.3. HLA Evolved Web Services

In the past few years, HLA has succeeded in many aspects, especially in the areas of interoperability and reuse, however, in pace with the rapid technical change and the future of the standard, HLA has to face to many new challenges.

5.3.1. Specification of HLA Evolved Web Services

As known, HLA has been developed within the defense simulation community, and more focus on the simulation in Local Area Network or in Virtual Private Network, thus disadvantages expose: lack of flexibility, lack of integration with business and etc. Meanwhile, Web Services has been developed within the commercial enterprise community, and has achieved a great success in business domain, which stems from the good characteristics of the technology itself, and widely recognized by enterprises and business organizations and effective support of the open source community. As the result, Web Services is the perfect option to help HLA overcome those disadvantages.

The general idea of HLA evolved Web services is illustrated in figure 8. We assume that a cooperative project has been launched. The information systems of the members run well within the federation. During this project, another enterprises want to join this project with different expectation, such as different cooperation time periods, different cooperation domains, different expected results from federation, etc. As the result, to rebuild the federation is impossible. Our solution is to add one more federate, *WebservicesFederate*, as shown in figure 8. This federate will provide various services, different access permissions, and the common API of existing federation. The candidates can used the common API and the service they prefer to general their own federate, and connect to the existing federation with different authorities via the wide area network.

For example, in figure 8, there are two enterprise X and Y want to participate in the existing project. Enterprise X is a supplier and enterprise Y is a client who is interested in the final product of this project. Thus, enterprise X has to know the workflow which related to his business, and synchronize its information with other stakeholders. While, enterprise Y only ask the information from the federation, so, it doesn't have to synchronize with other systems. As the result, enterprise will ask *WebservicesFederate* for the service with an authority of synchronization with other federates, but enterprise Y will only require the service with the lowest authority, which can manage to listen to the federation. Finally, no matter what kinds of services they got, they are connected to the existing federation via web services.

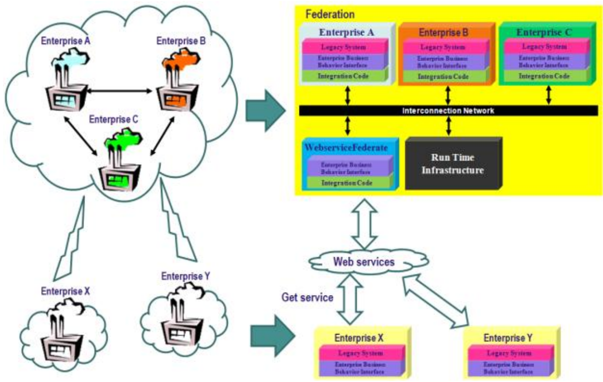


Figure 8 HLA Evolved Web Services

5.3.2. Architecture of HLA Evolved Web Services

The architecture of HLA Evolved Web Services is illustrated in figure 9. In our project, we choose an open source RTI, PORTICO (POTICO 2009), which doesn't provide Web-RTI functionality. Actually, PITCH RTI supports Web-RTI functionality (Möller 2007) , but it is not free, and as mentioned before, we want to initiate an open framework, so we choose PORTICO. So, we implement *WebservicesFederate* as a *bridge*, who takes charge of providing web services, connecting and synchronizing the federates outside the federation with the federates inside the federation.

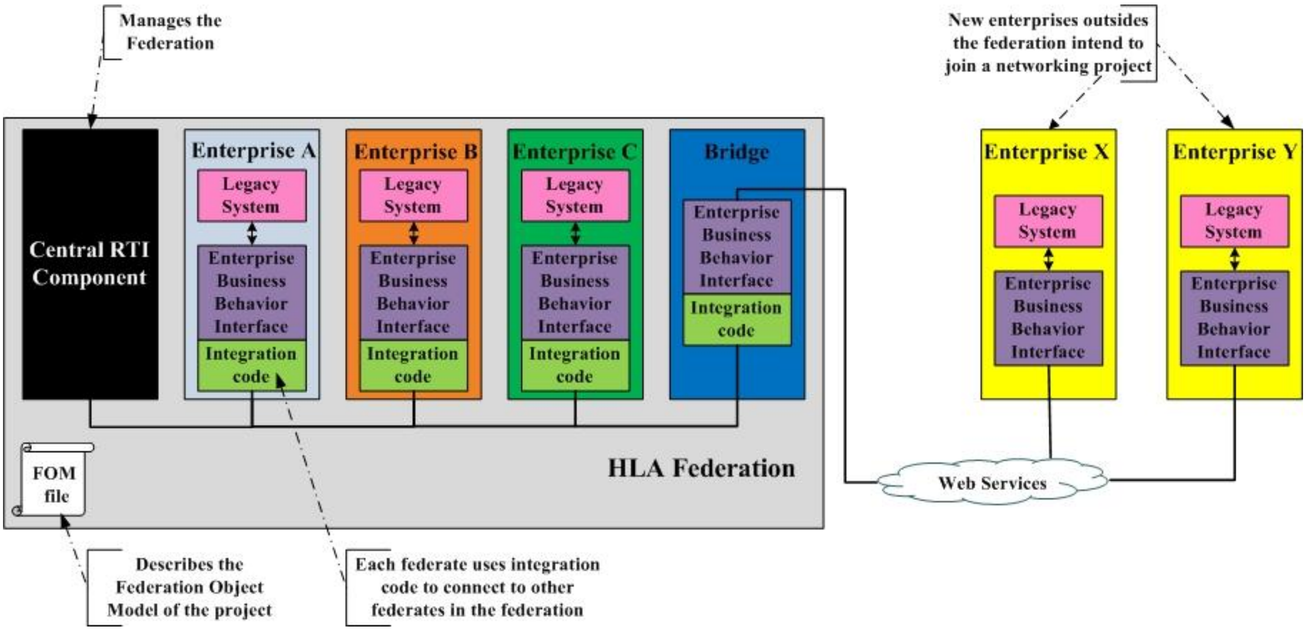


Figure 9 Architecture of HLA Evolved Web Services

As mentioned in section 5.1, after the harmonization of MDA and HLA FEDEP, we have *integration code* which provides a RTI independent API. And this API can be reused and published as common API. So, the candidates can reuse this API and follow the second scenario of model reversal, mentioned in section 5.2, to generate their own *Enterprise Business Behavior Interface* adapt to the common API. After that, the new federate outside the federation can send the information to *bridge* via the Web services interface and be synchronized by *bridge*.

5.3.3. Demonstration of HLA evolved web services

An implementation of the HLA evolved web services simulation has carried out by an internship student in our lab. It is based on POTICO RTI and Java. This simulation is implemented on Eclipse and can run in Windows NT or Unix system with JDK 1.6.0 (or higher) environment and ROTICO environment.

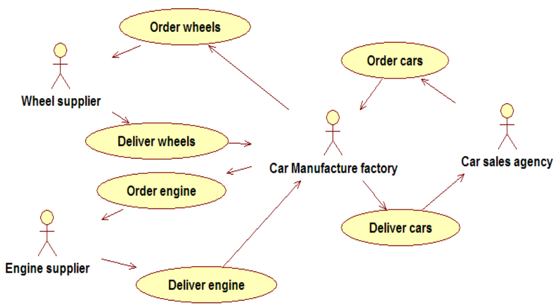


Figure 10 car manufacturing use case

This simulation describes a scenario of car manufacturing after a customer order arrives. This simulation simply simulates car manufacturer and suppliers of wheel and engine as figure 10 shows. Once the Car sales agency (*CSA*) receives an order from a client, it sends an order to the Car manufacture factory (*CMF*), then *CMF* sends orders to different suppliers, such as Wheel suppliers (*WS*) and Engine suppliers(*ES*), to get the parts or raw materials. When suppliers finish the products, they deliver them to the *CMF* where would assembles them and then deliver to the *CSA*. In this simulation, the status of the order is the concerned issue. *CSA* cares about when it can receive the cars it orders, and it concerns about the status of this order. The object class structure table shows in table 1.

Table 1: object class structure table

object class structure table		
Object root	WheelSupplier	dayToFinish
		currentState
		count
		price
	CarManufacturer	dayToFinish
		currentState
		count
		price
	EngineSupplier	dayToFinish
		currentState
		count
		price

We assume that *CSA* is the federate outside the federation, while, the *WS*, *ES* and *CMF* are the federates inside the federation. *CSA* is synchronized with *WS*, *ES* and *CMF* by *bridge*. Figure 11 illustrates the sequence diagram of this simulation. *CSA* requires the status of the order through web services, and *bridge* uses socket to connect the federation, and listen to the federation, waiting for the answer.

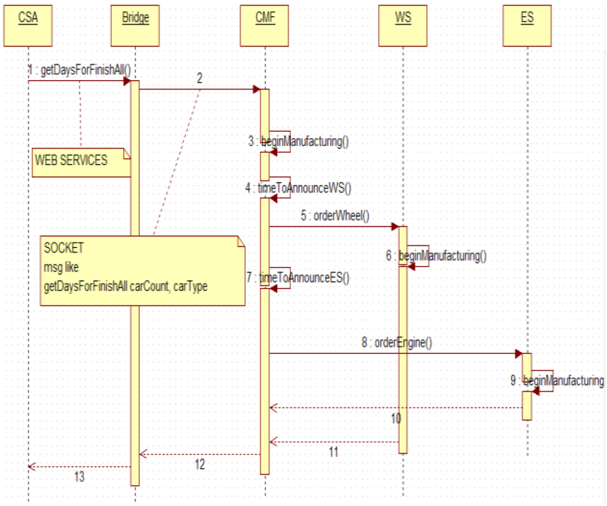
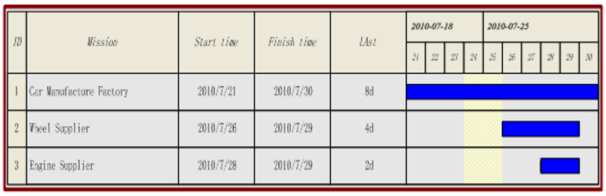


Figure 11 Sequence diagram of simulation

In this simulation, we assume that three seconds represent one manufacturing day. When *CSA* ask the probable manufacturing time for its order, it will receive a result, Gantt chart, as figure 12 shows. Then, when it requires the status of the order, it will receive the another Gantt chart with detail progress bar.



ID	Mission	Start time	Finish time	LAst	2010-07-18				2010-07-25					
					21	22	23	24	25	26	27	28	29	30
1	Car Manufacture Factory	2010/7/21	2010/7/30	8d										
2	Wheel Supplier	2010/7/26	2010/7/29	4d										
3	Engine Supplier	2010/7/28	2010/7/29	2d										

Figure 12 Simulation result

Now, the simulation runs well with laboratory data, except some gaps between our approach and the API provided by PORTICO. Recently, we are working on those gaps.

6. CONCLUSION

Based on the state-of-the-art, we have proposed a new systematic methodology, which is a valuable outcome of synthetic framework with harmonization of HLA FEDEP & MDA, Model reverse engineering and HLA Evolved Web Services. This methodology provides a new five steps process to develop models of simulation starting from conceptual enterprise models. In addition, it also bridges the gap from concepts to implementation in the field of enterprise modeling by offering a new standardized and reversible approach. And it also use the Web services to improve the performance of HLA. This methodology seems promising regarding to real enterprise information system requirement of distribution, federated interoperability and agility of adapt to dynamic context.

Up to now, this work is still in a research process. The methodology presented (each phase of HLA/MDA alignment and the model reversal process) still needs to be refined and detailed. An implementation of HLA Evolved Web Services has been carried out with an open source RTI, PORTICO. A case study will allow testing the proposed approach in an industrial context.

REFERENCES

Andreas T., 2002. Avoiding another Green Elephant — A Proposal for the Next Generation HLA based on the Model Driven Architecture. *Proceedings of the 2002 Fall Simulation Interoperability Workshop*, 02F-SIW-004.

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 15 June 2009]

Richardson L., Ruby S., 2007, *RESTful web services*, O'Reilly Media, Inc. p. 299

Chen D., Vallespir B., Doumeingts G., 1997, GRAI integrated methodology and its mapping onto generic enterprise reference architecture and methodology, *Computers in Industry*, Volume 33 Issue 2-3.

Doumeingts G., Ducq Y., 2001, Enterprise Modelling techniques to improve efficiency of enterprises, *International Journal of Production Planning and Control - Taylor & Francis*, Vol. 12, number 2, pp 146-163

Elvesæter B., Hahn A., Berre A., Neple T., 2007.Towards an Interoperability Framework for Model-Driven Development of Software Systems. *Interoperability of Enterprise Software and Applications*, 409-420.

Favre L., Martinez L., Pereira C., 2008. MDA-Based Reverse Engineering of Object Oriented Code. *SERA'08*, 153-160.

Gorka B., Larrucea X., Elvesæter B., Neple T., Beardsmore A., Friess M., 2007, *A Platform Independent Model for Service Oriented Architectures*, Enterprise Interoperability, Ed, London, S., pp. 23-32. ISBN 978-1-84628-713-8

IEEE std 1516.2-2000, 2000. *IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) - Federate Interface Specification*, New York: Institute of Electrical and Electronic Engineers.

IEEE std 1516.3-2003, 2003. *IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) - Federation Development and Execution Process (FEDEP)*, The Institute of Electrical and Electronic Engineer.

Jean B., 2005. On the Unification Power of Models. *Software and Systems Modeling*, volume 4, number 2, pp 171-188

Jouault F., Jean B., Mikaël B., 2009. Towards an advanced model-driven engineering toolbox. *Innovations in Systems and Software Engineering*, volume 5, pp 5-12.

Möller B., Clarence D., Mikael K., 2007, Developing Web Centric Federates and Federations using the HLA Evolved Web Services API. *Proceedings of 2007 Spring Simulation Interoperability Workshop*, 07S-SIW-107.

Möller B, Löf S., 2005, Mixing Service Oriented and High Level Architectures in Support of the GIG. *Proceedings of the 2005 Spring Simulation Interoperability Workshop*, 05S-SIW-064.

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].

OMG, 2010. *Architecture Driven Modernization (ADM): Knowledge Discovery Meta-Model (KDM) v1.2*, OMG, Document number: formal/2010-06-03. Available from: http://www.omg.org/spec/KDM/-1.2 [accessed 15 July 2010].

POTICO, 2009, *Developer Documentation*. Available from: http://porticoproject.org/index.php?title=-Developer_Documentation [accessed 15 July 2010]

Shawn P., 2003. The Next Step Applying the Model Driven Architecture to HLA. *Proceedings of the 2003 Spring. Workshop*, 03S-SIW -123.

Touzi J., 2007, *Aide à la conception de Système d'Information Collaboratif support de l'interopérabilité des entreprises*, PHD thesis, Institut National Polytechnique de Toulouse.

Trbovich S., Reading R., 2005. Simulation and Software Development for Capabilities Based Warfare: An Analysis of Harmonized Systems Engineering Processes. *Proceedings Spring Simulation Interoperability Workshop*, 05S-SIW-106.

Tu Z., Zacharewicz G., Chen D., 2010. Unified Reversible Life Cycle for Future Interoperable Enterprise Distributed Information Systems. *IESA 2010-Interoperability for Enterprise Software & Applications 2010*, 57-66, April 14-15, Coventry (UK).

Ullberg J., Chen D., Johnson P., 2007. Barriers Driven Methodology For Enterprise Interoperability. *IFIP International Federation for Information Processing*, 453-460.

Zacharewicz G., Chen D., Vallespir B., 2008. HLA Supported Federation Oriented Enterprise Interoperability, Application to Aerospace Enterprises. *Proceeedings of 2008 International Simulation Multiconference EuroSISO*, 08E-SIW-074, 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. 329-335, April 22-23, Beijing (China)

AUTHORS BIOGRAPHY

Zhiying TU is PHD student of University Bordeaux 1. His research subject is modelling and simulation of interoperable distribute enterprises. He recently focused on distribute simulation based on HLA and model reverse engineering.

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, WinterSim...) and SCS Simulation journals.

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.

Zhenchuan Zhang is Master student of University Bordeaux 1. Recently, he is doing his internship in IMS/LAPS. The subject of his internship is "Distributed enterprise information system interoperability based on HLA evolved SOA".

ADVANTAGE OF MOBILE TRAINING FOR COMPLEX SYSTEMS

Agostino G. Bruzzone, Marina Massei,
MISS DIPTEM University of Genoa, Italy
Eamil {agostino, massei}@itim.unige.it
URL www.itim.unige.it

ABSTRACT

This paper proposes the critical issues related to the innovative use of simulation for Mobile Training. The authors propose a methodology for developing simulators for training purposes; special attention is proposed on providing tailored solution for training activities to be conducted in different frameworks. The focus of such a research is to identify proper approach for development of a mobile training solution able to be successfully applied to trainees from different countries (i.e Mediterranean Area); many aspects, such as language, culture, procedures, have to be analyzed as well as possible benefits using simulators for training (i.e. cost reduction)

Fig.1 Mobile Training Solution based on 40' Shelter

1. INTRODUCTION

Training is a critical issue in many different framework; in fact currently new concept such as mobile training are evolving thanks to the support of new enabling technologies; therefore it is important to identify the aspects and requirements for these applications, from this point of view it is critical to proceed on the definition.
Vavoula and Sharples say that there are “three ways that learning can be considered mobile: in term of space, in different areas of life and with respect of time”. Today there is a growing importance in developing new methodologies for delivering training services: mobile training laboratories, mobile training programs, ad hoc equipment and trainers. In general sense, it is considered an efficient means of providing education and training even if affected by limitations (i.e. devices, design)

In fact, some difficulties rise designing models for specific training scenario, such as: formalization level of highly specialized expertise, competencies based on concrete know-how or on theoretical knowledge, etc. Usually, trainees need a mix of practical competencies, experience, know-how and theoretical competencies. According to Pieri and Diamantini a possible solution could be a blended learning model, which is a combination of different approaches and strategies to make teaching/learning process effective and “customized”. By using mobile training and lectures it is possible to get the most from advances in technology and traditional well known face to face training
Based on this possibility, training programs should be designed taking into account the integration of different educational methods, from traditional classes to distance learning, with special attention on lifelong learning.

2. MOBILE TRAINING DESIGN

In order to develop an effective model, useful for the development of training materials and the design of procedures for mobile training, it is fundamental to have a clear understanding of the mobile training as a mix of technology, human capacities and social interaction [Marguerite L. Koole].

To design effective mobile training modules it is important to assess the utilization of the following aspects within a mobile training system:

- Device: is important to assess these characteristics because mobile training device, such as a simulator, provide the trainee and the training task with an interface
- Trainee: take into account trainee's abilities, knowledge, motivations
- Social aspect: takes into account the processes of social interaction and cooperation. The rules of cooperation are fundamental in order to exchange information and acquire knowledge.

In fact only the interaction among the cited aspects provides the trainees with the proper feedback, which is the only way to acquire and maintain knowledge, experiences etc.

Therefore, a simulator for training purposes should:

- Be "user friendly" in order to increase success rates, because the trainee could concentrate on the tasks rather than the tool
- Make possible information exchange and collaboration among people with different goals and purposes
- Take into account the needs of trainees as people with different backgrounds, cultures and environments
- Be effective enabling trainees to interact with trainers, courses and virtual environments

Summarizing, an effective mobile training results from the integration of simulator, trainee and social aspects.

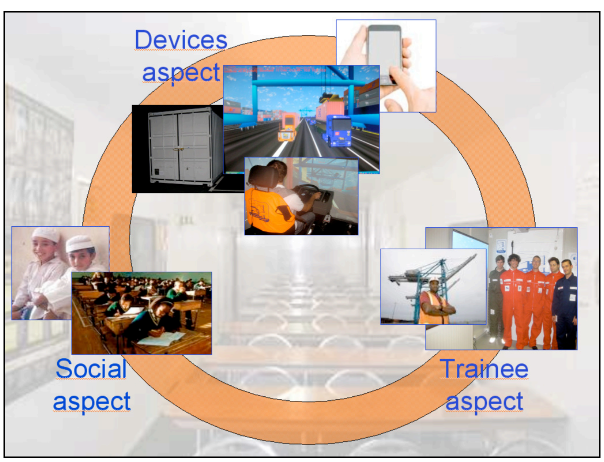


Fig. 2 Aspects to be considered for an effective Mobile Training

3. A MOBILE TRAINING EXAMPLE: VIRTUAL PORT FOR REAL OPERATORS

The authors experienced an innovative mobile training solution by developing a new training equipment for ports titled ST_VP (Virtual Port developed by Simulation Team); ST_VP is an HLA-based Real Time Distributed Simulation application for Training purposes, taking into account operators training and education, handling safety, operative efficiency.

Based on these conditions, an entire virtual world has been designed including surrounding areas, roads, container yard, different type of yard cranes, ship cranes, ships etc. In addition, ST_VP is based on modular approach that guarantee extensibility and interoperability for the dissemination of simulation techniques as Mobile Training Tool.

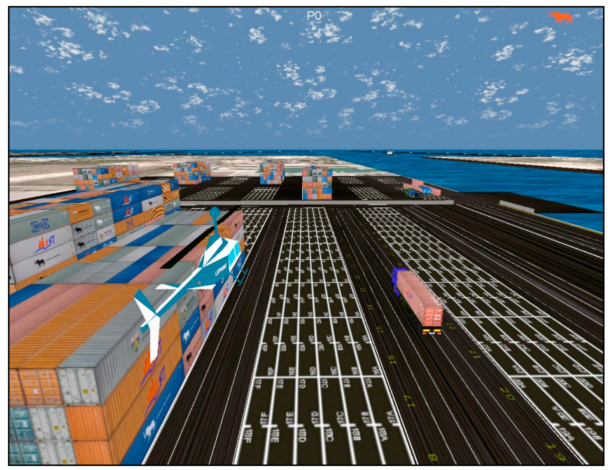


Fig.3 - Virtual Yard with Virtual Trucks and Virtual Copter

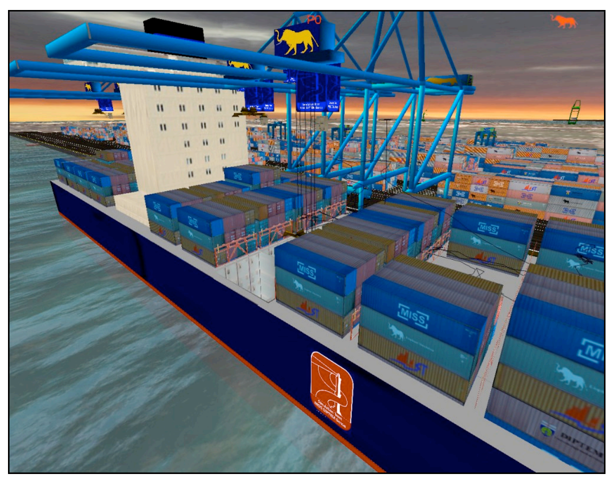


Fig. 4 - Virtual Portainer operating on Virtual Ship

The ST_VP structure is based on High Level Architecture and allows to connect multiple models representing different systems where to conduct the training, by this approach it becomes possible to generate scenarios with multiple objects intereacting. This aspect represents a real important concept in training simulation, moving from training on a single system to an environment for cooperative and competitive training. In fact ST_VP deals mostly with Ports, Terminals, Cranes and vehicles with different purposes: training single operators, training the operative squad in procedures, etc.

For these reason a wide range of configurations and operative applications are available, from operating stand alone on a single workstation to creating a federation of cranes interacting in a networks.

Therefore through simple MMI (man-machine interface) it is possible to configure a large set of position that can quickly changed from cranes to truck for creating complex cooperative scenarios; the ST_VP is designed to integrate scalable solutions (i.e. video games interfaces, laptops, workstations, 6 degrees of freedom mobile platforms,etc); by this approach it becomes possible to define cooperative operations where different teams are working concurrently in order to improve efficiency.

In order to support the mobile training several features have been introduced in ST_VP:

- Scalable Hardware Platforms and MMI: this allows to create different configurations able to be relocated quickly based on the current training needs; it is possible to create sessions with extended number of players very quickly.
- Mobile Training Framework: ST_VP is encapsulated in a real Container (40' high cube) that has been transformed in a mobile lab including all the ST_VP elements (computers, mobile platforms, MMIs, biomedical devices, interactive whiteboard, network infrastructure, cameras etc.); the ST_VP Container is designed to operate in extreme weather conditions and it is powered by flexible solution enabling it to operate in Europe, America or Africa; in addition the ST_VP Container have security solutions for protecting him against threats (i.e. firewalls, cameras, alarms, gps).
- Integration with blended education solutions: by this approach it is possible to remotely access the ST_VP Container, wherever in the world, seeing internal activities, communicating with the trainer/trainees and even downloading actions and exercise for remote analysis and evaluation; the trainees are enabled to access a e-learning platform based on Moodle for reviewing exercises and notes from the instructor.
- Tailoring Capabilities: the modular nature of ST_VP, the library of entities and the easy capability in changing the scenario allows to readapt the training to different ports and terminals with different cranes/vehicles, different procedures and different layouts.

In fact in the case of ST_VP it was decided to identify different type of elements supporting mobile training from two different point of view: Direct Supports devoted to help directly the mobility of the training infrastructure and framework service supporting mostly the adaptability of the simulator for being used in different place, so providing him opportunities and capabilities for being relocated; in this sense it is possible to classify in these two categories the following elements:

Direct Support

- Mobile Lab in a Standard 40' Container for easy Shipping
- HVAC ("Heating, Ventilating, and Air Conditioning") designed for facing extreme conditions for operating in any location
- Flexible Power able to support the different standards that are available in the different countries and continents
- Compact CAVE: cave 270 horizontal degrees and 135 vertical degrees operating within 2.2 x 2 meters
- Easy Packaging for shipping of the components in compact solutions inside (tailored configurations for specific users) or outside (lean configurations for basic use) of container

Framework Services

- HLA and Modular Approach for easy reconfiguration of scenarios and players
- Scenario Tailoring and Library supporting easy configuration on the specific training scenarios
- Integration with several user interfaces allowing to access both game engine and professional platforms for extending the number of players
- Blended Education Platform possibility to remote access to the services by trainers (remote instructor) and trainees (reviewing their activities and educational material)
- Biomedical Device Integration supporting direct measure of the trainee physical conditions and stress level during sessions provide support to complete quick comparison among different training sessions in different frameworks

The ST_VP federation includes several federates:

Portainer: allows the operator to practice a gantry crane in different scenarios. The operator can virtually load and unload container from a ship, in a virtual dock where different portainers work simultaneously.

Control & Debriefing: the trainer is provided with capability to navigate the virtual environment and to control the scenario configuration (i.e. set number of containers and operative straddle carriers, number and type of ships in port and number of trains) and to set weather conditions (i.e. time, weather condition, wind and sea condition). In addition the trainer is enabled to activate on-line with the on-going simulation simultaneous debriefing sessions for immediate review of results.

Truck, reach stacker, straddle carrier, wheel transtainer, rail transtainer, bridge crane, heavy crane federates: they dramatically improve the quality of the training session enabling each trainee to practice different vehicles/cranes in multipurpose operation.

The simulator automatically generates the missions: the user can define for every single operator which container has to be moved and the new position/destination to be reached.

The mission can be “multiple destinations”, in other words the trainee could be asked for moving the container from a yard allocation or a truck trailer and back again. Also, different levels of complexity can be selected, extra-move included, based on the necessities.

The adopted solution for the federation architecture enable the interaction among several vehicles.

Furthermore, the authors use extensively a network of experts in Simulation for Port Logistics as well as DIPTEM laboratories in order to guarantee a successful VV&A.

Fig. 5 Virtual Operation involving Virtual Ship and Virtual Truck

4. SIMULATION: AN AFFORDABLE MEANS OF PROVIDING TRAINING

The present study represents an interesting example in the field of Simulation for Mobile Training, the final users take advantage from low cost interactive distributed HLA-based environment which runs on PC equipped with simple devices (basic configuration) or becomes a mobile classroom once containerized with mobile platform and cave (full configuration).

In fact Ports represent a wide set of scenarios where different configurations are present: very different processes using often similar devices; in addition Ports are characterized by different size and needs, so it is very important to tailor the training; for instance ST_VP is currently in use in two Mobile Labs and at the same time are operating in light configuration for specific purposes in several continents.

In general sense, the “mobile” approach makes possible to extend the application of simulation as training support system in new areas, effectively facing interaction, cooperation, competition in a wide range of scenarios, with significant cost saving as well.

Fig. 6 Effective Mobile Training Solution: ST_VP

In the proposed case, ST_VP represents an innovative mobile training based on simulation; therefore with reference to port crane training, it is presented the benefit related to the use of a mobile training respect traditional solutions; the cases are represented by:

- Innovative Mobile training based on Simulation
- Traditional Mobile Pack with Crane Simulator
- Traditional Containerized Crane Simulator
- Fixed Solution for Crane Simulator
- Traditional Training on a Real Crane

The target functions to be evaluated are the number of ports to be tested and the number of people to be trained, in addition another important factor it is the time of direct training (operator interacting with the equipment) vs. the frontal lectures and observation of other people operating; another important factor is related to training costs. It is important to measure even the capability to reproduce a specific scenario; so these factors are estimated as following just considering the crane operation exercises:

$$T_{tot} = \begin{cases} \alpha < 1 & n_s \left[T_{tr} + n_p \left(T_{tvi} + T_{tri}\right)\right] + T_{st} + T_{ta} + T_{sh} \\ \alpha = 1 & n_s \left[T_{tr} + n_p T_{tri}\right] \end{cases}$$

$$N_{ports} = \frac{uc \cdot P}{T_{tot}}$$

$$Ex = T_{tr} \cdot \left(\frac{\alpha \cdot rcr + (1-\alpha) \cdot vcr \cdot pu}{n_p} \right)$$

$$C_{tot} = \begin{cases} \alpha < 1 & T_{tot}\left(Is + (\alpha \cdot Arcr + lc) + (1-\alpha) \cdot Avcr\right) + C_{st} + C_{ta} + C_{sh} \\ \alpha = 1 & T_{tot}\left(Is + Arcr + lc\right) \end{cases}$$

T_{tot}	Total Training Time for a Site
n_s	Number of sessions on a Site
α	Percentage of time on real crane respect Simulation
T_{tr}	training time
uc	use of the time frame for training (i.e. 8 hours per days, 5 days/week)
p_u	number of users concurrently training on the simulator
lc	production losses due to unavailability of the real crane
rcr	number of real cranes
vcr	number of virtual simulators
T_{tvi}	student transition time on simulator
T_{tri}	student transition time on real crane
T_{st}	setup time for the mobile solution
T_{ta}	time for tailoring scenario for the mobile solution
T_{sh}	time of shipping for the mobile solution
P	Period available for training
N_{ports}	Number of serviced Ports over P timeframe
Ex	Trainee time operating directly
n_p	number of people in a training session
C_{tot}	Total Costs on a Site including all sessions
Is	Instructor Cost
Arcr	Time Cost for using real crane
Avcr	Time Cost for using virtual crane
C_{st}	setup cost for the mobile solution
C_{ta}	cost for tailoring scenario for the mobile solution
C_{sh}	cost of shipping for the mobile solution

Table I - Comparison Scenario with Tailoring

		Mobile Training	Containerized Simulato	Pack Simulator	Fixed Simulation	Real Crane	
ns	Number of sessions on a Site	3	3	3	3	3	[sessions]
alfa	percentuage on real crane	0.5	0.5	0.5	0.5	1	
Pu	parallel trainees on simulator	4	1	1	1	1	[people]
Ttr	training time	40	40	40	40	40	[hours]
Ttvi	student transition time on simulator	2	2	2	2	0	[minutes]
Ttri	student transition time on real crane	0	0	0	0	2	[minutes]
Tst	setup time for the mobile solution	0.5	1	0.2	0	0	[days]
Tta	time for tailoring scenario for the mobile solution	1	20	20	20	0	[days]
Tsh	time of shipping for the mobile solution	3	3	3	3	0	[days]
rcr	number of real cranes	0	0	0	0	1	
vcr	number of simulated cranes	4	1	1	1	0	
P	Period available for training	6	6	6	6	6	[months]
np	number of people in a training session	4	4	4	4	4	
Is	Instructor Cost	50	50	50	50	50	[Euro/h]
Arcr	Time Cost for using real crane	171	171	171	171	171	[Euro/h]
Avcr	Time Cost for using virtual crane	175	263	44	219	175	[Euro/h]
Cst	setup cost for the mobile solution	1000	1000	1000	1000	0	[Euro/movement]
Cta	cost for tailoring scenario for the mobile solution	1000	20000	20000	20000	0	[Euro/scenario]
Csh	cost of shipping for the mobile solution	1000	2500	1000	64000	0	[Euro/movement]
lc	loss of productivity	0	0	0	0	2325	[Euro/hour]

Table II - Results on Scenario with Tailoring

		Mobile Training	Containerized Simulato	Pack Simulator	Fixed Simulation	Real Crane	
Ttot	Total Training Time for a Site	228.8	696.8	677.6	672.8	120.4	[hours]
Nports	Number of serviced Ports over P timeframe	4.62	1.52	1.56	1.57	8.77	[ports]
Ex	Trainee time operating directly	25	10	10	10	10	[hours]
Ctot	Total Costs on a Site including all sessions	54,099	209,682	128,753	250,015	306,566	[euro/site]

It is important to note that tailoring represent a critical factor for training effectiveness; therefore this factor is hard to be effective in real crane simulator: if the real crane works off line of the real production it not very realistic (i.e. not operating on a ship, not lanes of trucks) and introduces very high costs, viceversa operating on real procedure (real crane, real ship etc.) it involves high risk in damaging high value equipments as well as creating problems in operations; for instance in a sector where cranes have a cost of 1/10 of the ports the impact of maintenance costs related to training session was estimated in about 40 kEuro/year per crane; for traditional simulators tailoring is very limited too, due to the predefined scenario and the impossibility to create the procedures on-going in each terminal; therefore training is very peculiar and the above scenario is estimated based on mean values by subject matter expert estimations (Bruzzone, Vio, Capasso 2000; Bruzzone, Tremori, Capasso 2009).

Therefore it is interesting to present even the results without considering any requirement for tailoring that are reported in table III and IV.

Table III - Comparison Scenario without any Tailoring

		Mobile Training	Containerized Simulato	Pack Simulator	Fixed Simulation	Real Crane	
ns	Number of sessions on a Site	3	3	3	3	3	[sessions]
alfa	percentuage on real crane	0.5	0.5	0.5	0.5	1	
Pu	parallel trainees on simulator	4	1	1	1	1	[people]
Ttr	training time	40	40	40	40	40	[hours]
Ttvi	student transition time on simulator	2	2	2	2	0	[minutes]
Ttri	student transition time on real crane	0	0	0	0	2	[minutes]
Tst	setup time for the mobile solution	0.5	1	0.2	0	0	[days]
Tta	time for tailoring scenario for the mobile solution	0	0	0	0	0	[days]
Tsh	time of shipping for the mobile solution	3	3	3	3	0	[days]
rcr	number of real cranes	0	0	0	0	1	
vcr	number of simulated cranes	4	1	1	1	0	
P	Period available for training	6	6	6	6	6	[months]
np	number of people in a training session	4	4	4	4	4	
Is	Instructor Cost	50	50	50	50	50	[Euro/h]
Arcr	Time Cost for using real crane	171	171	171	171	171	[Euro/h]
Avcr	Time Cost for using virtual crane	175	263	44	219	175	[Euro/h]
Cst	setup cost for the mobile solution	1000	1000	1000	1000	0	[Euro/movement]
Cta	cost for tailoring scenario for the mobile solution	0	0	0	0	0	[Euro/scenario]
Csh	cost of shipping for the mobile solution	1000	2500	1000	64000	0	[Euro/movement]
lc	loss of productivity	0	0	0	0	2325	[Euro/hour]

Table II - Results on Scenario without any Tailoring

		Mobile Training	Containerized Simulator	Pack Simulator	Fixed Simulation	Real Crane	
Ttot	Total Training Time for a Site	204.8	216.8	197.6	192.8	120.4	[hours]
Nports	Number of serviced Ports over P timeframe	5.16	4.87	5.34	5.48	8.77	[ports]
Ex	Trainee time operating directly	25	10	10	10	10	[hours]
Ctot	Total Costs on a Site including all sessions	47,739	61,428	33,131	112,287	306,566	[euro/site]

It is evident the great benefits of innovative mobile training solution such as that one proposed by ST_VP respect traditional training on real crane and even respect old style simulators as summarized by the following figures.

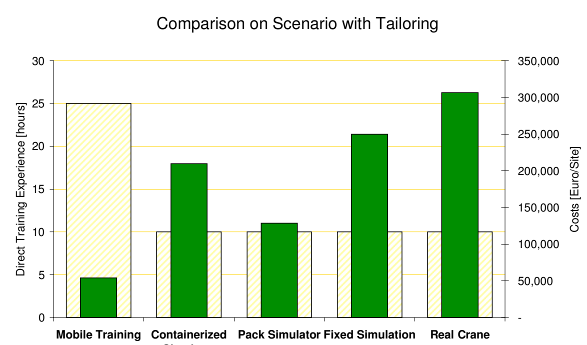


Fig. 7a Benefits of Mobile Training in Effectiveness and Costs with Scenario Tailoring

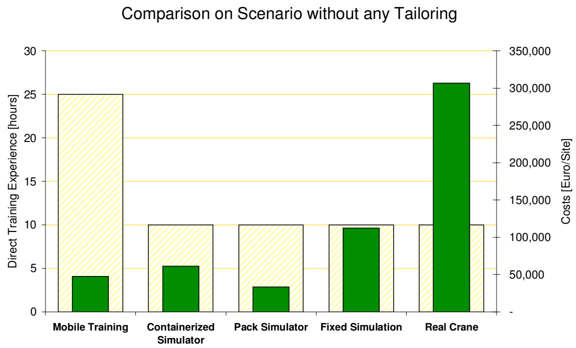


Fig. 7a Benefits of Mobile Training in Effectiveness and Costs without any Scenario Tailoring

CONCLUSION

This papers introduces the concepts of Mobile Training and provides an example of a simulator that is enabling these concepts; from this point of view it is important to identify the features and requirements that support the Mobile Training not only in term of direct support, but even as framework service guaranteeing usability within this paradigm.

In fact the example of ST_VP, developed by the Simulation Team, provides an unique opportunity, due to the high level of flexibility, to improve existing training procedures or to develop new ones, giving the authors the chance to tackle a very serious problem in the related application area. The approach outlined here provides an overview of the elements that supported a Mobile Training solution taking into account the positive follow up of innovative techniques applied to a sector that it is not very familiar with simulation such as maritime environments; the results obtained demonstrated significant benefits in term of training effectiveness; therefore the most interesting benefits obtained are related to the big increase in number of terminals, ports and operators that could benefit from this Mobile Training approach with a cost reduction. Finally the proposed research aims to deliver a training approach that have potential for further implementation in other sectors.

ACKNOWLEGMENT

The authors are very glad to thank Paolo Fadda, Gianfranco Fancello and Salvatore Capasso for their support in applying these concepts development within training programs for Port Activities.

REFERENCES

[1] Vavoula G.N. Sharples M.Kleos (2002) “A personal Mobile, Knowlwdge, and Learning Organisation System” In Milrad, M.Hoppe, H.U. and Kinshuk (Eds) in Proc.of the IEEE International Workshop on Wireless and Mobile Technologies in Education” WMTE’02 pp.155-156 Los Alimatos, USA

[2] Pieri M. Diamantini D.(2009) “From E-learning to Mobile Learning: New Opportunities” in “Mobile Learning Transforming the Delivery of Education and Training” Edited by Mohamed Ally, AU Press, ISBN 978-1-897425-44-2

[3] Koole M.L. (2009) “A Model for Framing Mobile Learning” in “Mobile Learning Transforming the Delivery of Education and Training” Edited by Mohamed Ally, AU Press, ISBN 978-1-897425-44-2

[4] Bluemel E. (1997) "Managing and Controlling Growing Harbour Terminals", SCS Europe BVBA, Ghent, Belgium

[5] Bontempi, Gambardella, Rizzoli (1997) "Simulation and Optimization for Management of Intermodal Terminals", Proc. of ESM97, Istanbul

[6] Bruzzone A.G. (1995) "Fuzzy Logic and Genetic Algorithms Applied to the Logistical and Organisational Aspects of Container Road Transports", Proc. of ESM95, Praha, June 5-7

[7] Bruzzone A.G., Kerckhoffs (1996) "Simulation in Industry", Genoa, Italy, October, Vol. I & II, ISBN 1-56555-099-4

[8] Bruzzone A.G., Giribone P. (1998) "Decision-Support Systems and Simulation for Logistics: Moving Forward for a Distributed, Real-Time, Interactive Simulation Environment", Proceedings of the Annual Simulation Symposium IEEE, Boston, 4-9 April

[9] Bruzzone Agostino (1999) "Port Terminal Simulators as Main Supports for Design, Training & Management", Proc. of Port Logistics99, Alexandria, Egypt, Feb 14-16

[10] Bruzzone A.G., Merkuryev Y.A., Mosca R. (1999) "Harbour Maritime & Industrial Logistics Modelling & Simulation", SCS Europe, Genoa, ISBN 1-56555-175-3

[11] Bruzzone A.G, Mosca R., Revetria R., Rapallo S. (2000) "Risk Analysis in Harbour Environments Using Simulation", International Journal of Safety Science, Vol 35, ISSN 0925-7535

[12] Bruzzone A.G., Vio F., Capasso S. (2000) "MOSLES: Modelling & Simulation for Transport and Logistic Educatonal Support", Proceedings ITEC2000, The Hague, April

[13] Bruzzone A.G., Tremori M., Capasso S. (2009) "Cybersar Training Procedures", DIPTEM Technical Report, Genoa, Italy

[14] Bruzzone A.G., Mosca R., Revetria R. (2002) "Cooperation in Maritime Training Process using Virtual Reality Based and HLA Compliant Simulation", Proceedings of XVIII International Port Conference, Alexandria Egypt, January 27-29

[15] Bruzzone A.G., P. Fadda, G. Fancello, E. Bocca, G. D'Errico, A. Tremori (2008) "Ship-To-Shore Gantry Crane Simulator Design ",*Proc. of HMS2008*,September 17-19, 2008, Campora S.Giovanni (CS),Italy
[16] De Ruit, Schuylebburg, Ottjes (1995) "Simulation of shipping traffic flow in the Maasvakte port area of Rotterdam", Proc. ESM95, Prague
[17] Fleming D.K. (1997) "World Container Port Ranking", Maritime Policy and Management, Vol. 24, No. 2, pp. 175-181
[18] Frankler E.G. (1987) "Port Planning and Development", John Viley and Sons, New York
[19] Hayuth Y., Pollatschek M.A., Roll Y. (1994) "Building a Port Simulator", SIMULATION, vol. 63, no. 3, pp. 179-189
[20] Koh P.H., Goh J.L.K., Ng H.S., Ng H.C. (1994) "Using Simulation to Preview Plans of a Container Port Operation", Proceedings of Winter Simulation Conference, Lake Buena Vista, Florida, December
[21] Merkuryev Y., Bruzzone A.G., Merkuryeva G., Novitsky L., Williams E. (2003) "Harbour Maritime and Multimodal Logistics Modelling & Simulation 2003", DIP Press, Riga, ISBN 9984-32-547-4 (400pp)
[22] Merkuriev Y., Bruzzone A.G., Novitsky L (1998) "Modelling and Simulation within a Maritime Environment", SCS Europe, Ghent, Belgium, ISBN 1-56555-132-X
[23] Nevins M.R., Macal C.M., Joines J. (1998) "A Discrete-Event Simulation Model for Seaport Operations", SIMULATION, vol. 70, no. 4, pp. 213-223, April
[24] Ottjes J.A., Hengst S., Tuturima W.H. (1994) "A Simulation Model of a Sailing Container Terminal Service in the Port of Rotterdam", Proc. ESM94, Barcelona
[25] Rizzoli A.E., Gambardella L.M., Bontempi G. (1997) "Simulation of an Intermodal Container Terminal to assist Management in the Decision-Making Process", Proc. of MODSIM9, International Congress on Modeling and Simulation, Hobart, Tasmania
[26] Teo Y.M. (1993) "Simulation and Graphics Animation in Port Design", Proceedings of ESM93, Lyon, France
[27] Thiers G., Janssens G. (1998) "A Port Simulation model as a Permanent Decision Instrument", SIMULATION, Vol. 71, no. 2, pp. 117-125, Villefranche L., Pecuchet J.P, Serin F. (1994) "Service Processes for Container Terminal Simulation", Proc. ESM94, Barcelona
[28] Vavoula, G., & Sharples, M. (2002). "KLeOS: A person, mobile, knowledge and learning organization system", *Proc. of the IEEE Int.Workshop on Mobile and Wireless Technologies in Education* August 29-30, Vaxjo, Sweden, 152-156

AUTHORS BIOGRAPHY

Agostino G. Bruzzone 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). He 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).

Marina Massei is currently enrolled in DIPTEM University of Genoa as member of the Simulation Team of Prof.Agostino Bruzzone. She was Finance Control & Administration Director of MAST an High Tech company working on M&S.
She had experiences as Project Manager in Modeling Logistics and Retail; she was involved in Master, PhD Programs as well as in Professional Course in PM and M&S.
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.

Simulation Integrated with Web 3.0 as Smart Support for Command and Control

Agostino G. Bruzzone
MISS Genoa
Genoa, Italy
Email agostino@itim.unige.it
URL www.itim.unige.it

Roberto Cianci
DIPTEM University of Genoa
Genoa, Italy
Email cianci@diptem.unige.it
URL www.diptem.unige.it

Francesco Longo
MSC-LES University of Calabria
Cosenza, Italy
Email f.longo@unical.it
URL www.msc-les.org

Giovanni Cantice
RAV 80 Italian Army
Cassino, Italy
Email giovanni.cantice@gmail.comhq.nato.int
URL www.esercito.difesa.it

Simonluca Poggi
SimulationTeam
Savona, Italy
Email poggi@simulationteam.com
URL www.simulationteam.com

ABSTRACT

This paper propose some preliminary research on the potential of the Web 3.0 concept into military training; the authors present an overview of this technology considering the application framework and developed the requirements for getting benefits of combining Simulation and Web 3.0; a proposal about a scenario and model is presented as case study to verify and validate this approach.

INTRODUCTION

The Web 3.0, as semantic web, allows to support quick and effective identification of resources and condition by correlating the info and data; in fact usually no single database or network source contains all the critical information for complex problems. A critical question it is if this capability could be applied to the training process of military simulation; based on subject matter experts operating in this area the answer results clearly positive; in fact to create a semantic infrastructure could results in an important support to simulate and facilitate data mining; in fact by this approach it becomes possible to show relationship among resources and services and, through intelligent support algorithms, to underline specific result by creating effective opportunities able to associate different contents with resources on the dbase and web (foir this context obviously mostly intranet). This is especially true for the intelligence field where the right gathering, correlation and spreading of the right information is determinant in a asymmetrical full spectrum military environment. Semantic technologies allows users to easily build new web sites and application to manage their needs, presenting various threads in a composite picture. The semantic web service/application could serve also as application integration mechanism by using meta-data that are interpreted in the meaning, realizing a common and universal understanding. In fact the combination of this approach with Intelligent Agents driving Computer Generated Forces (IA-CGF) could support quick and effective creation of complex scenarios in asymmetric warfare conditions for training. The authors are developing these concepts for a demonstration devoted to training and to complete VV&A with further potential for test and analysis and decision support applications.

SEMANTIC WEB

The first step it is to consider the innovative concepts behind Web 3.0: Berners-Lee, et al (2001) wrote "a new form of web content that is meaningful to computers will unleash a revolution of new possibilities." This is the concept related to the Semantic Web.
Tim Berners-Lee (well known for inventing the World Wide Web), already presented this idea at the very first World Wide Web conference in 1994. The potentials of the Semantic Web are also strongly reaffirmed in Shadbolt et al. (2006): the authors revisit the Semantic Web idea stating that it will mainly rely on key insights, tools and techniques developed during the last 50 years of research in the field of Artificial Intelligence. Nowadays all the disciplines (life sciences, environmental sciences,

business, commerce and defence to cite a few) ask for data (potentially available on the web) integration coming from different sources and the need for intelligent information combination (according to user requirements) is constantly growing. On the actual web (syntactic web) the main difficulty lies in the interpretation of the text being searched that is not semantically annotated. In 2006, Tim Berners-Lee said that the Semantic Web has to be regarded as one of the main components of the Web 3.0 (Shannon, 2006). The main difference among Web 1.0, 2.0 and 3.0 is based on bandwidth availability that allows the use of more packages, functions and applications for unlimited time with heavy multimedia applications (video web), that were before less manageable. From this point of view Web 3.0 allows users to extend previous communication capabilities by integrating them with multimedia contents and automatic search.

In effect the Web 3.0 should include multiple aspects; among others the most important are the advanced use of artificial intelligence techniques, the organization of information in databases (for information access) and Web 3D (three dimensional environment). Web 3.0 should adopt a multiplatform technology that can be easily integrated with different Operative Systems, architectures, browsers and should use many libraries and framework applications. Therefore the first step is to improve data interoperability levels (data web), then the combination of data mining and intelligent agents should support the intelligent integration of available information (in Web 2.0 this step is carried out by the user according to multiple syntactic searches and data re-organization). Following this process Web 3.0 becomes a Service Oriented Architecture that can also include a 3D application to create a set of three dimensional environments as demonstrated in systems such as "Second Life". What is?

In this paper the authors explore the possibility and the potentials to use Web 3.0 (specifically the Semantic Web) for training processes in military simulation with particular attention to Command and Control System services. In this context the Semantic Web can strongly support the user's requirements and predict what the user means by his query and deliver more relevant researches and results. A framework that includes a Semantic Web provides a strong potential in describing data and applications. The authors propose using these concepts in a framework for supporting Intelligent Agents Computer Generated Forces in generating threats for complex urban scenarios.

The state of the art on the Semantic Web reveals a wide spectrum of research topics and results as well. Semantic Web is currently a very exciting topic for researchers in different areas: from distributed information systems to artificial intelligence. One of the most important features of Semantic Web (as well as the current syntactic Web) is related to information search. When users search for information they would like to see only relevant results. They do not like to retrieve data that are unrelated to their search string. The major problem is that in the current Web it is difficult to find good research results due to a non-semantically annotated interpretation. Algorithms that interpret the syntactic content in order to return good search results are necessary. These algorithms can decide whether or not to take user context and the search history of users into consideration.

New approaches and techniques were developed in the last few years to better search results. For instance, Hsua and Wub (2006) introduced a new approach that determines the relevance of a Web page in its context, based on the relevancy context graph. These graphs estimate the distance and the relevancy degree between the retrieved document and the given topic. By calculating the word distributions of the general and topic-specific feature words, the method is expected to preserve the property of the relevancy context graph and reflect it on the word distributions. In contrast, Godoy and Amandi (2006) developed a document clustering algorithm, named WebDCC (Web Document Conceptual Clustering) in order to acquire user profiles, taking in consideration their search history that is properly clustered.

Another difficulty of current Web is represented by the integration of the various data sources in order to produce good search results. In fact information space is scattered over several sources that have different and non-standard structures, schemas and meta-data. The concept of community helps to overcome this limit. In fact, when one queries for a specific topic, the search can focus on the data provided by an existing community around that topic, excluding unrelated ones. Benatallah et al. (2006) underlined the problem of querying heterogeneously structured data and proposed a flexible and user-centric query matching algorithm. It exploits both community descriptions and peer relationships to find e-catalogs that best match a user query.

Another relevant aspect in Semantic Web is people trust. Consider for instance people authentication and reputation problems, users' privacy, confidential information management, etc. Artz and Gil (2007-a, 2007-b) studied trust feelings in content on the Semantic Web, identifying and analysing just what factors people use to decide what content they will trust and why. In this case simulation was very useful in order to study alternative models of content trust. Concerning people trust and specifically users' privacy during e-commerce activities, Jutla et al. (2006) propose a privacy management architecture, consisting of client-side and web-side architectural data components and services which inform the user of online privacy and trust within e-commerce tasks. Client-side ontology and data structures for representing user contexts are introduced. In the context of the Semantic Web, social networks are crucial to realize a web of trust, which enables the estimation of information credibility and trustworthiness. Matsuo et al. (2007) described a social network extraction system called POLYPHONET, that extracts relations of persons, detects groups of persons, and obtains keywords for a person by using several advanced techniques. Finally, social relations are classified in order to obtain and use person-to-word relations. A particular characteristic of Semantic Web is related to the extension of Web Services by representing explicit meanings. One aspect of the formal definition of Semantic Web Services is the description of the data that are processed by Semantic Web Services. The Web Service Modelling Framework, WSMF, (Fensel and Bussler, 2002) proposes a formal representation of Web Services, that allows access machine executable semantics. Note that usually different applications domains have different conceptualizations of their specific data sets. For instance, the health care industry where numerous providers use different standards to describe their application data. In particular Dogac et al. (2006) defined formal representations of healthcare data in order to develop Semantic Web Services.

Another important element of Semantic Web Service definition is the composition of Web Services. In fact an appropriate composition of different Web Service can produce the desired result for a client. Rao et al. (2006) studied an automatic Web Services composition, based on Linear Logic and theorem proving. In addition, it is necessary to specify how Semantic Web Services are executed; in particular a life cycle determines the order of Semantic Web Services use. When a request comes from a client, an appropriate set of Web Services have to be invoked in order to satisfy the client's requirements. Different researchers have studied a Web Services discovery approach. For instance, Bianchini et al. (2006) introduces a discovery approach based on a three-layer ontology, each layer describing Web Services more specifically in the form of an abstraction hierarchy.

MILITARY APPLICATIONS

The complexity of military applications comes from the necessity of integrating more options in just one complex application. In the case of a generic Command and Control (C2) system, such applications must guarantee:

- the visualization of a referenced and accurate cartography that usually includes a series of military data (in general units, means, tools, capability or references identifications)
- the possibility of sending standard messages
- the management of databases containing operative and logistics data (in a classified environment as well)
- the possibility of operations planning by scenarios/situations investigation and by correlating collected information in order to implement the received orders and transform them into possible actions/reactions.

The C2 system can be eventually completed with the integration/interoperability with other data/systems; however the system often is not able to provide a mid-level military decision maker with a complete general or particular framework, fundamental for a proper application of the decisional process and for a quick and coherent decision. This is mostly due to the implementation of partial architectures in the C2 Systems design because they only take into consideration a specialized referenced environment (ground, sea, air, cyberspace) and they do not consider other realities (by their inherent construction). Such situations can happen due to developer's choice, opportunity or blindness.

Indeed, it is a fact that database or network resources cannot contain all the needed information. The most recent tenets on information spreading/sharing such as those foreseen by NCW/NEC (Network Centric Warfare/ Network Enabled Capability) intend to mitigate this shortfall by applying circular methodologies for information spreading and sharing so that data can be managed and used (effectively and efficiently) by a large number of users. The possibility of modeling and simulating these different approaches for information spreading and sharing could make possible a more complete and detailed analysis of any possible scenario.

From this point of view Web 3.0 techniques represent an enormous potential, especially if we imagine its application in a functional modern C2 System, oriented to NCW/NEC techniques. The semantic techniques are used to multiply database queries in different functional areas and environments and to simplify data research by using IA (Intelligent Agents) and information correlation. Data mining can be used for simulating and showing resources and services relationships or for supporting the representation of specific actors on training simulators (using IA). In addition, by using ad-hoc algorithms, it is possible to create great opportunities by associating various resource contents of different realities via intranet and to represent dynamic and complete JCOP (Joint Common Operating Picture), that can be used to populate Constructive Simulator databases and reproduce real world scenarios. The resulting output may help to quickly predict possible actions/reactions or make hypotheses in the intelligence domain and in static and asymmetric environments, and lead to understanding the great importance of correct information reading, representation, and spreading in full spectrum environments. For example, analysis applied to an urban 3D environment in order to predict possible shooting directions used by a sniper or positions which can be used to trigger IEDs (improvised explosive device) or launch ambushes. Semantic capabilities could be used to manage local information from a commander on the ground in order to create its own order of battle, collect information, and to send needed information to his subordinates to reproduce a graphic planning or to search information on internet by using assisted Web 3.0 research benefits. Web 3.0 technology could be used as mechanism to integrate data and metadata of different systems.

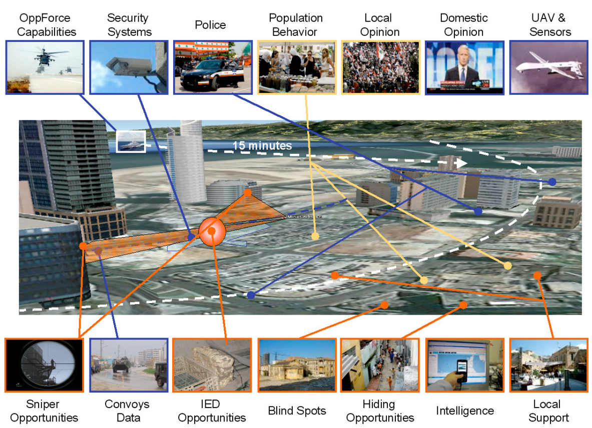


Figure 1- Queries to be elaborated by Intelligent Agents

MILITARY REQUIREMENTS WEB 3.0

As already mentioned, in the last years the development, deployment and interest for networked information systems for military applications are really increased. Referring to a multiple platform standard architecture, it is possible to observe new generations of client systems that integrate advanced communication paradigms and functionality; for instance information repository could be queried from clients with mechanism related to events and correlated in new high level events as complex information ontology and information management architecture and comprehensive security model.

Web 3.0 and related semantic web services have also received key elements from new software technology (CORBA, J2EE, C# and .NET), realizing new interfaces in a wide variety of technologies as hardware, components, internal components for operating systems, servers, etc. The possibility to map military architectures as GIG (Global Information Grid) or NCW/NEC (Network Centric Warfare/ Network Enabled Capability) with semantic web services is a real matter. In a hasty overview we can imagine a relative overlap between a general SOA (Service-Oriented Architecture) and GIG architectures as both normally use XML (eXtensible Markup Language) information representation and notification capabilities. Furthermore, some concerns should be identified in the web architecture weakness in terms of lack of support for real-time high availability delivering mechanisms or information services. To be more restrictive in the military environment there are other practical considerations such as the feasibility of scaling concept, size of development, level of communications and disruption that could be part of a military requirement.

A general GIG architecture should be planned as a "rapid targeting" or "rapid response", as an application that has to collect information and collect data from different sources. It makes clear the necessity of a meta-data repository to identify classes of servers and types of information that are available for querying as well as the subset of information that should be relevant to a generalized engagement/plan/decision. Keeping in mind that, in this case, the information sources are military computing applications or information systems, it is necessary also to identify military capabilities to be applied in standard formats. Standard formats are the key policies to the proposed approach because they allow the systems to be extended to new applications, new clients, new security policies and new systems.

In military systems, at the beginning the applications download initial data for maps, GPS, current tactical data, situation report, status report, etc., and only when a new event occur or the refresh replication mechanism is set, the data source "publishes" the new summary, normally in a data form suitable for a rapid transmission on the network. Large data object (as a high resolution picture) could be downloaded separately only if needed. In this case an important requirement is represented by the security policies: data should be filtered and enforced based on the

quality of mission. The same should happen regarding non-classified information that has to be introduced and downloaded in a specific operational map quadrant where a considerable number of classified information/data are present or, vice-versa, introduce only classified/unclassified data using a filtered bandwidth to avoid the unwanted information. In this case, the repository should collect all events and store them persistently in different DBs until they expire. During the validity period any applications might use rapidly and query the context also to associate the events to another map quadrant.
Additional "operational" capability/properties for this platform should be the ability to function robustly in a hostile environment: capability to re-configure the platform due to disruptions and provide the users with high performances in every network configuration (increase in number of clients, rates of clients, system failures, loading components, etc.).

Web 3.0 and C2 systems integration: critical issues

A more detailed list and description of the main problems in applying web 3.0 technologies in a C2 military system are reported as follows:

- Messages size: large sizes of data are destined to fail because of the problem with the bandwidth availability. In a traditional network with large bandwidth this is not a performance factor, whilst in a mobile tactical network among soldiers (as individuals) or low level command where the shared transmission capability is a decisive factor, it turns out to be the choke point of efficiency. Reduce the size and optimize the content of information becomes vital for this network as the standard Web System (WS) allows for negotiating a non-standard protocol to represent messages and would force the rules.
- Security enforcement: the servers, that embed security mechanism, do not guarantee the client about the enforcement of the message. The only guarantee is to implement an encryption policy or to send the document to a server that can respond to the client's expectations.
- Time critical events/message: in a military environment a lot of events/messages are important in a specific period of time; after this period their usefulness expires and they can be archived, disposed, erased. But in a particular situation (i.e. artillery coordinates) could be reused or turn into obsolete in less time than expected, hence the communication architecture must be able to deliver the right message at the right time, archive all valuable or history data, upload or expire other data.
- Quality of service: properties as security, robustness, performance, automatic configuration and other military specification must underlined as they are normally included in a GIG system. Scalability and performance are moreover important for the WS system.
- Reliable messaging: messages are normally a basic pillar for a military GIG system. A good WS must be include a notification application that can define a content-based publish-subscribe system and be compliant with the standard communication system properties.
- Scalability: the scalability needs have the same weight of time critical events delivering or architectural capabilities to favorite the decisional processes. Scalability becomes enormously important when we take into consideration the situation when nodes crash, messages are lost and, dynamic configuration problems present themselves, as in a neural network fundamental to maintaining the connection with the final scope.

INTEGRATION IA-CGF AND WEB 3.0

The authors are investigating the development of a framework, based on Web 3.0 technologies, that will be interrogated by IA-CGF units (Intelligent Agents Computer Generated Forces). The framework must provide the user with the evaluation of the best opportunities to set up asymmetric threats (i.e. ambushes, IED, snipers, etc.) in order to create an asymmetric scenario focused on effectiveness of the actions (including the impact of people behavior, human modifiers and network-related issues within a 3D environment). Figure 2 shows (in a possible architecture) how the IA-CGF interact with the Web 3.0 technologies as well as with multiple information sources.

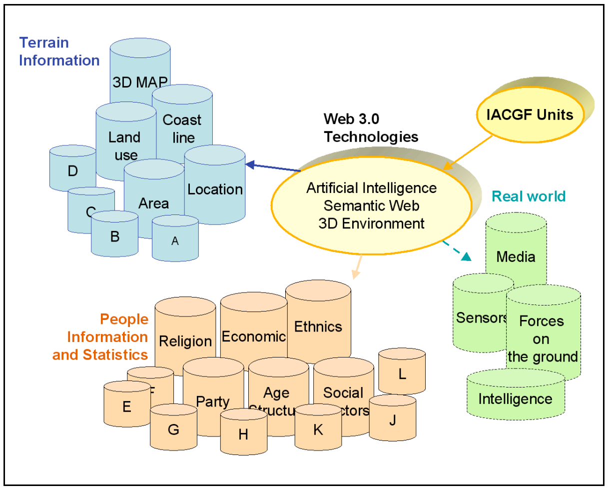


Figure 2: Intelligent Agents Navigating in Web 3.0

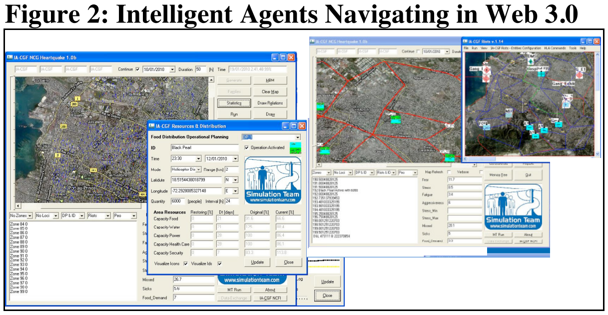


Figure 3: IA-CGF operating over complex urban scenarios

Figure 3 shows an example of a system (developed by the Simulation Team of the University of Genoa) implementing IA-CGF operating over complex urban scenarios.

Consider the case of an attack with snipers, the Web 3.0 technologies make possible an automatic identification of such an attack on the basis of a multiple factors investigation carried out over the social-terrain database by using the semantic queries. In this case, the following elements must be considered and included:

- OppForce Capabilities
- Security Systems
- Police
- Population Behavior
- Local Opinion
- Domestic Opinion
- UAV & Sensor
- Convoys Data
- Hiding Opportunities
- Intelligence
- Blind Spots
- Local Support
- Sniper Opportunities
- IED Opportunities

Noticeable fact is that IA-CGF are Computer Generated Forces managed by Intelligent Agents based on human behavior models (fear, aggressiveness, fatigue, stress, etc.). The IA-CGF Units are able to operate autonomously based on applicative sources of intelligence and human behaviors. In fact the IA-CGF units operate by collecting information from different sources based on Web 3.0 technologies (in particular Semantic techniques) and apply Artificial Intelligence algorithms in order to identify composite elements (i.e. synergy between a blind spot shooting and an hiding opportunity). In particular, IA-CGFs are expected to acquire multiple information from specific databases in a semantic way and then to use it to support decision making.

The present research is aimed to develop a demonstrator able to create scenarios that can be generated automatically by integrating IA-CGF and Web 3.0 technologies in reference to threats such as IED (Improvised Explosive Devices) allocation for a terrorist attack. The authors' goal is to provide IA-CGF units with all necessary information in order to establish the optimal IEDs position. A further step of integration could be to visualize all units in a distributed (web or GIG or other networks) 3D environment created by extracting information about the territory of interest from the users in order to provide such support not only in creating scenarios for training, but even for supporting users in AAR (After Action Review).

A wide range of information is requested to allocate IEDs, for instance:

- Territory Information: 3D Map, location, geographic coordinates, area, land boundaries, Coastline, maritime claims, climate, terrain, elevation extremes, land use, etc.
- People Information and statistics about the area: Population, Age Structure, Social, Ethnic, Religious, Economic and Political features, etc-
- Transportation: airports, heliports, pipelines, railways, roadways, waterways, ports and terminals, etc.
- Information from the real world: media and news, direct information from the field, sources of intelligence.

This gives an idea of the potential of this approach in terms of training; in addition it is evident that such architecture could be very effective in operative analysis and decision support on operational planning by integrating these technologies with real world data.

In addition, during the development of the proposed demonstrator for training, it will be possible to proceed extensively with VV&A (Verification, Validation and Accreditation). Once the demonstrator is successfully completed, it could be interesting, as a follow up, to consider adapting this approach to support decision making with integration even with real world data. Just one among possible further developments of this solution could be a decision support system and test and analysis tool based on these concepts.

CONCLUSIONS

Semantic Web and in general Web 3.0 is very challenging in improving search results and in integrating all data in a semantic way by using new techniques and approaches such as ontology of specific topics and domains. Taking into consideration sector complexity and data sensibility, Web 3.0 provides many opportunities for military training and for decisions making. In particular, the semantic techniques can be used for multiple database queries in different functional areas and environments, to simplify data research by using IA and information correlation, to improve data mining in order to simulate and show resources and services relations or to help the representation of specific actors on training simulators by using IA. The research in integrating IA-CGF with Web 3.0 for demonstrating the possibility to create automatically realistic threats for supporting training it is a very interesting opportunity; in this context it resulted critical the proper connection among military users and simulation experts for defining properly the development boundaries for being able to create a realistic and useful demonstrator with reasonable resources.

6. ACKNOWLEDGEMENTS

The authors are glad to thank Michele Turi for the support he provided in reviewing Military Requirements.

REFERENCES

- Artz D. and Gil Y., (2007-a), “A Survey of Trust in Computer Science and the Semantic Web”, Journal of Web Semantics, Volume 5, Issue 2
- Artz D. and Gil Y. (2007-b), “Towards Content Trust of Web Resources”, Journal of Web Semantics, Volume 5, Issue 4
- Bell D., Bussler C., Yang J., (2007) “The Semantic Web and Web Services”, Web Semantics: Science, Services and Agents on the World Wide Web 5 225–226
- Benatallah B., Hacid M., Paik H., Rey C.and Toumani F., (2006) “Towards Semantic-driven, Flexible and Scalable Framework for Peering and Querying e-Catalog Communities”, Information Systems archive, Volume 31, 266- 294, Issue 4 (June 2006), special issue: The semantic web and web services, Elsevier Science Ltd. Oxford, UK
- Berners-Lee T., Hendler J., Lassila O. (2001). The Semantic Web. *Scientific American Magazine*. Retrived, June 4, 2010, from http://www.scientificamerican.com/article.cfm?id= the-semantic-web
- Bianchini D., De Antonellis V., Pernici B., Plebani P., (2006). Ontology-based methodology for *e*-service discovery. Information Systems, Volume 31, Issues 4-5, pp 361-380.
- Bruzzone A.G., Cantice G., Morabito G., Mursia A., Sebastiani M., Tremori A. (2009) "CGF for NATO NEC C2 Maturity Model (N2C2M2) Evaluation", Proceedings of I/ITSEC2009, Orlando, November 30-December 4
- Bruzzone A., Giribone P., Mosca R. & Fassone B. (1999) "Project Management Tools: Simulation On Web Environment", Proceeding of Websim99, San Francisco, January 1999
- Bruzzone A., Revetria R. (2001) "Web Based Modeling as Support for Industrial Maintenance", Proceedings of FAIM2001, Dublin, July 16-18 2001
- Bruzzone A.G. (2009) "Serious Games for Training and Education on Defense Against Terrorism", Invited Paper al NATO M&S Group MSG-069 a Bruxelles sul tema "Serious Games for Training and Education on Defense Against Terrorism", Invited Speech at NATO M&S Group - MSG-069, Bruxelles, Belgium October 15, 16
- Bruzzone A.G. (2006) "Distributed Interoperable Simulation for Training", Invited Presentation on Simulare Meeting, Athens, February 16-18
- Bruzzone A.G., (2007) "Generated Forces in Interoperable Simulations devoted to Training, Planning and Operation Support", Keynote Speaker at EuroSIW2007, Santa Margherita Ligure, Italy, June
- Casey M., Pahl C. (2003) “Web Components and the Semantic Web”, Electronic Notes in Theoretical Computer Science 82 No. 5, URL: http://www.elsevier.nl/locate/entcs/volume82.html
- Curbera F., Mukhi N., and Weerawarana S., (2001) “On the Emergence of a Web Services Component Model”, Proceedings 6th Int, Workshop on Component- Oriented Programming WCOP2001

- DAML-S Coalition, (2002) "DAML-S: Web Services Description for the Semantic Web" In I. Horrocks and J. Hendler, editors, Proc. First International Semantic Web Conference ISWC 2002,L NCS 2342, Springer-Verlag
- Dogac, A., Laleci, G., Kirbas S., Kabak Y., Sinir S., Yildiz A., Gurcan Y. (2006) "Artemis: Deploying Semantically Enriched Web Services in the Healthcare Domain", Information Systems, Volume 31, Issue: 4-5, Pages: 321-339, Elsevier Science Ltd. Oxford, UK
- Evers M. (2008) "Virtual Worlds and Web 3.0: Examined, Compared, Analyzed", Business Wire Research and Markets Report, Dublin. Source: MarketWatch.com
- Fensel D., Bussler, C., (2002). The Web Service Modeling Framework WSMF. Electronic Commerce Research and Applications, 1(2), 2002.
- Godoy D. and Amandi A., (2006) "Modeling user interests by conceptual clustering", Information Systems**,** Volume 31 , 247 – 265, Issue 4, Special issue: The semantic web and web services, Elsevier Science Ltd. Oxford, UK
- Granlund R., Erik Berglund, Henrik Eriksson, (2000) "Designing web-based simulation for learning", Future Generation Computer Systems, Volume 17, Issue 2, Pages 171-185
- Heflin J.D. (2001) "Towards the semantic web: knowledge representation in a dynamic, distributed environment", Ph.D. Thesis, University of Maryland, College Park
- Hsua C., Wub F. (2006) "Topic-specific crawling on the Web with the measurements of the relevancy context graph", Information Systems, volume 31, 232–246, Elsevier Science Ltd. Oxford, UK
- Jutla D. N., Bodorik P. and Zhang Y. (2006) "PeCAN: An architecture for users' privacy-aware electronic commerce contexts on the semantic web" Information Systems Volume 31, Issues 4-5, July 2006, Pages 295-320, The Semantic Web and Web Services, Elsevier Science Ltd. Oxford, UK
- Longo F., B.M. (2007), "Critical Issues and Potential of SCORM & HLA integration for Supporting Training Simulation Interoperability", Proceedings of EUROSIW2007, Santa Margherita, Italy, June.
- Massei M., Bocca E., B.E. (2006) "Training Simulation as Competitive Advantage for Policy Definition in Logistics Networks", Proceedings of MSO2006, Gaborone, Botswana, September
- Matsuo Y., Mori J., Hamasaki M., (2007) "Polyphonet: An Advanced Social Network Extraction System from the Web" Web Semantics: Science, Services and Agents on the World Wide Web 5, 262–27, Elsevier Science Ltd. Oxford, UK
- Meirina C., Georgiy M. Levchuk, Sui Ruan, Krishna R. Pattipati, Robert L. Popp, (2006) "Normative framework and computational models for simulating and assessing command and control processes" Simulation Modelling Practice and Theory, Volume 14, Issue 4, May, Pages 454-479.
- Orvis K.L., Robert A. Wisher, Curtis J. Bonk, Tatana M. Olson (2002) "Communication patterns during synchronous Web-based military training in problem solving" Computers in Human Behavior, Volume 18, Issue 6, November 2002, Pages 783-795.
- Rao J., Kuengas P. and Matskin M. (2006) "Composition of Semantic Web Services using Linear Logic Theorem Proving", Information Systems, Volume 31, Issue 4, Pages 295-320, The Semantic Web and Web Services, Elsevier Science Ltd. Oxford, UK
- Sanchez E., "Fuzzy logic and the semantic web", Febbraio 2006, Elsevier.
- Shadbolt, N., Hall, W., Berners-Lee, T., (2006). The Semantic Web Revisited. *IEEE Intelligent Systems*, May-June 2006, 96-101.
- Shannon V., (2006). A 'more revolutionary' web. *The New York Times*. Retrived June 4, 2010 from http://www.nytimes.com/2006/05/23/technology/23iht-web.html?_r=1
- Shu-hsien Liao, "Case-based decision support system: Architecture for simulating military command and control" European Journal of operational Research, Volume 123, Issue 3, 16 June 2000, Pages 558-567
- Shu-Hsien Liao, Yen-Ping Ho, "knowledge-based architecture for implementing collaborative problem-solving methods in military e-training" Expert Systems with Applications, Volume 35, Issue 3, October 2008, Pages 976-990.
- Waller J.C., Natalie Foster, (2000) "Training via the web: a virtual instrument" Computers & Education, Volume 35, Issue 2, Pages 161-167.

STEERING PROJECTS THROUGH SIMULATION: AN INNOVATIVE TOOL TO IMPROVE PROJECTS GUIDANCE EFFICIENCY

Massimo de Falco[(a)], Luisa Falivene[(b)], Salvatore Miranda[(c)], Raffaele Iannone[(d)]

Department of Mechanical Engineering
University of Salerno
Fisciano, SA, 84084 - Italy

[(a)]mdefalco@unisa.it, [(b)]lfalivene@unisa.it, [(c)]smiranda@unisa.it, [(d)]riannone@unisa.it

ABSTRACT

A project plan hardly reflects what actually happens during the project mainly because it tends to be static, while nowadays projects are extremely dynamic. Moreover, the strong integration between the executing and the planning phases makes the projects guidance very hard, especially under strict time boundaries. The existing project management techniques are quite inadequate to handle these features. For this reason, the paper proposes an innovative tool able to guarantee an improvement of projects guidance efficiency by introducing the simulation into the time-cost trade-off analysis. The model on the basis of which the tool has been developed uses in an integrated manner different operational software: Microsoft's Project, Visual Basic for Application and Rockwell's Arena. The tool has been tested on a construction project in progress and has already proved its usefulness in the planning as well as in the executing phases.

Keywords: project management, project control process, stochastic network project, uncertainty management

1. INTRODUCTION AND LITERARY REVIEW

The flexibility required to coordinate the multiple feedback processes between the project variables during the execution phase makes nowadays projects highly dynamic and complex. Moreover, the lack of knowledge about the project, especially at the beginning stage (as the *learning curve* teaches), make necessary a systematic management of the project factors and variables since project success principally depends on it, both in terms of reliability of the results and time respect.

Technical, schedule, cost and political changes as well as mistakes that naturally occur during project execution make aleatory the duration of the activities in which the project has been subdivided.

The deep analysis of the existing project planning and control techniques has underlined their inadequacy in managing the present challenges (de Falco et al. 2008). These techniques have not been modified for several years and, therefore, are not able to manage the new critical aspects in actual projects.

Particularly, the deterministic assumptions of the Critical Path Method (CPM) (Kelley and Walker 1959, Kelley 1961) ignore the complexity associated to the uncertainty of the network activities. Except for specific cases, the decisions taken during the project execution on the bases of deterministic analyses do not guarantee high probabilities to complete the project respecting the project plan specifics. This because the conditions in which the project develops are intrinsically probabilistic.

On the contrary, the Program Evaluation and Review Technique (PERT) (Malcom et al. 1959, Elmaghraby 1977), even if based on a probabilistic procedure, is nevertheless limited since it reduces the solution space to a single critical path through the network, ignoring the effects of the complex interactions existing between the dependent sub-path. Except for singular cases, ongoing decisions taken through a PERT analysis do not assure the selection of the best project alternatives (Mummolo 1997).

During the execution phase, project plans have to be periodically re-evaluated and updated as soon new information are available. This creates a complex dynamic probabilistic problem whose final solution is a series of partially implemented plans, each one based on the best available information at the moment of the related evaluation (Conde 2009).

The simplifying hypotheses on which all the probabilistic approaches are based often compromise their reliability degree in representing the real problem. In these cases, turning to a simulative approach may accomplish interesting results for both time and cost management (Dawson 1995, Elkjiaer 2000).

The use of simulation models to represent projects guarantees different advantages:

- Virtual models are explicit and the assumptions are available for the analysis.

- The simulation models are able to unfailingly evaluate the logical consequences of an hypothesis.
- Innumerable factors may be simultaneously correlated.
- Virtual models can be simulated under controlled conditions, and, therefore, allow the analysts to obtain previsions about possible alternative choices.

Moreover, adopting a simulative approach in project management consents to consider different characteristics which can not be differently evaluated, such as: the dependencies between the activities durations; the possibility to decide the best project alternative in function of significant events occurring during project execution; the time-cost links for each activity of the network (Salvendy 2001).

In spite of that, simulation tools have found many applications in project planning and scheduling while their use as support to the control process, excluding some recent cases (Bowman 2006), is still very scarcely diffused (Artto et al. 2001).

The stated observations summarize the motivations of the present research work which proposes a simulative approach to guide projects with more effectiveness.

After a brief description of the methodology, which has been already presented in a previous paper (de Falco and Falivene 2009), the tool which has been ad hoc developed to implement the procedure will be illustrated in detail.

2. THE PROPOSED METHODOLOGY

The innovative methodology to guide projects proposed in the present paper is graphically represented in the figure that follows (Figure 1).

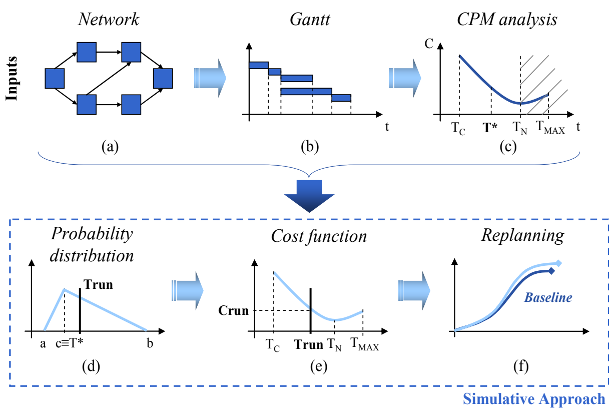


Figure 1: The simulative methodology to guide projects

The approach starts from different inputs coming from the classical project management tools.

Once the network diagram **(a)** and the related Gantt chart **(b)** have been built, the specific durations for each activity of the network (T* in the figure) which minimises the total cost of the whole project **(c)** can be determined through a classical CPM analysis. The network which derives from these durations allows the determination of a project baseline of reference.

At this point the simulative approach can start by introducing a variability to each duration T* through a specific probability distribution in order to consider the natural uncertainty of the activities duration.

At each iteration, for each activity, a duration value is sampled from the probability distribution function **(d)** and the relative cost value is updated **(e)**. On the basis of these values the critical path and the whole project duration can be identified **(f)**. After a sufficient number of repetitions a "baseline pencil", which represents the variation field of the project time-cost binomial, can be obtained.

The baseline pencil enables the determination of a probability distribution for the whole project duration and, therefore, the estimation of the probability of exceeding prefixed contractual due dates.

The proposed approach allows different advantages both in the planning and in the execution phase of a project. Particularly, during the execution phase, the data related to the completely performed activities are considered as deterministic inputs for the simulation model with the consequential reduction of the uncertainty associated to the project duration estimation.

3. THE TOOL

The logical model on the basis on which the tool has been built consists of three characteristic elements (Figure 2): the Input Module, the Simulation Model and the Output Module.

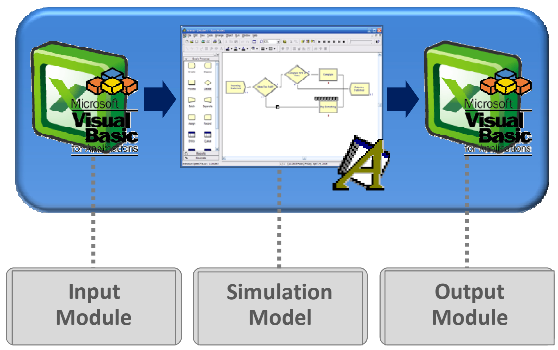


Figure 2: Graphical schematization of the logical model

These elements have been developed through the combination of different management software.

The Input Module allows the link with the project management software (Microsoft's Project) and has been developed through Visual Basic for Application in Excel. It consents to convert the data coming from the project management software into data readable for the Simulation Model.

The Simulation Model has been instead developed through the Rockwell Software's Arena. Particularly, a specific program code developed through Visual Basic for Application allows the automatic transfer of the information coming from the Input Module to the

Simulation Model which automatically builds the network diagram to start the simulation process.

Lastly, the simulation results are transferred to the Output Module which, in its turn, is able to convert in automatic the input data into information comprehensible for the user. The Output Module has been developed through VBA in Excel.

The following sub-paragraphs will illustrate the graphical aspect of the tool as well as its operation logic in implementing the proposed approach to manage projects.

3.1. The Input Module

The Input Module is the element that interacts with the project management software. The information coming from the project plan realized during the planning phase through Microsoft Project, are converted into data understandable for the Simulation Model as the button "Update" is activated.

Figure 3 shows the main table of the Input Module. Each of the column in the Excel sheet is dedicated to one of the project activity. In particular, the data of the project plan are stored into the yellow cells, whereas the white cells contain the values of the characteristic parameters calculated through the Excel sheet.

Particularly, among the data coming from the project plan there are:

- The total number of the activities in which the project has been split during the planning phase, *"N° Activities"*.
- The duration "*T**" identified through the preliminary CPM analysis for each activity of the project.
- The value chosen for the shape parameter k, which allows the definition of the time variation range Δ for each project activity, *"k"* (for the analytical expressions of Δ and *k* see (2) and (3) formulas).
- The characteristic parameters for the definition of the cost function for each of the project activity, "T_L"=Limit or Crash Time, "T_N"=Normal Time, "T_{max}"=Max Time, "C_L"=Limit Cost, "C_N"=Normal Cost, "C_{max}"=Max Cost.

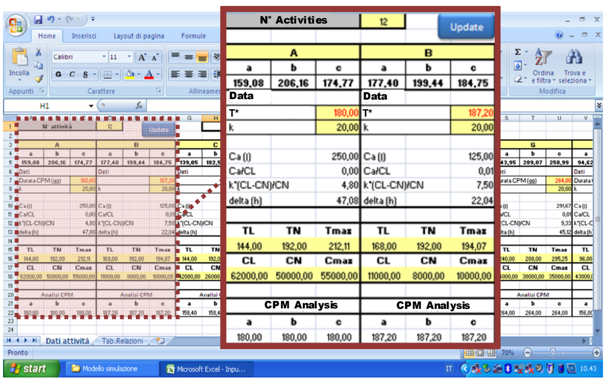


Figure 3: Input Module - Data related to the project activities

The Input Module elaborates these data and evaluates, for each one of the project activities, a series of characteristic parameters necessary to realize the simulative process. Specifically they are:

- The acceleration cost, *"Ca"*, which is automatically calculated by the Input Module through the following formula:

$$C_a = \frac{C_L - C_N}{T_N - T_L} \qquad (1)$$

- The time variation range for each activity duration, *"Δ"*, which has been defined through the following expression:

$$\Delta = \frac{T_N - T_L}{e^{k \cdot \frac{C_a}{C_L}}} \qquad (2)$$

in which:

$$k = k_1 \cdot \frac{C_L - C_N}{C_N} \qquad (3)$$

where k_1 is a positive constant and $(C_L-C_N)/C_N$ is the cost proportional increase of a generic project activity (de Falco and Falivene 2009).

- The characteristic parameters for the definition of the triangular probability distribution, *"a"*, *"b"* and *"c"* (see Figure 1d) which are automatically calculated by the Input Module, on the basis of specific hypothesis, through the formulas:

$$\begin{cases} a = c - \Delta / 3 \\ b = c + 2/3\Delta \end{cases} \qquad (4)$$

The condition (4) reflects the choice to consider for each activity the pessimistic event more likely than the optimistic event with a two to one ratio (de Falco and Falivene 2009).

Besides the mentioned data, the project management software has to transmit to the Input Module the information related to the logical links between the network activities in order to realize the subsequent simulation. Figure 4 shows how the Input Module presents the sheet in which the table summarizing the relations between the activities is reported.

These data in their original format cannot be read by the Simulation Model and, for this reason, have to be converted into information compatible with this model. To this aim, by activating the "Update" button (on the left side in the figure), the Input Module is able to transform the "relations table" into a "relations matrix". The column on the right of the matrix reports the number of predecessors for each activity, and the row under the matrix reports the number of activities which succeed each network activity. The delays of each network activity (column) in relation to the specific preceding activities (row) are reported in the matrix, if they are present.

At this point, all the information necessary to start the simulative process are available.

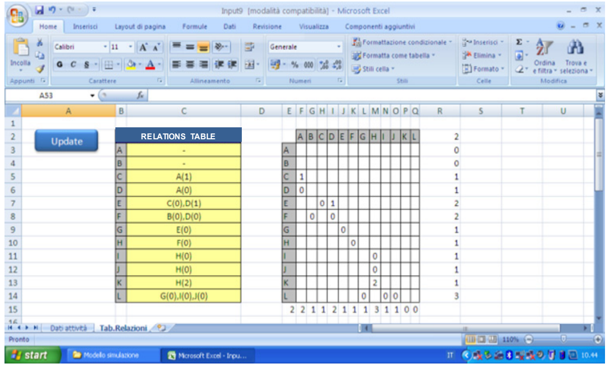


Figure 4: Conversion of the “relations table” into the “relations matrix”

3.2. The Simulation Model

The figure that follows (Figure 5) shows the main window of the Simulation Model.

In particular, the figure illustrates the screenshot of the tool in execution. The user can interface with a control panel that presents specific fields in which it is possible to insert the information needed to start the simulation.

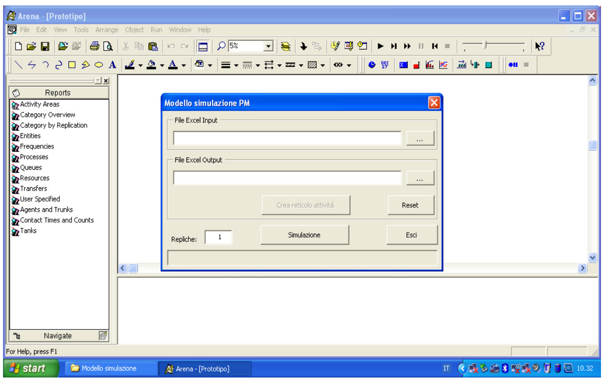


Figure 5: Main window of the tool

The first field, *“File Excel Input”*, allows the loading of the Excel file in which the input data are contained, that is the above described Input Module (Figure 6).

The *“File Excel Output”* field, in its turn, permits the loading of the output file, the Output Module, on which the results of the simulation will be written at the end of the run (Figure 7).

Once the two fields have been filled in, the activation of the button *“Create activity network”* allows the Simulation Model to automatically create the network related to the project (Figure 8).

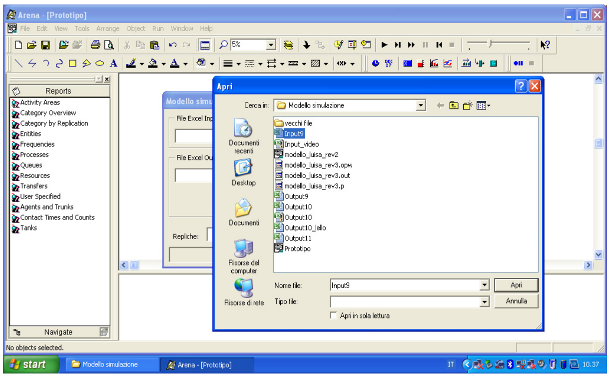


Figure 6: Loading of the input data into the Simulation Model

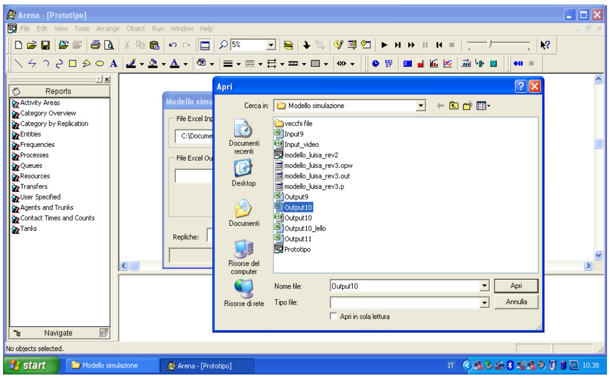


Figure 7: Loading of the file where the simulation results will be written

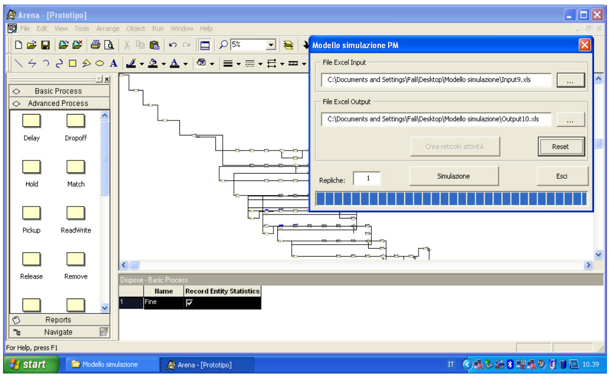


Figure 8: Creation of the project network through the Simulation Model

Particularly, the flow chart of Figure 9 illustrates the logical scheme that the Simulation Model follows in order to build the project network through the Arena Software. The figure shows the different Arena blocks used to represent the initial and the final nodes of the network, which are built once, as well as the project activities.

Moreover, the specific parameters coming from the Input Module (through the Readwrite 1 and Readwrite 2 blocks), generated (through the Delay, the Process and the Assign blocks) or transferred to the Output Module from the Simulation Model (through the Readwrite 3 block), are indicated under each Arena element.

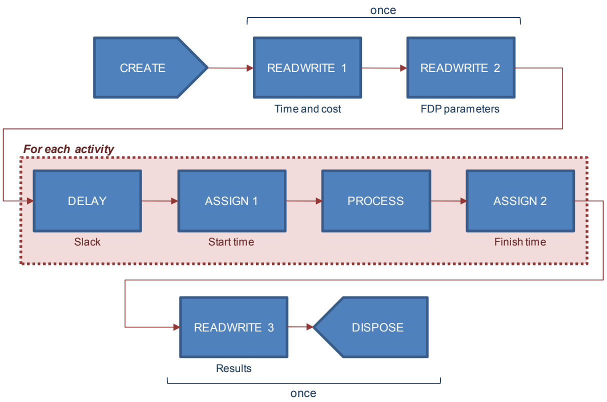


Figure 9: Logical scheme for the construction of the project network

At this point, the user has to choose and consequently write into the field *"Replications"* the number of simulation replications that wants to perform and, therefore, activate the button *"Simulation"* in order to start the simulation runs (Figure 10).

In the window of the figure it is possible to note the presence of a blue bar which allows the user to control the time advancement of the simulation in order to guarantee him a precise idea of the time to the completion of the process.

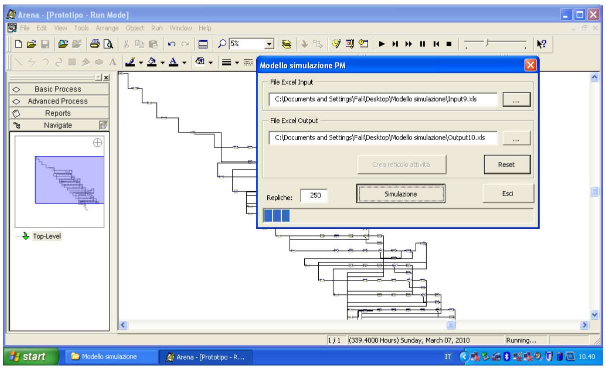


Figure 10: Choice of the number of replications and starting of the simulation process

3.3. The output module

As previously mentioned, the results of the simulation are automatically transferred to the Output Module, which has been developed through VBA under Excel.

The data coming from the Simulation Model are stored into an Excel sheet in the form of a matrix which columns represent the project activities and which rows report the data of each simulation run (Figure 11). Particularly, for each activity, the start times (*T_start*), the finish times (*T_finish*) and the costs (*Cost*) associated to the estimated duration at each iteration are loaded in the Output Module.

These data are reorganized by the Output Module and, for each activity and for each simulation replication, the unit cost is calculated (Figure 12).

For each simulation run, the unit cost for an activity has been determined by spreading on the time range (*T_finish-T_start*) the total cost associated to the duration estimated through the Simulation Model.

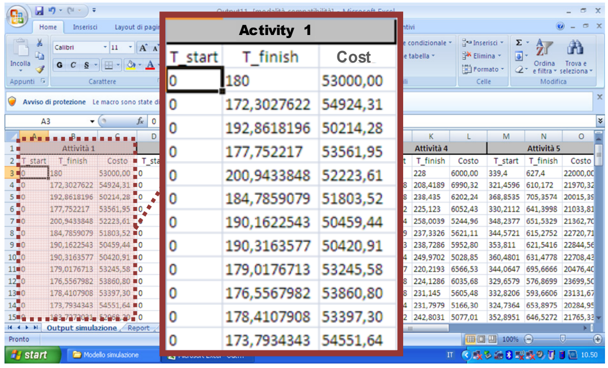


Figure 11: Simulation results in the Output Module

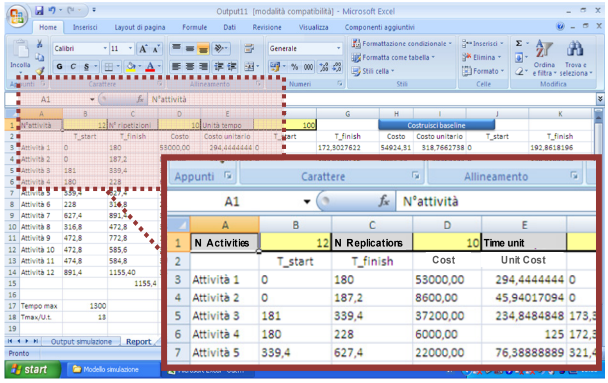


Figure 12: Estimation of the unit costs for the project activities

Also in this case, the data coming from the Simulation Model must be converted into data comprehensible for the users. For this reason, the Output Module provides the possibility to manage the available data in order to construct a Gantt diagram for each simulation replication realized (Figure 13).

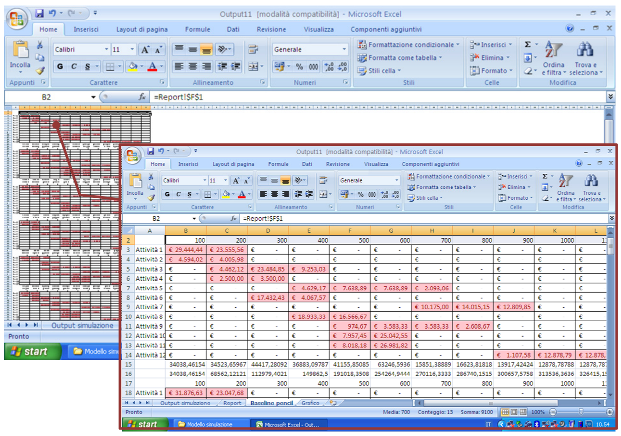


Figure 13: The Gantt diagrams related to the different simulation replications

The Figure 13 shows the different Gantt diagrams in the Excel sheet and a zoom on one of them in order to make their peculiar characteristics more evident. As it is possible to see in the figure, under each Gantt chart there are two rows which quote respectively the total costs and the cumulated costs of the project.

Through the cumulated costs, by activating the button *"Create Baseline"*, the Output Module is able to graphically represent a *"baseline pencil"* (de Falco and Falivene 2009) which portrays the variation field of the project time-cost binomial (Figure 14) and a related probability distribution function of the project completion time.

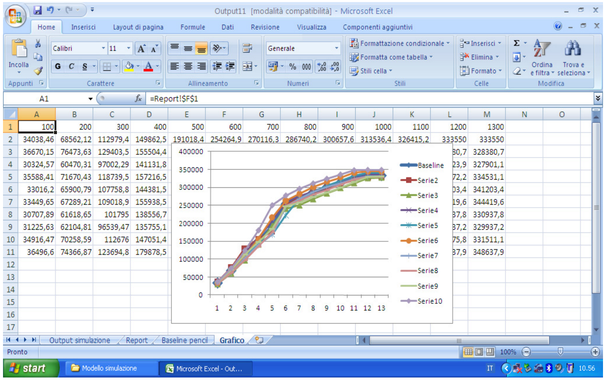


Figure 14: The baseline pencil

4. CONCLUSIONS AND FUTURE RESEARCH

The proposed approach to manage projects is a sort of probabilistic CPM that, by introducing simulation, consents the introduction into the classical analysis of the intrinsic aleatory of the network activities durations. This allows the increasing of the efficiency both of the planning phase and the execution phase thanks to the possibility of updating in real-time the data.

Particularly, during the project planning phase turning to the present methodology consents a series of interesting advantages:

- the generation of the baseline pencil and the related probability distribution function guarantees more confident estimations;
- the use of simulation allows managers an higher consciousness when submitting proposals or negotiating contracts with the clients;
- the approach determines a stronger consciousness of both real and perceived potentials of the chosen project proposal.

At the same time, the use of the methodology during the execution phase consents a more efficient control process since:

- it allows the real-time updating of the data after the control;
- it allows the identification, the testing and the evaluation of potential improvement strategies with project is in progress.

The ad hoc developed tool, thanks to its particular characteristics, gives users the effective total visibility of the project information heritage during its execution.

The real-time monitoring of the tendencies of particular relevant phenomena allows the efficient representation of the project dynamics and the production of interesting indicators for the ongoing monitoring, ex-post evaluations and impact evaluations.

Moreover, the goodness of this tool is more evident since it guarantees managers the availability of always updating data and the possibility of realizing continuous analyses on the tendency of relevant phenomena, such as analyses which give a complete and timely informative vision useful to make decisive choices and consequently evaluate their effects.

The simulative approach together with the developed tool are characterized by a great flexibility since the Input Module, the Simulation Model and the Output Module can be fitted on all kind of project.

The advantages of the proposed tool with respect to apparently similar systems such as spreadsheet model supported by simulation packages (i.e. Crystal Ball) can be resumed in the two following points:

1. easy management of the precedence rules among the activities;
2. quick realization of the simulation model reproducing the project activity network.

The simulation model is in fact automatically created by the tool, completely avoiding in this way the complex and time consuming phases of design, implementation and validation of the model. In addition, any variation on the activity durations is automatically managed, with the shifting of all the connected activities, making an effective real-time control of the project advancing possible.

The strength points evidenced by the proposed project management approach are several but the potential interesting extensions of the research are also numerous.

First of all, the cost functions associated to each project activity have been considered deterministic but in practice it is likely that the cost value related to a particular activity duration is considered variable according to a specific probability distribution, as for the time.

Finally, the tool has been tested on a construction project in progress and has already shown its potentialities but it would be opportune to follow a project from the drafting of the project plan to its completion in order to demonstrate its real degree of effectiveness in leading to a higher performing project management.

ACKNOWLEDGMENTS

Due to the recent cuts in research funds by the Italian Government, the Authors will not be able to continue the work presented in this paper; they will therefore gladly provide advice and all the necessary details to foreign institutions and private companies interested in pursuing this line of research.

REFERENCES

Artto, K.A., Lehtonen, J.M., Saranen, J., 2001. Managing projects front-end: incorporating a strategic early view to project management with simulation. *International Journal of Project Management*, 19, 255-264.

Bowman, R.A., 2006. Developing activity duration specifications limits for effective project control. *European Journal of Operational Research*, 174, 1194-1204.

Conde, E., 2009. A minmax regret approach to the critical path method with task interval times. *European Journal of Operational Research*, 197, 235-242.

Dawson, C.W., 1995. A dynamic sampling technique for the simulation of probabilistic and generalized activity networks. *Omega Int. J. Mgmt Sci.*, 23(5), 557-566.

de Falco, M., Falivene, L., Eldomiaty, T., 2008. Controlling fast-forwarding projects: criticalities and innovative decision-making perspectives. *Proceedings of the 22th IPMA World Congress*, 9-11 November, Rome (Italy).

de Falco, M., Falivene, L., 2009. Improving project control and guidance through a time-frame simulation approach. *Proceedings of the Advances in Production Management Systems (APMS) Conference*, 19-23 September, Bordeaux (France).

Elkjiaer, M., 2000. Stochastic Budget Simulation. *International Journal of Project Management*, 18, 139-147

Elmaghraby S.E., 1977. *Activity Networks: Project Planning and Control by Network Models*, John Wiley and Sons, New York.

Kelley, Jr. J.E., Walker, M.R., 1959. Critical-path planning and scheduling: An introduction. *Proceedings of the Eastern Joint Computer Conference*, pp. 160-173.

Kelley J.E., 1961. Critical path planning and scheduling–mathematical basis. *Operations Research*, 9, 296–320.

Malcom, D.G., Roseboom, J.H., Clark, C.E., Fazar, W., 1959. An application of a technique for research and development program evaluation. *Operations Research*, 7, 646-669.

Mummolo, G., 1997. Measuring uncertainty and criticality in network planning by PERT-path technique. *International Journal of Project Management*, 15, 377-387.

Salvendy, G., 2001. *Handbook of Industrial Engineering* – Third Edition. Wiley-IEEE

AUTHORS BIOGRAPHY

Massimo de Falco graduated cum laude in Aeronautic Engineering in 1991 and in 1996 he obtained his PhD in Operations Management at the Engineering Faculty of the University of Naples. In 1994 he was Visiting researcher at MIT, member of Sloan's Operations Management Group. In 1998 he became Associate Professor of Operations Management. He is lecturer in the MBA programs at LUISS University (Roma) and Polytechnic of Milan.

Luisa Falivene graduated cum laude in Mechanical Engineering in 2006. In 2010 she obtained her PhD degree at the University of Salerno (Italy) in Engineering and Economics of Innovation. Her research concentrates on the field of Project Management and Control.

Raffaele Iannone graduated cum laude in Mechanical Engineering in 2000. In 2005, he obtained his PhD degree in Engineering of Advanced Productive Systems. In 2005, he became assistant professor in operations management at the University of Salerno. His main research interests are in Supply Chain Management, Distributed Simulation and ERP Systems.

Salvatore Miranda graduated cum laude in Mechanical Engineering in 1997 and received his PhD degree in Industrial Plants from the University of Naples "Federico II" in 2002. From November 2002 he has been working at the University of Salerno, Italy as assistant professor. His research interests include production planning and control, supply chain simulation and automated production systems.

CONTROL STRATEGIES FOR A VARIABLE SPEED URBAN WIND TURBINE

A.Naamane, A. Mehdary, M.Trihi*

Laboratoire des Sciences de l'Information et des Systèmes (LSIS)
UMR6168
Avenue escadrille normandie niemen
13397 Marseille cedex 20
*Faculté des sciences aïn chock Casablanca Maroc

Email,: aziz.naamane@lsis.org; adil.mehdary@lsis.org; morad.trihi@fsac.ma

ABSTRACT

Different electric wind power conversion systems structures can be used based on the converters topologies. The main objective in all the structures is always the same: the wind energy at varying wind velocities has to be converted to electric power with the highest performances. In this paper, a double fed induction generator is used with different control strategies: a direct control without taking into account the coupling between the present currents and an indirect control based on state space approach. A comparison in term of robustness and stability between the two methods is presented through simulation results.

Keywords: wind-turbine, double fed induction generator, regulators, Park transform,

1. INTRODUCTION

Today, it is proved that fossil fuels continue to diminish and climate change poses an ever-increasing threat. Among the alternative power sources, wind power "is one of the most promising new energy sources" .Wind is appealing for several reasons. It is abundant, cheap, inexhaustible, widely distributed, clean, and climate-benign, a set of attributes that no other energy source can match. However, there are still many unsolved challenges in expanding wind power. In this paper, two approaches for control of a double fed induction generator (DFIG) are introduced.

Wind turbines use a doubly-fed induction generator (DFIG) consisting of a wound rotor induction generator and an AC/DC/AC IGBT-based PWM converter. The stator winding is connected directly to the 50 Hz grid while the rotor is fed at variable frequency through the AC/DC/AC converter.(Fig.1). The DFIG technology allows extracting maximum energy from the wind for low wind speeds by optimizing the turbine speed, while minimizing mechanical stresses on the turbine during gusts of wind. The optimum turbine speed producing maximum mechanical energy for a given wind speed is proportional to the wind speed. Another advantage of the DFIG technology is the ability for power electronic converters to generate or absorb reactive power, thus eliminating the need for installing capacitor banks as in the case of squirrel-cage induction generator [11].

This paper is organized as follows. Section II presents the modeling of the studied system. In Section III the two control approaches for variable speed wind turbine are presented with the simulation results. Finally, we close the paper by discussing ongoing and future challenges in the last Section.

2. MODELING SYSTEM

2.1. System representation

Figure 1 shows our wind conversion chain, which consists of a turbine, a generator (double fed induction generator) with a rotor winding, converters and PWM (Figure 1).

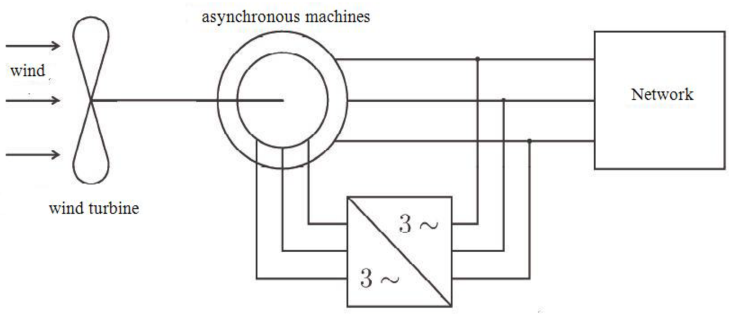


Figure1 : Doubly fed induction generator with converters

2.2. DFIG representation in Park reference

For the different equations used in this paper the following nomenclature will used:

P_m : Mechanical power captured by the wind turbine and transmitted to the rotor

P_s : Stator electrical power output

P_r: Rotor electrical power output
P_{gc}: C grid electrical power output
Q_s : Stator reactive power output
Q_r : Rotor reactive power output
Q_{gc}: C grid reactive power output
T_m: Mechanical torque applied to rotor
T_{em}: Electromagnetic torque applied to the rotor by the generator
ω_r : Rotational speed of rotors
V_s, V_r are Staor and Rotor voltages
V_{qs} , V_{ds} are the three-Phase supply voltages in d-q reference frame, respectively
I_{qs} , I_{ds} are the three-Phase stator currents in d-q reference frame, respectively
V_{qr} , V_{dr} are the three-Phase rotor voltages in d-q reference frame, respectively
I_{qr} , I_{dr} are the three-Phase rotor voltages in d-q reference frame, respectively
R_s, R_r are the stator and rotor resistances of machine per phase, respectively
φ_{qs},φ_{ds} are the three-Phase stator flux linkages in d-q reference frame, respectively
φ_{qr}, φ_{dr} are the three-Phase rotor voltages in d-q reference frame, respectively
L_s , L_r are the leakage inductances of stator and rotor windings, respectively
θ_s, θ_r are the stator and rotor flux angle, respectively

To model the DFIG, we establish the model of a wounded rotor induction machine, which will be obtained in the same manner as the model of squirrel-cage induction generator, except that the rotor voltages are not null.

General equations of synchronous wounded rotor machine are the following [6]:

$$[V_s]_3 = R_s[I_s]_3 + \frac{[d\varphi_s]_3}{dt} \qquad (1)$$
$$[V_r]_3 = R_r[I_r]_3 + \frac{[d\varphi_r]_3}{dt}$$

Using the transformed Park on flux and voltage equations of the DFIG, the model is presented in a two-phase rotating frame, with the following equations: [6] [8].

$$\begin{cases} V_{ds} = R_s I_{ds} + \frac{d\varphi_{ds}}{dt} - \dot{\theta}_s\varphi_{qs} \\ V_{qs} = R_s I_{qs} + \frac{d\varphi_{qs}}{dt} + \dot{\theta}_s\varphi_{ds} \\ V_{dr} = R_r I_{dr} + \frac{d\varphi_{dr}}{dt} - \dot{\theta}_s\varphi_{qr} \\ V_{qr} = R_r I_{qr} + \frac{d\varphi_{qr}}{dt} + \dot{\theta}_r\varphi_{dr} \end{cases} \qquad (2)$$

$$\begin{cases} \varphi_{ds} = L_s I_{ds} + M I_{dr} \\ \varphi_{qs} = L_s I_{qs} + M I_{qr} \\ \varphi_{dr} = L_r I_{dr} + M I_{ds} \\ \varphi_{qr} = L_r I_{qr} + M I_{qs} \end{cases} \qquad (3)$$

The electromagnetic torque is expressed by:

$$\Gamma_{em} = \frac{M}{L_s}(I_{qr}\varphi_{ds} - I_{dr}\varphi_{qs}) \qquad (4)$$

p is the number of pole of the DFIG.

By orienting the reference (d, q) so that the axis is aligned with the stator flux [8] [9], we obtain:

$$\varphi_{ds} = \varphi_s \quad \text{and} \quad \varphi_{qs} = 0 \qquad (5)$$

and the electromagnetic torque expression is:

$$\Gamma_{em} = \frac{M}{L_s} I_{qr}\varphi_s \qquad (6)$$

And the flux equation is:

$$\varphi_s = L_s I_{ds} + M I_{dr} \qquad (7)$$

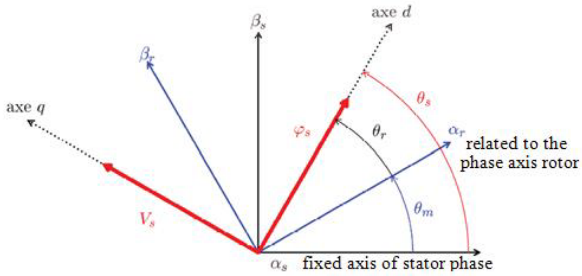


Figure 2: Orientation of the axis of the stator flux

Assuming that the grid voltage remains stable, with the value V_s, and as the flux is considered constant, the electromagnetic torque is proportional to the rotor current I_{qr} according to equation (6).

And by neglecting the stator windings resistance (in the case of for high power generators) the stator voltages equations become

$$V_{ds} = \frac{d\varphi_s}{dt} \qquad (8)$$
$$V_{qs} = \dot{\theta}_s\varphi_s$$

note $\dot{\theta}_s = \omega_s$ and $V_{qs} = V_s$

This gives the following expressions:

$$V_{ds} = 0 \qquad (9)$$
$$V_s = \omega_s\varphi_s$$

The relation between the stator and rotor currents is set from the equation (7)

$$I_{ds} = -\frac{M}{L_s} I_{dr} + \frac{\varphi_s}{L_s} \quad (10)$$

$$I_{qs} = -\frac{M}{L_s} I_{qr}$$

The stator active and reactive powers can be written as:

$$P = V_{ds} I_{ds} + V_{qs} I_{qs} \quad (11)$$
$$Q = V_{qs} I_{ds} - V_{ds} I_{qs}$$

The previous expressions are transformed to the following ones by using equation (9)

$$P = V_s I_{qs} \quad (12)$$

$$Q = V_s I_{ds}$$

and then by using equation (10) the powers are expressed as follows :

$$P = -V_s \frac{M}{L_s} I_{qr} \quad (13)$$
$$Q = -V_s \frac{M}{L_s} I_{dr} + \frac{\varphi_s V_s}{L_s}$$

with $V_s = \omega_s \varphi_s$

so

$$P = -V_s \frac{M}{L_s} I_{qr} \quad (14)$$
$$Q = -V_s \frac{M}{L_s} I_{dr} + \frac{V_s^2}{L_s \omega_s}$$

If the magnetizing inductance M is constant, we note that the active power P is proportional to the quadrature current and reactive power is mainly proportional to the rotor current (4).

To properly control the machine, we will set the relationship between the rotor currents and voltages applied to the machine.

Substituting in equation (3) currents by their value in equations (10) we obtain

$$\varphi_{dr} = \left(L_{r-} \frac{M^2}{L_s}\right) I_{dr} + \frac{V_s^2}{L_s \omega_s} \quad (15)$$

$$\varphi_{qr} = \left(L_{r-} \frac{M^2}{L_s}\right) I_{qr}$$

and replacing the flux in the relation (2) we obtain

$$V_{dr} = R_r I_{dr} + \left(L_{r-} \frac{M^2}{L_s}\right) \frac{dI_{dr}}{dt} - g\omega_s \left(L_{r-} \frac{M^2}{L_s}\right) I_{qr}$$

(17)

$$V_{qr} = R_r I_{dr} + \left(L_{r-} \frac{M^2}{L_s}\right) \frac{dI_{qr}}{dt} \quad g\omega_s \left(L_{r-} \frac{M^2}{L_s}\right) I_{qr} + g \frac{V_s^2}{L_s \omega_s}$$

Where g is the slip of the induction machine and

$$\omega_r = g\omega_s$$

In steady state operation the voltage expressions are:

$$V_{dr} = R_r I_{dr} - g\omega_s \left(L_{r-} \frac{M^2}{L_s}\right) I_{qr} \quad (16)$$
$$V_{qr} = R_r I_{dr} + g\omega_s \left(L_{r-} \frac{M^2}{L_s}\right) I_{dr} + g \frac{V_s^2}{L_s \omega_s}$$

It is to be noted from these expressions, one can generate two control strategies. As the powers and voltages are linked by a first order, and in the case where g could be neglected, one can control separately the active and reactive powers. Figure 3 represents a general case where the output depends only on the input without other blocks. So in the direct method we transform figure 4, to match the figure 3.

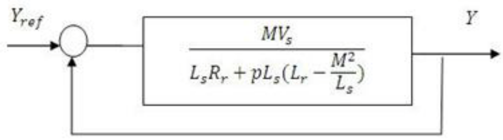


Figure 3: Block diagram of a decoupled relation

The following block diagram represent the different blocks showing the active and reactive powers as outputs and rotor voltages as inputs.

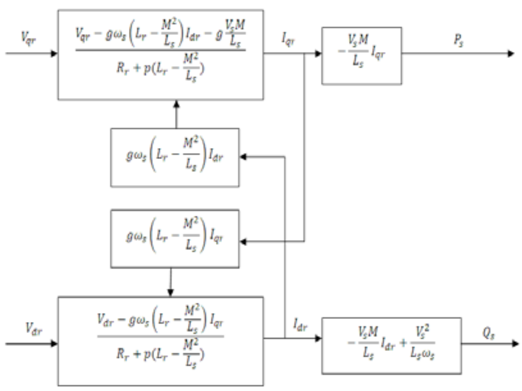


Figure 4: Block diagram of the DFIG taking into account the coupling

3. CONTROL STRATEGIES

In order to decouple the axes d and q, two terms ε_q et ε_d which represent coupling residues and perturbations are introduced. We assume that the term $L_r - \frac{M^2}{L_s}$ could be neglected if we add the constant term ν ($\nu = \frac{V_s^2}{L_s \omega_s}$), figure 5 shows the system after the assumptions considered above [9].

Regulators synthesis:

Both axes can be controlled separately; regulators R_q et R_d inputs are P_{ref} - Q_{ref}, and the diphased voltages as outputs. These latter are applied to the generator (DFIG) to allow the wind turbine to rotate at its maximum speed, so a Proportional & Integral Controllers are used. These PI regulators: U(p)/E(s)=$k_p + \frac{k_i}{p}$, followed as it is shown in figure 5 by the following transfer function :

$$H = \frac{MV_S}{L_S R_r + pL_S(L_r - \frac{M^2}{L_S})} \quad (17)$$

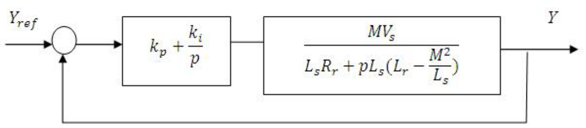


Figure 5: Block diagram for decoupled axes

Determination of parameters k_p et k_i *:*

The open loop transfer function is:

$$\boldsymbol{H} = (k_p + \frac{k_i}{p}) * \frac{MV_S}{L_S R_r + pL_S(L_r - \frac{M^2}{L_S})} \quad (18)$$

H can also be written as:

$$\boldsymbol{H} = \frac{(p + \frac{k_i}{k_p})}{\frac{p}{k_p}} * \frac{\frac{MV_S}{L_S\left(L_r - \frac{M^2}{L_S}\right)}}{p + \frac{L_S R_r}{L_S\left(L_r - \frac{M^2}{L_S}\right)}} \quad (19)$$

To eliminate the zeros in this transfer function we use poles compensation. This method is not the only one for the synthesis of a PI controller but it has the advantage to be rapidly implemented in the case of a first order transfer function [8] [9].

$$\frac{k_i}{k_p} = \frac{L_S R_r}{L_S\left(L_r - \frac{M^2}{L_S}\right)} \quad (20)$$

The transfer function becomes:

$$\boldsymbol{H} = \frac{k_p \frac{MV_S}{L_S\left(L_r - \frac{M^2}{L_S}\right)}}{p} \quad (21)$$

The closed loop function is therefore expressed as:

FTBF= $\frac{1}{1 + p\tau}$ with

$$\tau = \frac{1}{k_p} \frac{L_S\left(L_r - \frac{M^2}{L_S}\right)}{MV_S} \quad (22)$$

so

$$k_p = \frac{1}{\tau} \frac{L_S\left(L_r - \frac{M^2}{L_S}\right)}{MV_S} \text{ and } k_i = \frac{1}{\tau} \frac{L_S R_r}{MV_S} \quad (23)$$

The term τ is the system response time.

3.1. Indirect control

According to the equations (17), we represent then the system in the following form: [10]:

$$\dot{X} = AX + BU \quad (24)$$
$$Y = CX + DU$$

$$x = \begin{pmatrix} I_{qr} \\ I_{dr} \end{pmatrix}$$

and the rotor voltages are the inputs

$$. \, e = \begin{pmatrix} Vqr \\ Vdr \end{pmatrix}$$

The state matrix A is as

$$\begin{pmatrix} -\frac{R_r}{\left(L_r - \frac{M^2}{L_S}\right)} & -\frac{g\omega_s\left(L_r - \frac{M^2}{L_S}\right)}{\left(L_r - \frac{M^2}{L_S}\right)} \\ \frac{g\omega_s\left(L_r - \frac{M^2}{L_S}\right)}{\left(L_r - \frac{M^2}{L_S}\right)} & -\frac{R_r}{\left(L_r - \frac{M^2}{L_S}\right)} \end{pmatrix} \quad (25)$$

The control matrix B is in the form

$$\frac{1}{\left(L_r - \frac{M^2}{L_S}\right)} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (26)$$

The observation matrix C is the identity matrix ($Y = C * X$), the input-output coupling matrix is zero.

Regulators synthesis:

The goal is to determine a command which allows the poles of the closed loop transfer function to be properly placed [10] according to the desired specifications. The poles of the transfer function are the eigenvalues of the state matrix; the goal is to set a linear state feedback control as it is shown in figure 6. The matrix K is established so that coupling between the currents is neutralized.

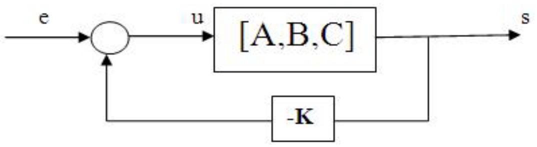


Figure 6: control state feedback

With K= $\begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix}$

The state equation is thus written in the following form:

$$\frac{dx}{dt} = Ax + B(e - Kx) = (A - BK)x + Be \qquad (27)$$

System dynamics is described by the new transfer matrix *L=A-BK*, its eigenvalues are chosen according to the performances demand. So by using the characteristic polynomial P(λ) and resolving the equations system, the matrix K is determined.

$$P(\lambda) = det(\lambda I - A - BK) \qquad (28)$$

$$K = \begin{pmatrix} 1 - R_r & g\omega_s\left(L_r - \frac{M^2}{L_s}\right) \\ g\omega_s\left(L_r - \frac{M^2}{L_s}\right) & 1 - R_r \end{pmatrix} \qquad (29)$$

The resulted control structure is presented in figure 7. The currents I_{qr} and I_{dr} are decoupled and defined using two PI power regulators..

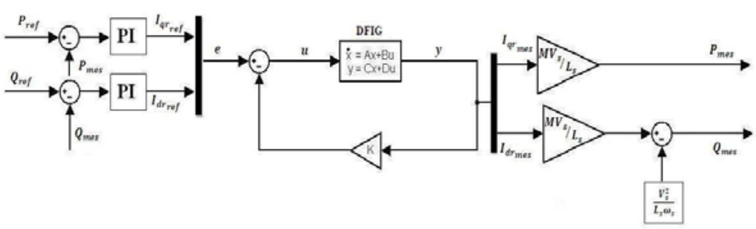


Figure 7: Indirect control structure

3.2. Performance analysis

This section gives the comparison simulation results between the two control strategies in term of robustness and performances. For these simulations a Double Fed Induction Generator of 1.5 KW is considered.

a- Following up the control signal:

In this first test step variations of active and reactive powers are applied while the machine is driven at fixed speed.

Test conditions:

Machine trained at 1450 rpm

at t = 3s: P_{ref} goes from 0 to -2000W
at t = 2s: Q_{ref} goes from 0 to 1000 VAR

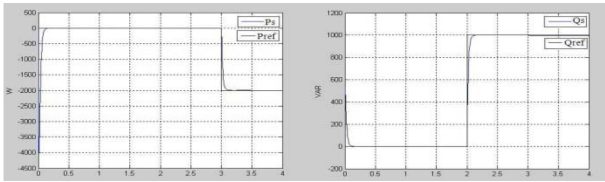


Figure 8: reference tracking active power and reactive (direct control)

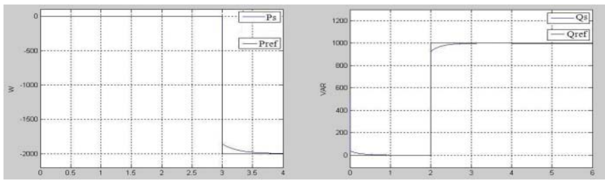


Figure 9: reference tracking of active and reactive power (indirect control)

b- Sensitivity to disturbances:

This test enables us to verify the behavior of output powers when the rotational speed of the machine varies abruptly.
Test conditions:
machine being driven at 1350 rpm
active power setpoint fixed at -1000 W
fixed reactive power setpoint fixed at 1000 VAR
at t = 2s speed changes suddenly from 1350 to 1450 rpm.

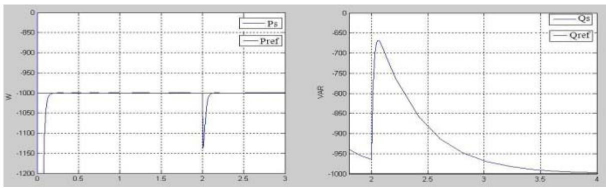


Figure 10: Effect on powers of a sudden change in speed (direct control)

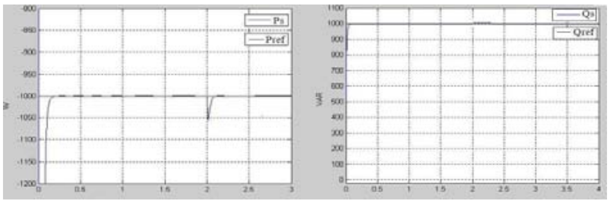


Figure 11: Effect on the powers of a sudden change in speed (indirect control)

c- Robustness:

To test robustness, we vary the model parameters of the used DFIG. In fact in the calculations of regulators constants, the different parameters are assumed fixed. Or in a real system, these parameters can vary due to different physical phenomena (inductor saturation resistance heat etc…), and also due to measurements inaccuracies.

Test conditions:

t = 2s: resistors R_s and R_r multiplied by 2
the inductance L_s, L_r and M divided by 2
machine being driven at 1350 rpm.
at t = 3.5 s: Q_{ref} goes from 0 to -1000 VAR
at t = 3.5 s: P_{ref} goes from 0 to 1000W

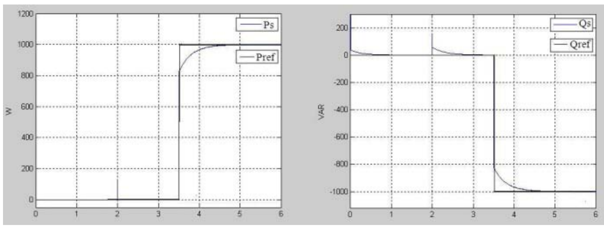


Figure 12: Effect of DFIG parameter variations on P_s and Q_s (indirect control)

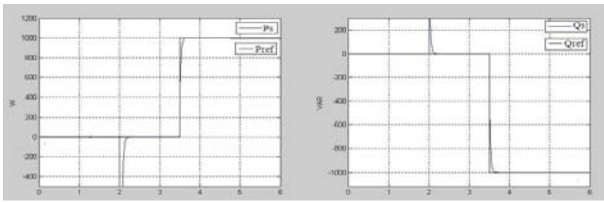


Figure 13: Effect of DFIG parameter variations on P_s and Q_s (direct control)

It is to be noted that rotation speed variation and the generator parameters change have a weak impact on indirect control method. This control development has enabled us to highlight several interesting aspects for further study on the whole power wind production. It is obvious that direct method is easier to implement than the indirect one. According to the simulation results the indirect method is more robust and less sensitive to perturbations.

CONCLUSION

The purpose of this work is to investigate control approaches of a Double Fed Induction Generator: a direct control without taking into account the coupling between the present currents and an indirect control based on state space approach. A comparison in term of robustness and stability between the two methods is presented through simulation results.

As the whole presented work is based upon simulation, we are currently working on an experimental bench [12] in order to validate experimentally the obtained simulation results. As a future work, we intend to propose a real-time implementation of the proposed control using the system model and Real-Time Workshop using dSPACE hardware.

REFERENCES

[1] A.MIRECKI, "*Etude comparative de chaînes de conversion d'énergie dédiées à une éolienne de petite puissance*", thèse 2005

[2] B.Beltran, " *Maximisation de la puissance produite par une génératrice asynchrone Double alimentation d'une éolienne par mode glissant d'ordre supérieur*", JCGE'08 Lyon 17&18/12/08

[3] B.Boukhezzar, "*Sur les stratégies de commande pour l'optimisation et la régulation de puissance des éoliennes a vitesse variable*", thèse 2006

[4] Shashikanth Suryanarayanan and Amit Dixit, "*Control of Large Wind Turbines: Review and Suggested Approach to Multivariable Design*"

[6] H. Razik, "*Modélisation de la machine asynchrone à cage. Application à la simulation de moteurs à cage défaillante. Spectres des courants des moteurs défaillants*", Université Henry Poincaré

[7] L.C.Henriksen, "*Model Predictive Control of a Wind Turbine*", Technical University of Denmark 2007

[8]A.Boyette, "*Contrôle-commande d'un générateur asynchrone a double alimentation avec système de stockage pour la production éolienne*", thèse 2006

[9] Fréderic Poitiers, "*Etude et commande de génératrice asynchrones pour l'utilisation de l'énergie éolienne*", école polytechnique de l'université de Nantes 2003

[10] Philipe de Larminat, *"Automatique: commande des systems lineaire"* Edition Hermes 1996

[11] Holdsworth L, Wu XG, Ekanayake JB, Jenkins N. *Comparison of fixed speed and doubly-fed induction wind turbines during power system disturbances*. IEE Proceedings: Generation, Transmission, Distribution, 2003, 3: 343-352.

[12] A.Naamane, N.K M'sirdi "MACSyME: Modelling, Analysis and Control for Systems with Multiple Energy sources" , *in : Springer , "Book Proceeding of SEB 2009", Springer Berlin Heidelberg, Vol. ISBN 978-3-642-03453-4 , Ch. Sustainability in Energy and Buildings,http://www.springerlink.com/content/v23046h301v7015n/* Mai 2009 .

A PAIRWISE COMPARISON-BASED MODEL OF UNCERTAINTY IN MULTI-PERSON MULTI-CRITERIA DECISION PROBLEMS

Nils Burgmann[a], Jana Görs[b], Graham Horton[a]

[a] Otto-von-Guericke-University, Magdeburg, Germany
[b] Zephram GbR, Magdeburg, Germany

Corresponding author email: jana.goers@zephram.de

1. SUMMARY

This paper describes a simple model for the local uncertainty in a multi-person multi-criteria decision problem (MPMCDP). The model is motivated by the authors' experiences in the first stage of corporate innovation processes, which are characterized by a very large number of ideas, for which little or no clarifying information is available. Both the uncertainty model and the ranking algorithm are based on pairwise comparisons of alternatives in order to minimise the costs of processing ideas and to improve the reliability of the results. The model allows uncertainty to be detected cheaply and suggests an efficient method for its reduction.

2. INTRODUCTION

Innovation is a key factor in establishing and maintaining competitiveness. Many companies use innovation processes as a structured approach to obtaining new products and services. This process begins with the generation of a pool of ideas, may include activities such as evaluation, research, development and prototyping, and culminates in the launch of a new product. The process is often formulated using a Stage-Gate paradigm (Cooper, Edgett and Kleinschmidt 2001) (Belliveau, Griffin and Somermeyer 2002). In the first stage of this process, ideas for new or improved products are generated, and at the subsequent first gate, these ideas are evaluated, and those with the highest potential are selected to be investigated further.

Typical characteristics of the first gate of an innovation process are that a very large number of ideas can be involved, several decision-makers participate in the selection, and there are many selection criteria to be considered. We thus have to solve a large Multi-Person, Multi-Criteria Decision Problem (MPMCDP).

A major constraint of the early stages of an innovation process is that little or no background information is available on the ideas; the decision-makers must therefore make their judgements based on superficial descriptions of the ideas, which may only consist of a single sentence. Under such circumstances, decision-makers may consider different arguments and thus disagree in their judgements. This leads to uncertainty in the overall selection result. Uncertainty is stated to be a major difficulty in the innovation process (Leifer *et al* 2000). Some approaches to managing uncertainty in innovation processes are described in (Luoma, Paasi, Strong and Zhou, 2009). However, these approaches are only applicable in the later stages of the innovation process.

Selection can be subject to two types of error. With a rejection error, the selection procedure rejects an idea which – if it had been pursued – would have been successful. This results in lost opportunities. With an acceptance error, the selection procedure erroneously identifies an idea as (one of) the best in the pool. This leads to economic losses, since resources are invested in these ideas, which, however, ultimately prove to be unsuccessful in the market.

Clearly, additional information on the ideas such as technical feasibility, necessary investments and market attractiveness would aid the decision-makers in making better judgements. However, obtaining this information can be very expensive, since it requires experts and may involve substantial projects such as engineering studies or market research. It would therefore be prohibitively expensive and very inefficient to develop all the ideas in the first stage before evaluating them.

We are therefore faced with a dilemma: we need to identify the best ideas in the pool, but we cannot afford to generate the information that is needed to do so.

In the second and third authors' innovation consulting practice, this situation occurs frequently; because of the lack of additional information, raw ideas are debated controversially and receive widely varying evaluations. This observation formed the motivation to develop the solution described in this paper.

Based on a ranking algorithm which uses pairwise comparisons of alternatives, we have created a model which localises the uncertainty in the judgements. This allows us to restrict expensive further development of alternatives to those that are affected by the comparisons with the highest degree of uncertainty. The resulting procedure yields an overall decision that is more reliable than if no additional information is used, but at the same time can produce substantial savings compared to the brute-force approach in which all alternatives are developed before evaluation. The model is appropriate for use as a computer-based tool.

3. SCIENTIFIC PROBLEM

We are given an MPMCDP which corresponds to the first gate of an innovation process. This MPMCDP may typically involve from 5 to 10 decision-makers, from 3 to 15 decision criteria and up to several hundred alternatives. In the case of a product innovation application, the decision-makers represent various roles within a corporation such as management, marketing and engineering, the criteria include factors such as market attractiveness, technical feasibility and strategic fit, and the alternatives are raw ideas for new products that have been generated using, for example, market trends, creativity techniques or customer input.

The objective is to select a small number (typically 1 to 5) of alternatives that rank highest with respect to the set of criteria and include the opinions of all the decision-makers.

In the context of a given solution method for this MPMCDP, we wish to create a model of the uncertainty that is present in this selection. This model should localise the uncertainty and thus provide a clue as to which alternatives need to be developed further, with the expectation that a second evaluation of the developed alternatives will yield a selection result with a higher degree of certainty. Furthermore, the model should be simple to understand and to implement in a computer-based innovation management tool.

4. THE UNCERTAINTY MODEL

Our starting point for the model is the MPMCDP algorithm of Chelvier et al (Chelvier, Krull and Horton 2009) (Chelvier, Horton, Krull and Rauch-Gebbensleben 2009). This algorithm is based on pairwise comparisons of alternatives with respect to individual criteria. The results of these comparisons are then used to construct a Markov chain, in which each variable represents one of the alternatives and the arc weights represent the aggregated judgements. This Markov chain is then solved numerically using a Gauss-Seidel or Jacobi method (Stewart, 1994) to obtain a probability vector whose values are used to compute a ranking of the alternatives. The algorithm is similar in derivation and structure to the well-known PageRank algorithm (Page, Brin, Motwani and Winograd, 1999) which is used by the Internet search engine Google to compute rankings for web pages. The algorithm allows different coefficients to be assigned to represent varying degrees of importance both of the individual criteria and also of the judges.

Decision methods based on pairwise comparisons have a number of advantages over the more common scoring methods, which require decision-makers to assign numerical values to alternatives. These advantages are discussed, for example, in (Saaty and Sodenkamp 2008) and (Saaty 2008). Cooper *et al* also suggests that pairwise comparisons may be more appropriate than scoring when little information is available (Cooper, Edgett and Kleinschmidt 2001).

Our model is based on the assumption that decision-makers will give contradictory judgements when the information available for each of the alternatives under comparison is insufficient, since each decision-maker will base their judgement on the information available to themselves or on the arguments which occur to them spontaneously. On the other hand, if sufficient information is available, the decision-makers' decision will be unanimous, since we assume that decision-makers judge rationally and that a sufficient amount of information on the two alternatives under comparison will yield an unambiguous preference. (Our model ignores differences in personal taste, which will yield contradictory decisions regardless of the amount of information provided for each alternative.)

We denote the decision-makers by D_i, i=1..I, the alternatives by A_j, j=1..J and the evaluation criteria by C_k, k=1..K.

We now consider a pairwise comparison of the alternatives A_{j1} and A_{j2} with respect to criterion C_k. We collect a total of I judgements from the decision makers, of which I_{12} prefer A_{j1} over A_{j2} and I_{21} prefer A_{j2} over A_{j1}. Since these are the only two judgements allowed (alternatives are not allowed to be judged as equivalent), we have $I_{12} + I_{21} = I$. If $I_{12} = I$ and $I_{21} = 0$ or vice versa, then we have unanimous judgements; any other result we refer to as controversial. A unanimous result has a degree of controversy of 0 (all decision-makers make the same judgement), and the maximum possible degree of controversy holds for $I_{12} = I_{21} = I/2$ (equal numbers of decision-makers prefer A_{j1} over A_{j2} and *vice versa*.)

The central assumption of our model is that controversy can be interpreted as uncertainty, since we assume that a controversial result results from a lack of sufficient information, and that individual judgements are then essentially random. Similarly, we deem unanimity to correspond to a maximally certain result. In this case, we assume that all decision-makers have enough information (or prior knowledge) to determine the correct judgement. This assumption is intuitive, since the larger the majority of decision-makers that favours one alternative over another, the more certain this result one would assume this result to be.

With this assumption, we can measure the degree of uncertainty (controversy) inherent in each pairwise comparison simply by counting judgements. Comparisons with a high degree of uncertainty are assumed to be lacking in information. In practice, the two alternatives which receive a controversial comparison must then be enhanced with additional information that is relevant to the comparison. Conversely, those judgements which have a low uncertainty can be accepted as they are. After the additional information has been provided, the controversial (i.e. uncertain) comparisons can be repeated; if the additional information is sufficient, then we can expect to obtain a unanimous (i.e. certain) result.

This approach requires the specification of one parameter, namely the threshold that distinguishes controversial comparisons from uncontroversial ones. In an application with 10 decision-makers, for example, the choice might be made to treat judgements that are 5-5 or 6-4 as controversial, and all others as uncontroversial.

5. EXPERIMENTAL STUDY

In order to gain some experience with the model, we performed a small experiment. We considered a hypothetical secretary with several years experience who is considering starting her own business. We generated seven business ideas and gave them to a group of 18 non-experts for evaluation. The evaluation criterion "*Level of investment needed*" was used. Only one evaluation criterion was used, since additional criteria would have increased the workload for the test subjects without providing any further insights into the model. Of course, in a real-life situation, many more important criteria would have to be considered. The following business ideas were used for the experiment:

A. Renting out a conference room to small businesses
B. Writing business presentations
C. Healthy workplace service
D. Monthly rental of works of art to hotels, restaurants and cafes
E. Training for improved customer service
F. Writing book summaries
G. Dry cleaning delivery service

The experiment consisted of the following steps:

1. The decision-makers use pairwise comparisons of alternatives with respect to the evaluation criterion to generate individual rankings.
2. The comparisons are aggregated to form a Markov chain. The Markov chain is solved to obtain an overall ranking.
3. The controversial comparisons are identified.
4. The decision-makers are given additional information on the alternatives which are involved in controversial comparisons.
5. The decision-makers repeat the comparisons which were controversial using the new information.
6. A new Markov chain is constructed and solved to obtain a revised ranking.

For simplicity, each decision-maker was assigned the same coefficient. Again, in practice, different decision-makers may be assigned different coefficients to reflect their varying levels of experience or expertise. The choice of coefficient has no effect on our results.

The solution of the Markov chain was computed and is shown in Table 1. In this example, alternative *F* "*Writing book summaries*" has the highest probability value and was therefore ranked first, and alternative *A* "*Renting out a conference room to small businesses*" received the lowest probability value and was therefore ranked last.

Table 1: Computed ranking for the seven alternatives

Value	Idea #
0.583	*F*
0.215	*B*
0.068	*E*
0.054	*C*
0.043	*G*
0.024	*D*
0.013	*A*

Table 2 shows the distributions of the judgements obtained. The alternatives in the rows were preferred over those in the columns. Thus, for example, five decision-makers preferred alternative *D* over alternative *C*, and 13 preferred alternative *C* over alternative *D*. We considered a comparison to be controversial if it came out as 9-9 or 10-8. Using this value, only one of the 42 comparisons – namely between alternatives *E* and *C* – was controversial.

Table 2: Distribution of individual judgements

	A	*B*	*C*	*D*	*E*	*F*	*G*
A	-	1	4	5	4	0	4
B	17	-	15	16	15	4	16
C	14	3	-	13	9	1	13
D	13	2	5	-	7	0	7
E	14	3	9	11	-	3	11
F	18	14	17	18	15	-	16
G	14	2	5	11	7	2	-

In step 4 of the experiment, the decision-makers were given information on the investments that would need to be made for alternatives *E* and *C*. This resulted in a comparison which was 16-2 in favour of *E*.

After repeating the Markov chain computation, a new ranking was obtained, as shown in Table 3. Alternatives *G* and *C* have exchanged positions, otherwise, the ranking remains unchanged. It is worth noting that, of the seven alternatives, only two had to be developed further in order to obtain a result which satisfied the certainty criterion.

Table 3: Final ranking

Value	Idea #
0.582	*F*
0.215	*B*
0.081	*E*
0.044	*G*
0.041	*C*
0.024	*D*
0.013	*A*

It is difficult to make general predictions about the development savings that our approach can make possible. In the extreme case that all alternatives are contained in at least one uncertain comparison, then no savings are possible, because all alternatives must be developed. On the other hand, if all comparisons are unanimous or nearly unanimous, the model tells us that our selection has a high degree of certainty, and this is achieved without any development cost.

There are a few situations in which the unanimity of a comparison can be predicted. An alternative which is obviously superior to the others will yield unanimous comparisons. The same is true for an alternative which is obviously inferior to all the others. On the other hand, an evaluation criterion which is ambiguous will tend to yield uncertain comparisons, since the decision-makers may base their judgements on different interpretations of this criterion.

6. CONCLUSIONS

This paper presents a new model for localising uncertainty in multi-person multi-criteria decision problems. The model is designed for use with a ranking algorithm that utilises pairwise comparisons of alternatives. It is based on the assumption that controversial comparisons are a symptom of uncertainty.

The model is useful in cases with localised uncertainty, i.e. in which decision-makers disagree on a pairwise comparison for a limited number of alternatives. In such cases, the model suggests which alternatives need to be developed in order to reduce the uncertainty in the ranking. In the extreme case that many or all alternatives are part of uncertain comparisons, then no development savings are possible. In the other extreme case, in which all comparisons are unanimous or nearly unanimous, the model determines that no development is necessary.

The model is still at an early stage of development, and experience with different MPMCDPs needs to be gathered. In addition, we would like to develop an overall uncertainty score for a ranking which aggregates the uncertainties of the individual comparisons in an appropriate manner.

Our approach assumes that uncertainty is caused by a lack of information. However, experience shows that ambiguously formulated evaluation criteria can have the same effect. We therefore plan to develop a method for detecting these ambiguities in order to prompt a re-formulation.

Uncertainty is a significant problem in practice, since the cost of making an incorrect selection in the early stages of the innovation process can be considerable. We believe that our approach can make a contribution to the efficiency and the reliability of innovation processes.

7. REFERENCES

Belliveau, P., Griffin, A., and Somermeyer, S., 2002. The PDMA ToolBook for New Product Development, Wiley.

Chelvier, R., Krull, C., Horton, G., 2009. Two Heuristics for Improving the Efficiency of a Markov Chain Based Decision Making Method. *The International Workshop on Modelling & Applied Simulation,* September 2009, Santa Cruz de Teneriffe, Spain.

Chelvier, R., Horton, G., Krull, C., Rauch-Gebbensleben, B., 2009: Decision Making with a Random Walk in a Discrete Time Markov Chain. *Proceedings of 6th Vienna International Conference on Mathematical Modelling,* February 2009, Vienna, Austria.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. 2001. Portfolio Management for New Products. Perseus Books.

Leifer, R., McDermott C. M., O'Connor, G. C., Peters, L. S., Rice, M. P., Veryzer, R. W., 2000. Radical Innovation - How mature companies can outsmart upstarts, Harvard Business School Press.

Luoma, T., Paasi, J., Strong, R., Zhou, R. 2009. Uncertainty Management in Service Innovation. Proceedings of the 1st ISPIM Innovation Symposium, Singapore, 14-17 December 2009.

Page, L., Brin, S., Motwani, R., and Winograd, T., 1999. The PageRank Citation Ranking: Bringing Order to the Web. Technical Report, Stanford InfoLab, Stanford University.

Saaty, T.L., Sodenkamp M., 2008. Making decisions in hierarchic and network systems. *Int. J. Applied Decision Sciences*, 1 (1), 24–79.

Saaty, T.L., 2008. Relative Measurement and Its Generalization in Decision Making: Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors: The Analytic Hierarchy/Network Process. *Rev. R. Acad. Cien. Serie A. Mat.,* 102 (2), 251–318.

Stewart W.J., 1994: Introduction to the Numerical Solution of Markov Chains, Princeton University Press, Princeton, NJ.

PRECISION TRACKING MOTION CONTROL OF AN XY MICROPOSITIONING STAGE DRIVEN BY STEPPER MOTOR

Tlijani Hatem [(a)], Benrejeb Mohamed [(b)],

[(a)]Institut Supérieur des Sciences Appliquées et de Technologie de Gafsa (ISSAT), Tunisie
[(b)]Ecole Nationale d'Ingénieurs de Tunis (ENIT), Laboratoire de Recherche en Automatique(LA.R.A), Tunisie

[(a)]tlha74@yahoo.fr, [(b)]Mohamed.Benrejeb@enit.rnu.tn

ABSTRACT
This paper aims to present a motion coordination control approach for two axis XY stage system driven by PM stepper motors. The proposed coordinated control approach is based on a three tier composite control structure feed-forward, feedback PID and force disturbance observer. The contouring performance of a biaxial system is studied for circular trajectories and has been improved by using the proposed control scheme.To demonstrate the effectiveness of the control system design based on analysis of stage dynamics, typical results of system performance experiments micropositionning motion with high precision tracking and accurate positioning was obtained by simulations developed using Matlab and Simulink software.

Keywords: motion coordination, disturbance observer, XY stage, stepper motor, position control.

1. INTRODUCTION

Precision manufacturing has been steadily gathering momentum and attention over the last century in terms of research, development, and application to product innovation. The driving force in this development appears to arise from requirements for much higher performance of motion precision multi-axis machine. Today, ultra-precision multi-axis machine under computer control has a resolution and positioning accuracy in an order better than micrometers. In the new millennium, ultra-precision manufacture is poised to progress further and to enter the nanometer scale regime (nanotechnology) such as Micro-electro-mechanical Systems (MEMS).

In the past decade, dc motors have been widely used in these systems as high-performance drives due to the relative ease in controlling them. This ease of control is due to the fact that the system equations describing a dc motor are linear. However, there are still disadvantages in using such motors for positioning systems. In fact the mechanical commutators and brush assembly make them much more expensive. Besides, they may produce undesired sparks, which are not allowed in some applications. In particular, for high speed repetitive motion, the brushes are subject to excessive mechanical wear and consequently lead to a decrease in performance. These inherent disadvantages have prompted continual attempts to find better solution instead of dc motor. An attempt was made to use stepper motor. In fact their ability to provide accurate control over speed and position combined with their small size and relatively low cost make stepper motors a popular choice in a range of applications (F. Nollet, T. Floquet and W. Perruquetti, 2008). In particular, permanent magnet stepper motors deliver higher peak torque per unit weight and have a higher torque to inertia ratio than dc motors. Furthermore, they are more reliable and, require less maintenance, however, using the stepper motor in an open loop configuration results in low performance (G.Grellet, G. Clerc, 2000). Due to technological breakthroughs in digital signal processors, continuous time closed-loop control laws for position regulation and to go ahead and consider feedback for these incremental actuators. The Performance here is much better than the open-loop situation.(Zribi and Chiasson, 1991) considered the position control of stepper motors by exact feedback linearization.(Bodson, Chiasson, Novotnak, and Rekowski, 1993), reported on an experimental implementation of a feedback linearizing controller that guarantees position trajectory tracking by using field-weakening techniques and a speed observer. Accordingly, this study proposes a coordinated control approach based on a three tier composite control structure feed-forward, feedback PID and force disturbance observer (T.K. Kiong, L.T.Heng and H.Sunan, 2008, H.Tlijani and M. Benrejeb, 2010). The proposed approach control is applied to a biaxial XY stage motorized by two rotary bipolar PM stepper motors.

This paper is organized as follows. The stepper model and studied system description are given in Section 2. In section 3, feed-forward, feedback PID are given, and a force observer to suppress force disturbances arising from friction due to movement on motors is designed to solve the position tracking problem. The simulation and results for contouring circle trajectory under the control of the proposed scheme is described in Section 4.

2. STUDIED SYSTEM DESCRIPTION

Among the various configurations of long travel and high precision multi-axis machine, one of the most popular is known as the moving gantry. In this configuration, two motors which are mounted on two slides move a load simultaneously in tandem. This gantry system consists of four sub-assemblies, the X and Y-axis sub-assemblies, the planar platform, two bipolar stepper motors, and the end effectors. The system is equipped with a high power density due to the dual drives, and it can yield high speed motion with no significant lateral offset when the actuators are appropriately coordinated and synchronized in motion. The main challenges to address in order to harness the full potential of this configuration are mainly in the control system. In addition to precision motion control of the individual motor, efficient synchronization among them is crucially important to minimize the positional offsets which may arise due to different drive and motor characteristics, non-uniform load distribution of the gantry and attached end-effectors. The studied system is given in, figure.1.

The stepper motors, supplied with the gantry system, provide precise movements in response to electrical voltage pulses. These actuators are permanent magnet stepper motors with two-phases labeled as α and β. The electrical and mechanical parameters of the two stepping motors are given in appendix. The dynamic equations describing the used stepping motors are composed of three non-linear differential equations (1), (2) and (3) (H.Tlijani, K.B.Saad and M.Benrejeb, 2009, H.Tlijani, B.B.Salah and M. Benrejeb, 2005). These equations give a relation between the stator currents, the voltages and the mechanical quantities: torque, speed and angular position.

$$U_\alpha = Ri_\alpha + L\frac{di_\alpha}{dt} - \Omega\, k\sin(p\theta) \qquad (1)$$

$$U_\beta = Ri_\beta + L\frac{di_\beta}{dt} + \Omega\, k\cos(p\theta) \qquad (2)$$

$$C_r = -k\left[i_\alpha \sin(p\theta) - i_\beta\cos(p\theta)\right] \qquad (3)$$

$$-D\frac{d\theta}{dt} - J\frac{d^2\theta}{dt^2} - C_f\, sign(\Omega)$$

$$where\ \frac{d\theta}{dt} = \Omega$$

There are various configurations of gantry stages; many of them are intrinsically similar. A typical gantry stage may be considered as a two-degree of freedom (2DF) servo-mechanism, which can be adequately described by the schematics in fig.1.

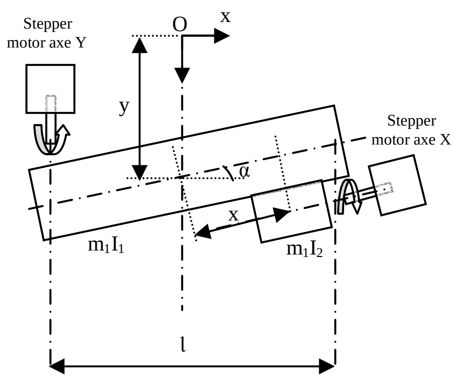


Figure 1: Motion reference using gantry system

A servomotor carries a gantry on which a slider holding the load is mounted. One motor yields a linear displacement y, while the other yields a linear displacement x. and also the dynamic loading present due to the translation of the slider along the gantry. The central point of the gantry is thus constrained to move along the dashed line with two degrees of freedom. The displacement of this central point from the origin O is denoted by y. The gantry may also rotate about an axis perpendicular to the plane of fig.1 due to the deviation between x and y, and this rotational angle is denoted by α.The slider motion relative to the gantry is represented by x. It is also assumed that the gantry is symmetric and the distance from its central point to the slider mass center is denoted by d. With this formulation of the gantry stage, it is imminent to proceed with the dynamic modeling of the gantry stage. This model has been introduced by (T.K.Kiong, L.T.Heng and H.Sunan, 2008, Chuan. Shi, Peqing.Ye, Qiang.Lv, 2006).

Let m_1, m_2 denote the mass of the gantry and slider respectively, l denotes the length of the gantry arm, I_1, I_2 denote the moment of inertia of the gantry arm and slider respectively, we assume that

$$I_1 = m_1(l/2)^2,\ I_2 = m_2(l/2+x)^2$$

$$X = \begin{bmatrix} x & \alpha & y \end{bmatrix}^T$$

$$y = y_2 + (y_1 - y_2)/2$$

The positions of $m_i, i = 1,2$ are given by

$$x_1 = 0$$

$$y_1 = y$$

$$x_2 = x\cos\alpha + d\sin\alpha$$

$$y_2 = y + d\cos\alpha - x\sin\alpha$$

Which lead to the corresponding velocities.

$$v_1 = \begin{bmatrix} 0 \\ \dot{y} \end{bmatrix}$$

$$v_2 = \begin{bmatrix} \dot{x}\cos\alpha - x\dot{\alpha}\sin\alpha + d\dot{\alpha}\cos\alpha \\ \dot{y} - d\dot{\alpha}\sin\alpha - \dot{x}\sin\alpha - x\dot{\alpha}\cos\alpha \end{bmatrix}$$

Thus, the total kinetic energy may be computed as

$$\begin{aligned} K &= \frac{1}{2}m_1 v_1^T v_1 + \frac{1}{2}m_2 v_2^T v_2 + \frac{1}{2}(I_1 + I_2)\dot{\alpha}^2 \\ &= \frac{1}{2}(m_1 + m_2)\dot{y}^2 + \frac{1}{2}(I_1 + I_2 + m_2 x^2 + m_2 d^2)\dot{\alpha}^2 \\ &+ \frac{1}{2}m_2\dot{x}^2 - \dot{y}\dot{\alpha}m_2(d\sin\alpha + y\cos\alpha) - \dot{x}\dot{y}m_2\sin\alpha \\ &+ \dot{\alpha}\dot{x}m_2 d \end{aligned}$$

This can be further written as

$$K = \frac{1}{2}\dot{X}^T I \dot{X} \qquad (4)$$

Where I is the inertia matrix given by:

$$I = \begin{bmatrix} m_1 + m_2 & -m_2 d\sin\alpha - m_2 x\cos\alpha & -m_2\sin\alpha \\ -m_2 d\sin\alpha - m_2 x\cos\alpha & I_1 + I_2 + m_2 x^2 + m_2 d^2 & m_2 d \\ -m_2\sin\alpha & m_2 d & m_2 \end{bmatrix}$$

The elements of the Coriolis and centrifugal matrix A can be derived from

$$A_{ij} = \sum_{k=1}^{3} a_{ijk}\dot{q}_k$$

Where $\dot{q}_1$, $\dot{q}_2$ and $\dot{q}_3$ represents the derivative of x, α and y respectively, and c_{ijk} are computed as:

$$a_{ijk} = \frac{1}{2}\left(\frac{\partial e_{ij}(q)}{\partial q_k} + \frac{\partial e_{ik}(q)}{\partial q_j} + \frac{\partial e_{jk}(q)}{\partial q_i}\right) \qquad (5)$$

Where e_{ij} represents the element of the inertia matrix I .Substituting the assumed inertia equation into equation (5), matrix A can be expressed as

$$A = \begin{bmatrix} 0 & x\dot{\alpha}s\alpha - d\dot{\alpha}c\alpha - \dot{x}c\alpha & -\dot{\alpha}c\alpha \\ x\dot{\alpha}s\alpha - d\dot{\alpha}c\alpha - \dot{x}c\alpha & (xs\alpha - dc\alpha)\dot{y} - (\frac{l}{2} + 2x)\dot{x} & (\frac{l}{2} + 2x)\dot{\alpha} - \dot{y}c\alpha \\ -\dot{\alpha}c\alpha & (\frac{l}{2} + 2x)\alpha - \dot{y}c\alpha & 0 \end{bmatrix}$$

Where, $\cos\alpha = c\alpha, \sin\alpha = s\alpha$, finally the dynamic model is expressed as

$$I\ddot{X} + A\dot{X} + BF = BU \qquad (6)$$

Where

$$B = \begin{bmatrix} 1 & 1 & 0 \\ l\cos\alpha & -l\cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$F = \begin{bmatrix} F_x & F_\alpha & F_y \end{bmatrix}^T, U = \begin{bmatrix} u_x & u_\alpha & u_y \end{bmatrix}^T$$

F_x, F_α, F_y are the frictional forces, and u_x, u_α, u_y are the generated mechanical forces.

3. COORDINATED CONTROL SCHEME

In the proposed coordinated control approach, motors are assigned to horizontal axe x and vertical axe y. A supervisory motion program drives the axis through these actuators which share an identical commanded trajectory pre-planned. Each servo loop then has the responsibility of keeping the actual trajectory as closely as possible to the commanded trajectory, since each of motors has its own individual servo loop. Presuming they have tight servo loops, this method provides a tight and smooth link between the motors.Figure.2 provides a block diagram of the proposed coordinated control scheme.

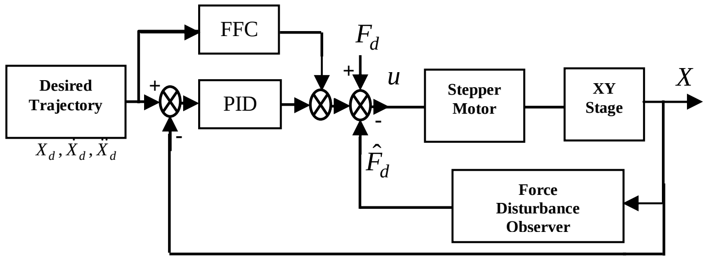


Figure 2: Block diagram of the positioning control system

As shown in the above scheme, a three tier composite control structure is adopted: feed-forward control, feedback control and force disturbance observer. This design possesses several important and useful features. First, it incorporates a feed-forward component to facilitate a high speed response. The feed-forward component addresses model based characteristics relating to the stepper motors. Second, an optimal PID feedback controller is designed and intended to provide optimal command response and stability properties. Third, since the achievable performance of a precision positioning system is unavoidably and very significantly limited by the amount of disturbances present, and the uniformity of their distribution among the motors, a disturbance observer is augmented to the composite controller structure. It provides a fast response to load disturbances and other exogenous signals acting asymmetrically on the two motors. This feature is especially useful since load disturbances are major factors affecting the control performance, especially when the motors jointly carry a dynamical and asymmetrical load such as an additional servo system running across the system (Yo.Tomita, K.Makino and M.Sugimine, 1996). It is used to estimate the actual disturbance, deduced from a disturbance observer, to compensate for the disturbances force.

3.1. Feed-forward Control

The servo system at equation (3) can thus be alternatively described by:

$$\ddot{\theta} = -\frac{D}{J}\dot{\theta} - \frac{k}{J}\left[i_\alpha \sin(p\theta) - i_\beta \cos(p\theta)\right] - \frac{C_r}{J} - \frac{C_f}{J} sign(\dot{\theta})$$
$$\ddot{\theta} = -\frac{D}{J}\dot{\theta} + \frac{k}{J}\left[-i_\alpha \sin(p\theta) + i_\beta \cos(p\theta)\right] + \frac{k}{J} g(\theta, \dot{\theta})$$

$$g(\theta,\dot{\theta}) = -[\frac{C_r}{k} + \frac{C_f}{k} sign(\dot{\theta})]$$

$g(\theta,\dot{\theta})$ is assumed to be a smooth non-linear function which may be unknown. With the tracking error e defined as: $e = \theta_d - \theta$, we have :

$$\ddot{e} = -\frac{D}{J}\dot{e} - \frac{k}{J}\left[-i_\alpha \sin(p\theta) + i_\beta \cos(p\theta)\right] - \frac{k}{J} g(\theta,\dot{\theta}) \quad (7)$$
$$-\frac{k}{J}[-\frac{J}{k}\ddot{\theta}_d - \frac{D}{k}\dot{\theta}_d]$$

the system state variables are assigned as:

$$x_1 = \int_0^t e(t)dt, \quad x_2 = e, \quad x_3 = \dot{e}$$

denoting: $X = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T$,the equation (7) can then be put into the equivalent state space form:

$$\dot{X} = \Lambda X + \Gamma[-i_\alpha \sin(p\theta) + i_\beta \cos(p\theta)] + \Gamma g(\theta,\dot{\theta}) + \Gamma[-\frac{J}{k}\ddot{\theta}_d - \frac{D}{J}\dot{\theta}_d] \; (8)$$

$$\Lambda = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -\frac{D_{r0}}{J} \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 0 \\ 0 \\ -\frac{k_\phi}{J} \end{bmatrix}$$

The design of the feed-forward control law is straightforward. From equation (8) the term:

$$\Gamma[-\frac{J}{k}\ddot{\theta}_d - \frac{D}{J}\dot{\theta}_d]$$

may be neutralised using a feed-forward control term in the control signal. The feed-forward control is thus designed as:

$$u_{FF}(t) = \frac{J}{k}\ddot{\theta}_d + \frac{D}{J}\dot{\theta}_d \qquad (9)$$

Clearly, the reference position trajectory must be continuous and twice differentiable; otherwise a pre-compensator to filter the reference signal will be necessary. The only parameters required for the design of the feed-forward control are the parameters of the second-order linear model.

3.2. Feedback Control

In the composite control system, PID is used as the feedback control term. While the simplicity in a PID structure is appealing, it is also often proclaimed as the reason for poor control performance whenever it occurs. In this design, advanced optimum control theory is applied to tune PID control gains. The PID feedback controller is designed using the Linear Quadratic Regulator (LQR) technique for optimal and robust performance of the nominal system. The nominal portion of the system is given by:

$$\dot{X} = \Lambda X + \Gamma[-i_\alpha \sin(p\theta_m) + i_\beta \cos(p\theta_m)]$$

$$u_{PID} = \Psi X = \psi_1 x_1 + \psi_2 x_2 + \psi_3 x_3$$

This is a PID control structure which utilizes a full-state feedback is well known in modern optimal control theory. The PID control is given by

$$u_{PID} = -R^{-1}\Gamma^T \Delta X \quad (10)$$

Where Δ is the positive definite solution of the Riccati equation :

$$\Lambda^T \Delta + \Delta\Lambda - \Delta\Gamma R^{-1}\Gamma^T \Delta + Q = 0$$

3.3. Design of the Disturbance Observer

As the achievable, performance of a precision positioning is unavoidably and significantly limited by the amount and the uniformity of disturbances, among the motors. A disturbance observer is augmented to the composite control structure to provide a fast response to load disturbances and other exogenous signals acting asymmetrically on the actuators. In figure.3, X, u, F and

$\hat{F}_d$ denote the position signal, control signal, actual and observed disturbance force signal associated with the axis system. H and H_n denote respectively the actual system, and the nominal system.

$$H_n(p) = \frac{a_0}{p^l(p^{m-1} + a_1 p^{m-l-1} + \ldots + a_{m-l-1}p + a_{m-l})},$$

Here, a third order model will be used, $l = 1, m = 2$:

$$H_n(p) = \frac{K}{p(Tp+1)} \quad (10)$$

The disturbance observer incorporates the inverse of the nominal system, and thus a low pass filter G is required to make the disturbance observer proper and practically realisable. For the choice of a second order model H_n , a suitable filter is:

$$G(p) = \frac{g_2}{p^2 + g_1 p + g_2} \quad (11)$$

the disturbance observer is equivalent to an additional disturbance compensator C_{obser} , which closes a fast inner loop. It can be shown that:

$$C_{obser} = \frac{G}{1-G} H_n^{-1} \quad (12)$$

For the choice of H_n and G , it follows that

$$C_{obser} = \frac{g_2(Tp+1)}{K(p+g_1)} \quad (13)$$

Therefore, C_{obser} can be considered as a lead/lag compensator by appropriately designing g_1 and g_2 relatively to K and T .

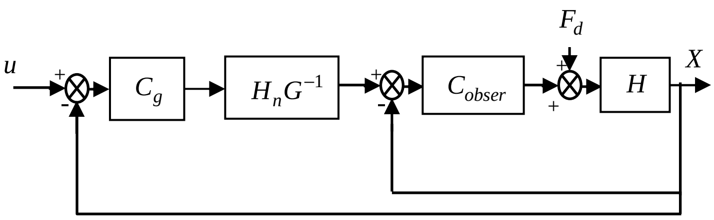


Figure 3: Control system with disturbance observer

The disturbance observer can be designed in many ways. One possible approach is given as follows:

• Identify the nominal model K and T , based on which the outer loop controller C_f can be designed to achieve a desired command response.
• Adjust g_1 and g_2 of the disturbance compensator C_{obser} to satisfy requirements for robustness and disturbance suppression characteristics. The system sensitivity function and the system transmission function can thus be set independently.
• Carry out simulation and fine tuning till the performance is acceptable.

4. SIMULATION AND RESULTS

Real-time simulation is carried out on biaxial system with one motor and digital encoders each along the x and y direction. The control task in the simulation is to execute planar motion with circular trajectory as straightly and precisely as possible. In the simulation, the two motors have the same dynamical properties. Simulation results are provided to illustrate the effectiveness of the control scheme. The figures (4,5) illustrate respectively the position, and the velocity of axis x and y necessary for the synchronization of biaxial motion. The tracking performance achieved from the use of composite control is given in figure.7, showing that a maximum tracking error of less than 0.01 mm is achieved in the generation of circle profile figure 6.

5. CONCLUSION

This study has presented the coordinated control approach based on three composite controllers, feedback component (PID), a feedforward component (FFC) and a force disturbance observer component. It has shown that is satisfied the desired velocity and acceleration of a stepper motor in the control mode of the incremental motion. It address, also several important challenges to the design of precision motion coordination for two axis XY stage system driven by PM stepper motors. To provide evidence on the effectiveness of the control system design simulation results are obtained by using Matlab and Simulink. It has also shown that the proposed method is more suitable than the conventional method.

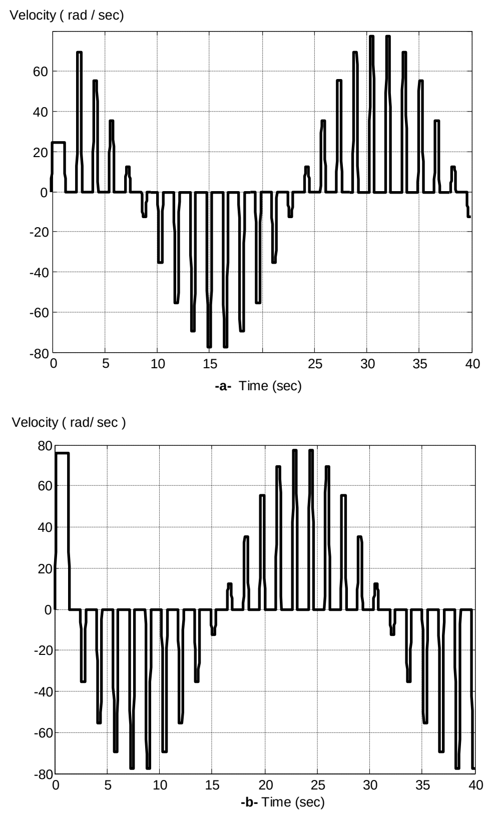


Figure 4: Velocity profile for circular motion:
a) axe x – b) axe y

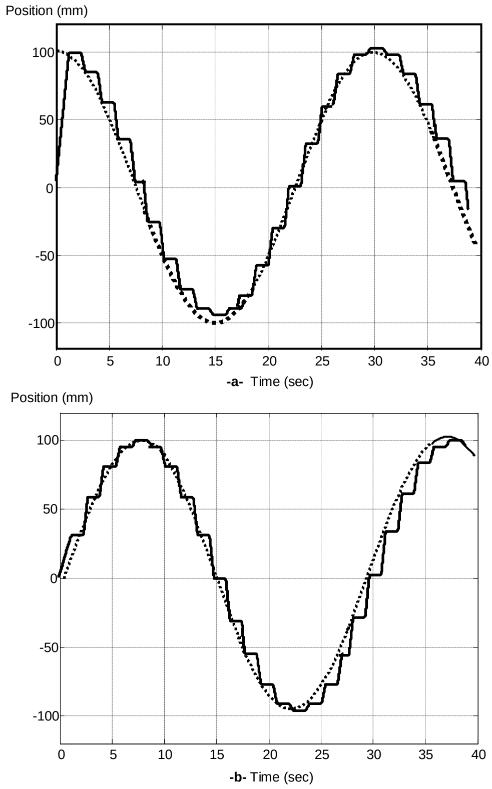


Figure 5: Position profile for circular motion:
a) axe x - b) axe y, desired position (dotted line)

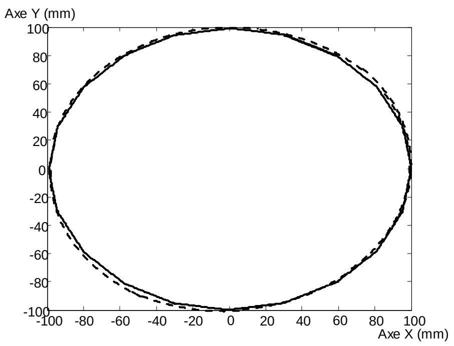


Figure 6: Circular tracking motion: desired circle profile (dotted line)

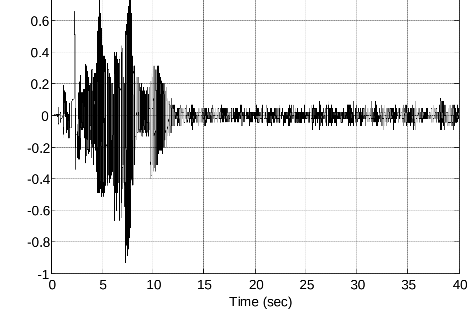


Figure 7: Tracking error position for circular motion

REFERENCES

H.Tlijani, K. Ben Saad, M. Benrejeb, Incremental Actuators Motion Control, Using a Fuzzy PID Controller, *International Review of Automatic Control*, vol.2N.1, p.p.121-127, 2009.

T.K. Kiong, L.T. Heng, H.Sunan. Precision Motion Control Design and Implementation. Springer, Verlag London Limited, 2008.

Chuan. Shi, Peqing. Ye, Qiang. Lv, Study on Control Algorithm for Continuous Segments Trajectory Interpolation. *Proceeding of the 1st International Symposium on Digital Manufacture*, Oct.2006, 994-998.

H.Tlijani, M. Benrejeb, B.Ben Salah, Approches de commande pour le contournage des primitives et des trajectoires complexes utilisant deux moteurs pas à pas couplés, *Quatrième Conférence Internationale IEEE, Journée de la Technologie Electrique et d'Automatique, Hammamet, 2006.*

H.Tlijani, B.Ben Salah, M. Benrejeb, On coordination of multi axis system movements using stepping actuators, *Third IEEE International Conference on Systems, Signals and Device*s, Sousse, 2005.

H.Tlijani, M. Benrejeb, Sliding coordinated control for moving gantry using incremental actuators, *Sixiéme Conférence Internationale IEEE, Journée de la Technologie Electrique et d'Automatique*, Hammamet, 2010.

B.Ben Salah, M. Benrejeb, A digital control for improving the position resolution of Permanent-magnet stepping motors, Systems Analysis Model.Simulation, Edition Taylor & Francis, Vol. 43, N° 2, pp.189-200, 2003.

K.K Shyu, C.K.Lai Incremental Motion Control of Synchronous Reluctance Motor Via Multisegment Sliding Mode Control Method, *IEEE Transactions on control systems technology*, vol.10, N° 2, March 2002.

G.Grellet, G. Clerc, Actionneurs Electriques. Editions Eyrolles, Paris, 2000

Yo. Tomita, K. Makino, M. Sugimine. High Response XY Stage System Driven by Parallel Linear Motors, Annals of the ClRP vol. 45, 359-362, 1996.

F. Nollet, T. Floquet_, W. Perruquetti, Observer-based second order sliding mode control laws for stepper motors *Control Engineering Practice* vol.16, 429-443, 2008.

Bodson. M, Chiasson.J, Novotnak, R., Rekowski, R, High performance nonlinear feedback control of a permanent magnet stepper motor, *IEEE Transactions on Control Systems Technology*, vol.1, 5-14, 1993.

Zribi. M, Chiasson. J, Position control of a PM stepper motor by exact linearization. *IEEE Transactions on Automatic Control*, vol.36, 620-625, 1991.

A SIMULATION-BASED FRAMEWORK FOR CONSTRUCTION PROJECT INFORMATION MANAGEMENT

Elmira Moghani [(a)], Simaan AbouRizk [(b)]

[(a)]PhD Candidate, Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada
[(b)]Professor, Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada

[(a)] moghani@ualberta.ca [(b)] abourizk@ualberta.ca

ABSTRACT
Modern capital projects generate large amounts of documentation in several different formats and controlled by many different project parties: everything from design drawings to cost estimates, schedules, equipment information, change orders, and work logs. Accessing this documentation can be both difficult and time-consuming. Project managers clearly need efficient methods to manage project information; the current methods focus on capturing and integrating information through a database and CAD drawings, but they are limited in scope, not easy to re-use in other projects, and do not capture the construction process. The solution is a well-structured simulation-based system that can dynamically capture, store, process, and access all project information, from the planning stage through the construction process to the completed project, including all changes to the original plan. The motivation for a simulation-based foundation for this approach stems from its ability to model the dynamic processes involved rather than just the static information.

Keywords: project information, construction process, High Level Architecture, Distributed simulation

1. INTRODUCTION

Complex capital projects can involve thousands of workers, hundreds of millions of dollars, and many years of work. These types of projects generate large amounts of documentation in many different formats and controlled by many different project parties: everything from design drawings to cost estimates, schedules, equipment information, change orders, and work logs. During the construction period, a company generates reports in certain areas, such as workers' time cards and the actual as-built drawings of the project, and projects are also almost never completed exactly as planned; keeping track of changes in the building process is very complicated, adding yet another layer of documentation. However, even this level of documentation is incomplete, especially when it comes to recording the actual building process.

Accessing this documentation can also be both difficult and time-consuming. Because the existing documentation is stored based on the interest of each party (e.g., drawings may be kept with the design consultant, the foreman might have worker records, and the project manager might have the financial information), the project information is disorganized and reconstructing a record of how the project was completed may require hiring experts. This can be a problem, particularly when it comes to settling claims and dealing with quality control, both of which can mean losses of millions of dollars (Akinci 2004).

It should therefore be unsurprising that project managers and contractors are always looking for efficient ways to collect and store construction histories and retrieve them effectively (Hendrickson and Au 1989). Different studies have focused on developing a computer based information system to integrate the collection, processing and transmission of information (Lock 1993). A database management system (DBMS) has become a common solution to overcome some of the limitations of data sharing (Mazerolle and Alkass 1993; Bowler 1994; Dawood et al. 2002). But the current methods do not capture the construction process; existing methods usually focus on capturing and storing physical information from computer-aided design (CAD) drawings, or the scope of the project from as-planned and as-built documents, rather than capturing the actual building process. They are also limited in scope, focusing on small parts rather than the entire project. For example, project managers who are interested in equipment control try to capture information such as working hours, breakdown or fuel consumption of onsite equipment, but estimators are interested in collecting the workers' daily log or material cost (Navon 2005; Fayek et al. 1998). Even integrating the project information database using the Critical Path Method (CPM) does not support capturing the complete operation; a CPM network only describes tasks to a certain level of information, such as duration and resources.

The solution is a well-structured simulation-based system that can dynamically capture, store, process, and access all project information, from the planning stage through the building process to the completed project, including all changes to the original plan. Ideally, the system should integrate this information and present it to the user, providing an overview of the project at every stage for comparison with the project as planned. It should be able to incorporate process models, product models, resource models, and static information in one system. A user should be able to see the resources that completed a given scope of work in a given period of time under the influence of external factors (e.g. weather).

This kind of documentation will be helpful for claims, control purposes, reproducing actual product drawings, operation and maintenance, and planning for future similar projects. It will also increase the learning process in the involved organizations for future projects by enabling access to the information of previous projects (FIATECH 2009). The solution must be easy to use from a manager's perspective, and it must also be reusable for different types and sizes of construction projects.

One promising method for implementing the solution is computer simulation. Computer simulation was introduced to the area of construction research by Halpin (1977) with his proposed CYCLONE system. The proposed system and all derivatives were used in construction research to model real construction processes, leading construction managers to more efficient use of materials, manpower, and equipment (Paulson et al. 1987; Ioannou 1989; Martinez and Ioannou 1994; Hajjar and AbouRizk 2002). Computer simulation has been successfully implemented in the construction industry for many purposes (Sawhney 1994), and has already been effectively used in managing information in the construction domain by integrating database management systems with a simulation model (Moghani et al. 2009). In this approach, a construction manager can use the collected data to re-run a simulation process model and examine the effect of new data on project performance. However, current modeling approaches cannot integrate or assimilate information from different sources and by different participants in an efficient and organized manner. For a more complex project that needs different information from different participants or software, distributed simulation is a new simulation technique that can facilitate modeling effort and integration in the simulation environment.

AbouRizk introduced High Level Architecture-based distributed simulation to simulate construction projects (2006), facilitating integration, collaboration, and reusability of the simulation model. High Level Architecture (HLA) was developed in 1995 by the Department of Defense (DoD) as an advanced technique for integrating simulation models to support reuse and interoperation of simulation models and reduce the cost and effort of modeling in simulation projects (Fujimoto 2003). Under the HLA standard, different developers can build individual components (federates) of one system (a federation), maintaining interoperability between them. This approach allows us to standardize the integration process between different computer software, simulation systems, and from different users; it is therefore a promising implementation for the proposed information management solution.

This study aims to develop an interoperable and reusable simulation-based framework to integrate project information including as-planned, process, and as-built information to enable construction managers to create the real history of the project from planning to completion. This framework will utilize HLA and a software application framework approach (Froehlich et al. 1998) for its implementation to be a reusable, extensible and modular model, with flexibility in representing different modeling approaches and data forms in the same simulation. In this framework a 3D CAD model will be integrated with process simulation model to automatically capture design information and use it for simulation purposes, and a relational database will be the medium for integrating as-built information with the simulation model and updating the process accordingly. The project is under development using the Construction Synthetic Environment (COSYE) as its base; it is focused on repetitive construction such as tunnelling, which lends itself better to simulation planning (AbouRizk and Ruwanpura 1999; Fernando et al. 2003; Al-Bataineh 2008; Marzok et al. 2008). The case studies used or planned are existing tunnelling projects in Edmonton, Alberta, Canada.

2. HLA-BASED TUNNEL SIMULATION FRAMEWORK

COSYE (<u>Co</u>nstruction <u>Sy</u>nthetic <u>E</u>nvironment), the simulation environment used as a base for this research, was developed based on HLA standards to facilitate modeling more complex projects, such as tunnel construction, which involves many activities, equipment, materials and human resources.

The HLA standards consist of three main components: the HLA rules (IEEE 1516 2000), the interface specifications (IEEE 1516.1 2000), and the Object Model Template (OMT) (IEEE 1516.2 2000). The HLA rules outline the creation of a federation and federates and cover all their responsibilities to ensure a consistent implementation and proper interactions. The interface specification defines the functional interfaces between federates and the run time infrastructure (RTI). RTI software provides interface services that support federates' interactions and federation management, such as transferring the responsibility of updating an attribute between federates, managing data distribution and assisting with time management in the federation. Any software can interact through the RTI as a single federate. The Object Model Template is a standard for defining and documenting the form, type, and structure of data shared within a simulation, and it consists of three different object models: the Federation Object Model (FOM), the Simulation Object Model (SOM) and the Management Object Model (MOM).

The COSYE Framework is a software application that supports development of federations in Microsoft Visual Studio. In this Environment, the RTI server is a

.NET implementation of the HLA RTI and runs as a windows server, and the COSYE OMT Editor is used to develop and edit OMT documents in Visual Studio (AbouRizk and Hague 2009)

Based on HLA standards, COSYE enables the integration of different computer software and different simulation techniques in a single environment to model all the processes involved in large scale projects. HLA services in COSYE also enable collaboration of different experts to develop various parts of a simulation model. For instance, in the tunnel federation, one may simulate the tunnel construction process while the other focuses on material supply simulation.

In this project, the tunnel federation simulates the whole construction process of a utility tunnel including excavation and lining of the working shaft, the tunnel and the retrieval shaft. Using HLA requires division of the simulation into different federates and the HLA architecture thus helps to create modular, reusable, and extensible modeling elements.

2.1. Brief Overview of Tunnelling

Constructing a utility tunnel occurs in different phases. First, to get access to the tunnel excavation depth, a vertical shaft called a construction shaft is usually excavated. This shaft is the main access during the construction process for lowering the equipment into the tunnel or removing the excavated dirt. Shaft construction is usually done in sections with the depths dependent on the soil type and geometry of the shaft. The excavation and lining processes are done for each section sequentially. After finishing the excavation and lining for shaft sections, the tunnel construction process can start.

Tunnel construction methods vary depending on geotechnical information, availability of equipment, and the geometry of a tunnel. For long tunnels or tunnels with a large diameter, a Tunnel Boring Machine (TBM) is typically utilized for excavation and lining processes while for short or small tunnels, hand excavation is a better option. Before starting the tunnel construction, the tail tunnel and undercut area (an enlargement at the bottom of the shaft used for staging material handling and dirt removal operations) are excavated. The tunnel construction activities are divided into:

1. Excavation of the tunnel
1. Removing the dirt from the excavation area and transferring it to the construction shaft (using muck carts).
2. Hoisting the dirt to the ground level (using a crane, clamshell bucket, gantry, etc.) and transfer it outside the construction area
3. Lowering down the liners and transferring them to the excavation face.
4. Lining the tunnel
5. Extending construction and utility services
6. Excavating and supporting the removal shaft in a case of using TBM.

Figure 1 shows a typical utility tunnel layout.

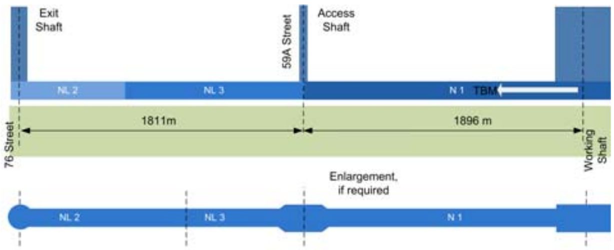


Figure 1: Layout of a Utility Tunnel

2.2. Developing the Tunnelling Product Model

In order to be able to create a generic federation and integrate the CAD model with the simulation, a conceptual project model consisting of product, process, management, environment, and resource models was developed (Figure 2).

The tunnel construction method for utility facilities was investigated, and all the information regarding activities and required resources was collected. In order to develop a product model, the CAD drawings were reviewed and the attributes were defined and added to the project model. It is important to mention that these models are for utility tunnels, and they mostly focus on the TBM (tunnel boring machine) tunnelling method.

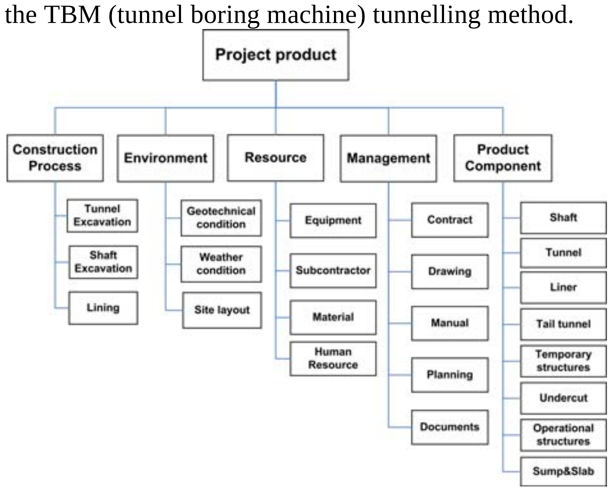


Figure 2: Tunnel Construction Project Model

2.3. Developing a Federation Object Model (FOM) for the Tunneling Federation based on the Product and Process Model

The first and most important task in developing the simulation model based on High Level Architecture (HLA) is to define the HLA federation object model. The FOM is composed of a group of interrelated components specifying information about classes of objects, interactions, attributes, and their parameters. The tunnelling federation object model was developed based on conceptual project models and is still being improved. Figure 3 shows the FOM in the tunnel federation.

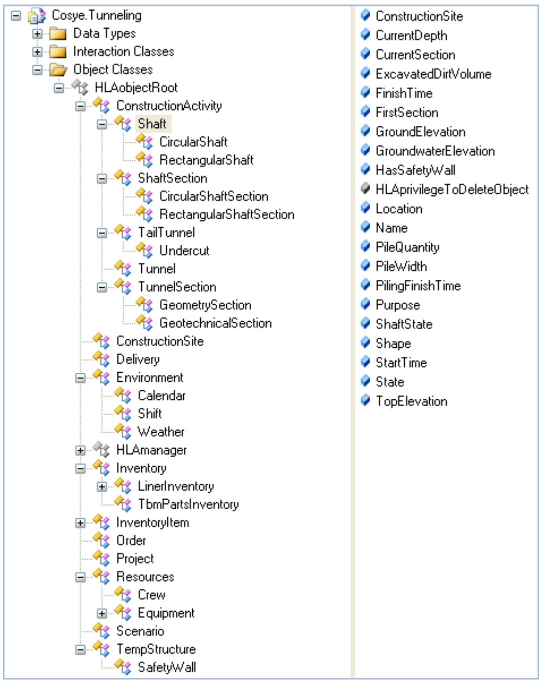


Figure 3: Tunnel Federation Object Model

2.4. Overview of Tunnelling Federation

Figure 4 shows the proposed framework for the simulation model, based on available studies and expert knowledge gathering. Most federates were developed based on a case study for tunnel construction in the City of Edmonton using a specific construction method, and will be used as a base for further development.

The scenario setup federate is designed for the user to configure different tunnel projects and scenarios. In the current federation, the user inputs all necessary parameters at this federate: shift length, shift start time, coffee break duration, lunch break duration, project start date, work status on weekends or holidays; project setting such as number of shafts, number of tunnels, and their attributes like tunnel length, section length, etc.; resource setting refers to equipment information and crew information. This information will be passed to other relevant federates through RTI.

The shaft federate simulates the shaft construction process including preparation, excavation, and lining. The tunnel federate is designed to simulate the complete process of constructing a tunnel including excavation, lining, resetting TBM, and TBM breakdown; it also covers other common activities, like extending utilities and surveying.

The dirt removal federate models the process of removing dirt from the tunnel face to the undercut, dumping dirt from undercut to ground, and loading carts with materials.

The supplier federate has the same responsibility as a supplier contractor: receive a new order from the contractor through the procurement federate, schedule for delivery of a new order and send a response to the procurement federate.

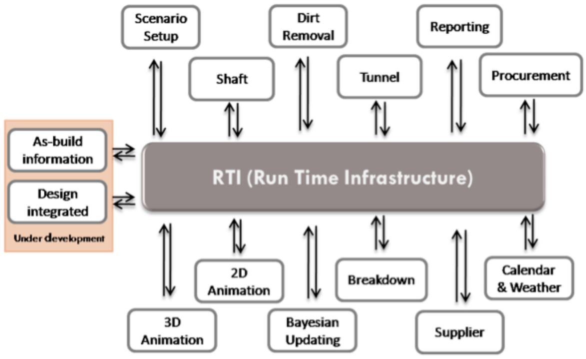


Figure 4: Tunnel Federation

There can be more than one supplier which can join the federation and receive orders from the procurement federate. The procurement federate plays the same role of the procurement office in the tunnelling contractor group. It tracks how many liners are in the inventory; if that level reaches a specific threshold, the procurement federate places an order for the concrete lining segments to the supplier.

The Bayesian updating federate is designed to apply a Bayesian updating method for predicting machine breakdown, TBM advance rate, tunnel productivity, scheduling and cost.

The visualization federate displays a 2D and 3D animation of the construction process as the simulation is running. It is a real time visualization of the entire process and shows the different states of the TBM during tunnel construction, the excavation and lining of the tunnel, as well as traveling muck carts in the tunnel. In current federation, all the mentioned federates are fully functional for the NEST tunnel case study in Edmonton, Canada.

Future development will involve connecting a 3D CAD drawing to the simulation model and extracting as-planned geometry and geotechnical information. This federate will remove the need for manual data entry, especially when new revisions of drawings are released.

Capturing and storing all the as-built information will be based on automatic data transfer from the site and data will be stored in the SQL Server database. Figure 5 illustrates an example of relationship in the database.

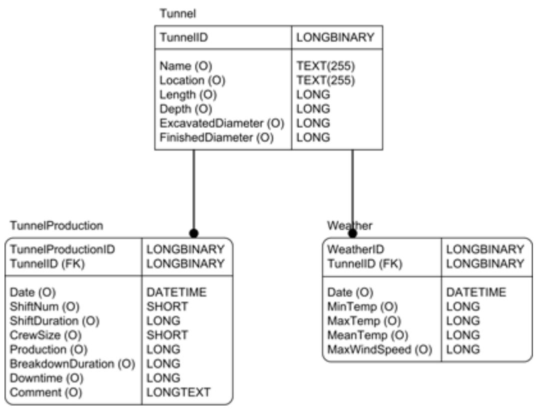


Figure 5: ER Diagram of Database

3. INFORMATION MANAGEMENT STRUCTURE

To fulfill the objective of this research, a vital part of the implementation process is to study the possible ways of representing information within the COSYE environment and create a conceptual structure for the data management plan. We determined the required information by answering the following questions:

- What kind of information will be useful for decision making process and different participants in the project?
- What kind of information from the planning stage will be required for developing the simulation model?
- What kind of information from the jobsite will change the inputs of the simulation model and how can we store it to make it exchangeable with original inputs?
- How can we make the history of projects reusable for future projects or simulation models?

Answering these questions determines the information required from a tunnelling construction project in order for the simulation to be representative. Finally, to collect and integrate on-site process information (as-built) and to be able to store the as-planned data and their relationships throughout a simulation model, an accurate knowledge management system needs to be developed. For this purpose, a comprehensive relational database is proposed to store planning data, actual as-built data, and the history of project process changes. A view of our vision is given in Figure 6.

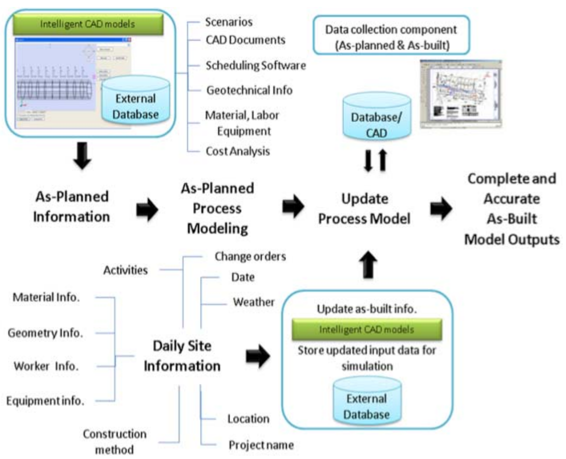


Figure 6: Integrated Framework for Tunnelling Construction

4. FUTURE WORK

In the future, new federates such as a design information federate will be developed to facilitate information transfer from CAD systems to the simulation model, and a relational database will be designed and connected to the federation to store as-built information as the project advances.

To make changes to the process model during the construction phase without a need to re-create the entire model (e.g., adding one removal shaft or changing the construction method of the tunnel), generic and reusable federates are required. To develop such federates; the current FOM needs to expand to provide all the object classes, attributes, and interactions for different tunnel construction methods, activities, equipments, and the required information from the planning and construction stage.

The simulation model and design inputs will then be updated based on new process and information (reconstructed from the beginning of the project but as if the project is going on now – i.e., simulate the simulation) to mirror any changes in design and construction, create an accurate model and verify it. The product/process data integration will be evaluated and adjusted if necessary.

Once the NEST case application is functional, the work will be repeated with another tunnel construction project to identify areas where the framework must be adjusted and an application framework approach as a basis of this research will be utilized to provide a generic framework with modularity, reusability and extensibility.

5. CONCLUSION

The ability to access detailed information of a project along with the actual construction process (as it was built) will be highly beneficial for the management process for different purposes such as quality control, claims, process improvement and future project planning. This paper proposes the development of a generic framework for information management in repetitive construction projects such as tunneling to incorporate process model, product model, resource model, and static information in one system. Deploying HLA-based distributed simulation and using a software application framework enables us to standardize the integration process between simulation systems, different computer software, and project models and supports interoperability, reusability, and extensibility of this framework. While all COSYE development to date has taken place in an academic environment and COSYE itself has not yet been commercialized, development of a full federation typically takes between one half to one full man-year (1,000 to 2,000 hours).

With this system, project managers will have full access to the project data. Moreover, they will have an accurate model at the end of the project for reviewing purposes, documentation for future work, and educational purposes. The system will save time and money, improve training, and increase operational efficiency.

ACKNOWLEDGMENTS

This study is supported by National Science and Engineering Research Council (NSERC) Industrial Research Chair in Construction Engineering and Management research.

REFERENCES

AbouRizk, S.M., Hague, S. 2009, An Overview of the COSYE Environment for Construction Simulation, *Proceedings of the Winter Simulation Conference*, pp. 2624-2634. Austin (TX, USA)

AbouRizk, S.M., 2006. Collaborative Simulation Framework for Multi-user Support in Construction. *Discovery Grant Proposal.*

AbouRizk, S.M., and Mather, K., 2000. Simplifying Simulation Modeling Through Integration with 3-CAD. *Journal of Construction Engineering and Management*, 126(6), 475-483.

AbouRizk, S.M., Ruwanpura, J.Y., Er. K.C., and Fernando, S., 1999. Special Purpose Simulation Template for Utility Tunnel Construction. *Proceedings of the 31st Winter Simulation Conference*, pp. 948-955. Dec. 05-08, Phoenix (Arizona, USA).

Akinci, B., 2004, Using Sensor Systems and Standard Project Models to Capture and Model Project History for Building Commissioning and Facility Management. *Facility Area Network Workshop*, Feb. 26-27, Civil Engineering Research Lab, Urbana-Champaign, IL

Al-Bataineh, M.T., 2008. *Scenario-Based Planning for Tunneling Construction*. Thesis (PhD). University of Alberta.

Dawood, N., Sriprasert, E., Mallasi, Z., and Hobbs, B., 2003. Development of an integrated information resource base for 4D/VR construction processes simulation. *Automation in Construction*, 12(2), 123-131.

Fayek, A., AbouRizk, S.M., and Boyd, B., 1998. Implementation of automated site data collection with a medium-sized contractor. *Proceedings of the International Computing Congress on Computing in Civil Engineering*, pp. 454–457. Boston (Massachusetts, USA)

Fernando, S., Mohamed, Y., AbouRizk, S.M., and Ruwanpura, J.Y., 2003. A review of simulation applications for varying demands in tunneling. *Proceedings of the ASCE Construction Research Congress*, Honolulu (Hawaii, USA).

FIATECH., 2009. *Element 9: Lifecycle Data Management & Information Integration.* Available from: http://fiatech.org/lifecycle-data-management-information-integration.html?showall=1 [Accessed Feb. 22nd, 2010]

Froehlich, G., Hoover, J., Liu, L., and Sorenson, P., 1998. Designing Object-Oriented Frameworks. *CRC Handbook of Object Technology*.

Fujimoto, R.M., 2003. Distributed simulation systems. *Proceedings of the Winter Simulation Conference*, pp. 124-134. New Orleans (LA, USA)

Hajjar, D., and AbouRizk, S.M., 2002. Unified Modeling Methodology for Construction Simulation. *Journal of Construction Engineering and Management*, 128(2) , 74-185.

Halpin, D.W., 1977. CYCLONE: Method for Modeling of Job Site Processes. Journal of the Construction Division, ASCE, 103(3) , 489-499.

Hendrickson, C., and Au, T., 1989. *Project Management for Construction*. Pittsburgh, PA: Prentice Hall.

IEEE Std 1516, *2000. IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) – Framework and Rules*

IEEE Std 1516.1, *2000. IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) - Federate Interface Specification*

IEEE Std 1516.2, *2000. IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) – Object Model Template (OMT) Specification*

Ioannou, P.G., 1989. *UM-CYCLONE User's Guide*. Ann Arbor, Michigan, Department of Civil Engineering, The University of Michigan.

Martinez, J., and Ioannou, P.G., 1994. General Purpose Simulation with Stroboscope. *Proceedings of the 26th Winter Simulation Conference*, pp. 1159-1166, Orlando (Florida, USA)

Marzok, M., Abdallah, M., and El-Said, M., 2008. Tunnel_Sim: Decision Support Tool for Planning Tunnel Construction Using Computer Simulation. *Proceedings of the 2008 Winter Simulation Conference*, pp. 2504-2511. Miami (Florida, USA)

Mazerolle, M., and Alkass, S., 1993. An Integrated System to Facilitate the Analysis of Construction Claims. *The Fifth International Conference on Computing in Civil and Building Engineering*, pp. 1509-1516. Anaheim (California, USA)

Moghani, E., Salehi, M., and Taghaddos, H., 2009. Simulation Based Schedule Enhancement of Tower Cranes. *Proceedings of the 2nd International/ 8th Construction Special Conference*. May 27-30, St. John's (Newfoundland and Labrador, Canada)

Navon, R. 2005. Automated project performance control of construction projects. *Automation in Construction*, 14(4), 467-476.

Paulson, B.C., Chan, W.T., and Koo, C.C., 1987. Construction Operation Simulation by Microcomputer. *Journal of Construction Engineering and Management*, 113(2) , 302-314.

Sawhney, A., 1994., *Simulation-based planning for construction.* Thesis (PhD). University of Alberta.

POINTWISE COMPLEATNESES AND POINTWISE DEGENERACY OF STANDARD AND POSITIVE HYBRID LINEAR SYSTEMS DESCRIBED BY THE GENERAL MODEL

Tadeusz Kaczorek

Faculty of Electrical Engineering
Bialystok University of Technology
Bialystok, Poland

kaczorek@isep.pw.edu.pl

ABSTRACT
Necessary and sufficient conditions for the pointwise completeness and pointwise degeneracy of the standard and positive hybrid linear systems described by the general model are established. It is shown that the standard general model is always pointwise complete and it is not pointwise degenerated and the positive general model is pointwise complete if and only if its matrix A_2 is diagonal.

Keywords: pointwise completeness, degeneracy, standard, positive, general model, hybrid system

1. INTRODUCTION

In positive systems inputs, state variables and outputs take only non-negative values. Examples of positive systems are industrial processes involving chemical reactors, heat exchangers and distillation columns, storage systems, compartmental systems, water and atmospheric pollution models. A variety of models having positive linear behavior can be found in engineering, management science, economics, social sciences, biology and medicine, etc.

Positive linear systems are defined on cones and not on linear spaces. Therefore, the theory of positive systems is more complicated and less advanced. An overview of state of the art in positive systems theory is given in the monographs (Farina and Rinaldi 2000; Kaczorek 2002).

The most popular models of two-dimensional (2D) linear systems are the discrete models introduced by Roesser (1975), Fornasini and Marchesini (1976, 1978), and Kurek (1985). The models have been extended for positive systems. An overview of positive 2D system theory has been given in the monograph (Kaczorek 2002).

A dynamical system described by homogenous equation is called pointwise complete if every given final state of the system can be reached by suitable choice of its initial state. A system which is not pointwise complete, is called pointwise degenerated.

The pointwise completeness and pointwise degeneracy belong to the basic concepts of the modern control theory of 2D linear systems and they play important role specially in positive 2D linear systems.

The pointwise completeness and pointwise degeneracy of linear continuous-time system with delays have been investigated in (Olbrot 1972; Popov 1972), of discrete-time and continuous-time systems of fractional order in (Busłowicz 2008; Kaczorek and Busłowicz 2009) and of positive discrete-time systems with delays in (Busłowicz, Kociszewski and Trzasko 2006; Choundhury 1972). The pointwise completeness of linear discrete-time cone-systems with delays has been analyzed in (Popov 1972). The pointwise completeness and pointwise degeneracy of standard and positive linear systems with state-feedbacks have been investigated in (Kaczorek 2010a; Kaczorek 2009).

The pointwise completeness and pointwise degeneracy of 2D standard and positive Fornasini-Marchesini models have been addressed in (Kaczorek 2010b).

Positive 2D hybrid linear systems have been introduced in (Kaczorek 1998; Kaczorek 2007; Kaczorek 2002) and positive fractional 2D hybrid linear systems in (Kaczorek 2007). Comparison of different method of solution to 2D linear hybrid systems has been given in (Kaczorek, Marchenko and Sajewski 2008). Realization problem for positive 2D hybrid systems has been addressed in (Kaczorek 2008b).

In this paper the pointwise completeness and pointwise degeneracy of standard and positive hybrid linear systems described by the general model will be addressed.

The structure of the paper is the following. In section 2 the pointwise completeness and the pointwise degeneracy of the standard general model is investigated. Necessary and sufficient conditions for the positivity and the pointwise completeness, pointwise degeneracy of the general model are established in section 3. Concluding remarks are given in section 4.

In the paper the following notation will be used. The set of $n \times m$ real matrices will be denoted by $\Re^{n\times m}$ and $\Re^n = \Re^{n\times 1}$. The set of $n \times m$ real matrices with nonnegative entries will be denoted by $\Re_+^{n\times m}$ and $\Re_+^n = \Re_+^{n\times 1}$. The $n \times n$ identity matrix will be denoted by I_n and the transpose will be denoted by T.

2. POINTWISE COMPLETENESS AND POINTWISE DEGENERACY OF STANDARD GENERAL MODEL

Consider the autonomous general model

$$\dot{x}(t,i+1) = A_0 x(t,i) + A_1 \dot{x}(t,i) + A_2 x(t,i+1)$$
$$t \in \Re_+ = [0,+\infty],\ \ i \in Z_+ = \{0,1,...\} \tag{2.1}$$

where $\dot{x}(t,i) = \dfrac{\partial x(t,i)}{\partial t}$, $x(t,i) \in \Re^n$, $u(t,i) \in \Re^m$, $y(t,i) \in \Re^p$ are the state, input and output vectors.
Boundary conditions for (2.1) are given by

$$x_1(0,i) = x_i,\ i \in Z_+ \text{ and } x(t,0) = x_{t0}, \dot{x}(t,0) = x_{t1},\ t \in \Re_+ \tag{2.2}$$

Definition 2.1. The general model (2.1) is called pointwise complete at the point (t_f, q) if for every final state $x_f \in \Re^n$ there exist boundary conditions (2.2) such that $x(t_f, q) = x_f$.

Theorem 2.1. The general model (2.1) is always pointwise complete at the point (t_f, q) for any $t_f > 0$ and $q = 1$.

Proof. From (2.1) for $i = 0$ we have

$$\dot{x}(t,1) = A_2 x(t,1) + F(t,0) \tag{2.3}$$

where

$$F(t,0) = A_0 x(t,0) + A_1 \dot{x}(t,0) = A_0 x_{t0} + A_1 x_{t1}. \tag{2.4}$$

Assuming $x_{t0} = 0$, $x_{t1} = 0$ we obtain $F(t,0) = 0$ and from (2.3)

$$x(t,1) = e^{A_2 t} x(0,1). \tag{2.5}$$

Substituting $t = t_f$ and $q = 1$ we obtain

$$x_f = e^{A_2 t_f} x(0,1) \tag{2.6}$$

and

$$x(0,1) = e^{-A_2 t_f} x_f. \tag{2.7}$$

Therefore, for any final state x_f there exist boundary conditions $x_{t0} = 0$, $x_{t1} = 0$ and $x_1 = e^{-A_2 t_f} x_f$ such that $x(t_f, 1) = x_f$ since the matrix $e^{-A_2 t_f}$ exists for any matrix A_2 and any $t_f > 0$. □

From theorem 2.1 we have the following corollaries.

Corollary 2.1. Any general model (2.1) is pointwise complete at the point (t_f, 1) for arbitrary $t_f > 0$.

Corollary 2.2. The pointwise completeness of the general model at the point (t_f, 1) is independent of the matrices A_0 and A_1 of the model.

Definition 2.2. The general model (2.1) is called pointwise degenerated at the point (t_f, q) in the direction v if there exist a nonzero vector $v \in \Re^n$ such that for all boundary conditions (2.2) the solution of the model for $t = t_f$, $i = q$ satisfies the condition $v^T x(t_f, q) = 0$.

Theorem 2.2. The general model (2.1) is not pointwise degenerated at the point (t_f, 1) for any $t_f > 0$.

Proof. Using the solution of (2.3)

$$x(t,1) = e^{A_2 t} x(0,1) + \int_0^t e^{A_2(t-\tau)} F(\tau,0) d\tau \tag{2.8a}$$

we obtain

$$v^T x(t,1) = v^T e^{A_2 t} x(0,1) + \int_0^t v^T e^{A_2(t-\tau)} F(\tau,0) d\tau \tag{2.8b}$$

where $F(t,0)$ is defined by (2.4). From (2.8b) it follows that does not exist a non-zero vector $v \in \Re^n$ such that for all boundary conditions (2.2) $v^T x(t_f, 1) = 0$ since the matrix $e^{A_2 t_f}$ is nonsingular for every matrix A_2 and $t_f > 0$. □

Example 2.1. Consider the general model (2.1) with the matrices

$$A_0 = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix},\quad A_1 = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix},\quad A_2 = \begin{bmatrix} -1 & 0 \\ 1 & -2 \end{bmatrix}. \tag{2.9}$$

Find the boundary conditions (2.2) at the point $(t_f, q) = (1,1)$ for $x_f = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

Taking into account that the eigenvalues of A_2 are $\lambda_1 = -1$, $\lambda_2 = -2$ and using the Sylvester formula we obtain

$$\begin{aligned} e^{-A_2 t} &= \frac{A_2 - \lambda_2 I_n}{\lambda_1 - \lambda_2} e^{-\lambda_1 t} + \frac{A_2 - \lambda_1 I_n}{\lambda_2 - \lambda_1} e^{-\lambda_2 t} \\ &= \begin{bmatrix} e^t & 0 \\ e^t - e^{2t} & e^{2t} \end{bmatrix} \end{aligned} \tag{2.10}$$

from (2.7) we have the desired boundary conditions

$$\begin{aligned} x(0,1) = e^{-A_2 t_f} x_f &= \left. \begin{bmatrix} e^{t_f} & 0 \\ e^{t_f} - e^{2t_f} & e^{2t_f} \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right|_{t_f = 1} \\ &= \begin{bmatrix} 2e \\ 2e + e^2 \end{bmatrix} \end{aligned} \tag{2.11}$$

and $x(t,0) = 0$, $\dot{x}(t,0) = 0$, $t \geq 0$.

The above conditions can be extended as follows.

From (2.1) for $i = 1$ we have

$$\dot{x}(t,2) = A_2 x(t,2) + F(t,1) \tag{2.12}$$

where

$$F(t,1) = A_0 x(t,1) + A_1 \dot{x}(t,1)\,. \qquad (2.13)$$

Substitution of (2.3) and (2.5) for $F(t,0) = 0$ into (2.13) yields

$$F(t,1) = (A_0 + A_1 A_2) x(t,1) = (A_0 + A_1 A_2) e^{A_2 t} x(0,1)\,. \quad (2.14)$$

Assuming $x(0,1) = 0$ we obtain $F(t,1) = 0$ and from (2.12)

$$x(t,2) = e^{A_2 t} x(0,2) \qquad (2.15)$$

Continuing this procedure for $i = 2,\dots,q-1$ we obtain the following theorem, which is an extension of Theorem 2.1.

Theorem 2.3. The general model (2.1) is always pointwise complete at the point (t_f, q), $t_f > 0$, $q \in N = \{1,2,...\}$ for any matrices A_k, $k = 0,1,2$.

Theorem 2.2 can be also extended for any point (t_f, q).

3. POINTWISE DEGENERACY AND POINTWISE DEGENERACY OF THE POSITIVE GENERAL MODEL

Definition 3.1. The model (2.1) is called positive if $x(t,i) \in \Re_+^n$, $t \in \Re_+$, $i \in Z_+$ for any boundary conditions

$$x_{t0} \in \Re_+^n\,,\ x_{t1} \in \Re_+^n\,,\ t \in \Re_+\,,\ x_i \in \Re_+^n\,,\ i \in Z_+ \quad (3.1)$$

Theorem 3.1. The general model (2.1) is positive if and only if

$$A_2 \in M_n \qquad (3.2a)$$

$$A_0, A_1 \in \Re_+^{n\times n},\ \ A = A_0 + A_1 A_2 \in \Re_+^{n\times n} \qquad (3.2b)$$

where M_n is the set of $n \times n$ Metzler matrices (with nonnegative off-diagonal entries).

Proof. Necessity. Necessity of $A_0 \in \Re_+^{n\times n}$ and $A_1 \in \Re_+^{n\times n}$ follows immediately from (2.4) since $F(t,0) \in \Re_+^n$, $t \in \Re_+$ and x_{t0}, x_{t1} are arbitrary. From (2.5) it follows that $A_2 \in M_n$ since $e^{A_2 t} \in \Re_+^{n\times n}$ only if A_2 is a Metzler matrix, $x(t,1) \in \Re_+^n$, $t \in \Re_+$ and $x(0,1)$ is arbitrary. From (2.12) it follows that $F(t,1) \in \Re_+^n$, $t \in \Re_+$ for any $x(0,1) \in \Re_+^n$ only if $A = A_0 + A_1 A_2 \in \Re_+^{n\times n}$. The proof of sufficiency is similar to the one given in (Kaczorek 2002: pp.255). □

Definition 3.2. The positive general model (2.1) is called pointwise complete at the point (t_f, q) if for every final state $x_f \in \Re_+^n$ there exist boundary conditions (3.1) such that

$$x(t_f, q) = x_f\,,\ t_f > 0\,,\ q \in N = \{1,2,...\}\,. \qquad (3.3)$$

It is assumed that $x_{t0} = 0$ and $x_{t1} = 0$ for $t \in \Re_+$.

Theorem 3.2. The positive general model (2.1) is pointwise complete at the point (t_f, 1) if and only if the matrix A_2 is diagonal.

Proof. In a similar way as in proof of Theorem 2.1 we may obtain the equation (2.7). It is well-known (Kaczorek 2002) that $e^{A_2 t} \in \Re_+^{n\times n}$, $t \in \Re_+$ if and only if A_2 is a Metzler matrix. Hence $e^{-A_2 t_f} \in \Re_+^{n\times n}$ if and only if A_2 is a diagonal matrix. In this case for arbitrary $x_f \in \Re_+^n$ if and only if $x(0,1) \in \Re_+^n$. □

In a similar way as for standard general model we can prove the following theorem.

Theorem 3.3. The positive general model (2.1) is pointwise complete at the point (t_f, q) $t_f > 0$, $q \in N = \{1,2,...\}$ if and only if the matrix A_2 is diagonal.

From Theorem 3.3 we have the following corollary.

Corollary 3.1. The pointwise completeness of the positive general model (2.1) is independent of the matrices A_0 and A_1 of the model.

Definition 3.3. The positive general model (2.1) is called pointwise degenerated at the point (t_f, q) if there exists at least one final state $x_f \in \Re_+^n$ such that $x(t_f, q) \neq x_f$ for all $x(0,i) \in \Re_+^n$ and $x(t,0) = 0$, $\dot{x}(t,0) = 0$, $t \in \Re_+$.

Theorem 3.4. The positive general model (2.1) is pointwise degenerated at the point (t_f, q) if the matrix $A_2 \in M_n$ is not diagonal.

Proof. In a similar way as in proof of Theorem 2.1 we may obtain the equality (2.7) which can be satisfied for $x_f \in \Re_+^n$ and $x(0,1) \in \Re_+^n$ if and only if the Metzler matrix A_2 is diagonal. The proof for $q > 1$ is similar. □

These considerations can be easily extended for $x(0,i) \in \Re_+^n$, $x(t,0) \in \Re_+^n$ and $\dot{x}(t,0) \in \Re_+^n$, $t \in \Re_+$.

Example 3.1. Consider the general model (2.1) with the matrices

$$A_0 = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix},\quad A_1 = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix},\quad A_2 = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}. \quad (3.4)$$

The model is positive since the matrices A_0 and A_1 have nonnegative entries and

$$\begin{aligned} A &= A_0 + A_1 A_2 \\ &= \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \in \Re_+^{2\times 2} \end{aligned} \quad (3.5)$$

The matrix A_2 is diagonal and the positive model with (3.4) by Theorem 3.2 is pointwise complete at the point (t_f, 1), $t_f \geq 0$. Using (2.7) we obtain

$$x(0,1) = e^{-A_2 t_f} x_f = \begin{bmatrix} e^{t_f} & 0 \\ 0 & e^{2t_f} \end{bmatrix} x_f \in \Re_+^2 \quad (3.6)$$

for any $x_f \in \Re_+^2$ and $t_f \in \Re_+$.

Example 3.2. Consider the general model (2.1) with the matrices (2.9). The model is positive since A_0 and A_1 have nonnegative entries, $A_2 \in M_n$ and

$$A = A_0 + A_1 A_2 = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \in \Re_+^{2\times 2}. \quad (3.7)$$

Let $x_f = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Using (2.10) and (2.7) we obtain the vector

$$x(0,1) = e^{-A_2 t_f} x_f = \begin{bmatrix} e^{t_f} & 0 \\ e^{t_f} - e^{2t_f} & e^{2t_f} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} e^{t_f} \\ e^{t_f} - e^{2t_f} \end{bmatrix} \quad (3.8)$$

with nonnegative second component for $t_f > 0$. Therefore, the model is pointwise degenerated at the point (t_f, 1). The same result follows from Theorem 3.4 since the matrix A_2 is not diagonal. Note that the vector $x(0,1)$ given by (2.11) for $x_f = [2 \quad 3]^T$ has positive components.

4. CONCLUDING REMARKS

The pointwise completeness and pointwise degeneracy of the standard and positive hybrid linear systems described by the general model have been addressed. Necessary and sufficient conditions for the pointwise completeness and pointwise degeneracy have been established. It has been shown that the standard general model (2.1) is always pointwise complete at the point (t_f, q) for any t_f and $q \geq 1$ and it is not pointwise degenerated at any point. Necessary and sufficient conditions for the positivity of general model (2.1) have been established. The positive general model is pointwise complete at the point (t_f, q) for $t > 0$, $q \geq 1$ if and only if the Metzler matrix A_2 of the model is diagonal. The considerations have been illustrated by numerical examples. These considerations can be extended for linear hybrid systems with delays. An extension of these considerations for fractional hybrid systems is an open problem.

ACKNOWLEDGMENTS

This work was supported by Ministry of Science and Higher Education in Poland under work No NN514 1939 33.

REFERENCES

Busłowicz M., 2008. *Pointwise completeness and pointwise degeneracy of linear discrete-time systems of fractional order*. Zesz. Nauk. Pol. Śląskiej, Automatyka, No. 151: pp. 19-24 (in Polish).

Busłowicz M., Kociszewski R., Trzasko W., 2006. *Pointwise completeness and pointwise degeneracy of positive discrete-time systems with delays*. Zesz. Nauk. Pol. Śląskiej, Automatyka, No. 145: pp. 55-56 (in Polish).

Choundhury A. K., 1972. *Necessary and sufficient conditions of pointwise completeness of linear time-invariant delay-differential systems*. Int. J. Control, Vol. 16, No. 6: pp. 1083-1100.

Farina L., and Rinaldi S., 2000. *Positive Linear Systems; Theory and Applications*. New York: J. Wiley.

Fornasini E., and Marchesini G., 1976. *State-space realization theory of two-dimensional filters*. IEEE Trans. Autom. Contr., AC-21: pp. 484-491.

Fornasini E., and Marchesini G., 1978. *Double indexed dynamical systems*. Math. Sys. Theory, Vol.12: pp. 59-72.

Kaczorek T., 1985. *Two-Dimensional Linear Systems*. Berlin: Springer Verlag.

Kaczorek T., 2002. *Positive 1D and 2D systems*. London: Springer-Verlag.

Kaczorek T., 2010a. *Pointwise completeness and pointwise degeneracy of standard and positive linear systems with state-feedbacks*. JAMRIS, Vol. 4, No. 1: pp. 3-7.

Kaczorek T., 2009. *Pointwise completeness and pointwise degeneracy of standard and positive fractional linear systems with state-feedbacks*. Archives of Control Sciences, Vol. 19: pp. 295-306.

Kaczorek T., 2010b. *Pointwise completeness and pointwise degeneracy of 2D standard and positive Fornasini-Marchesini models*. COMPEL (submitted)

Kaczorek T., 1998. *Reachability and minimum energy control of positive 2D continuous-discrete systems*. Bull. Pol. Ac. Techn. Sci., Vol. 46, No. 1: pp. 85-93.

Kaczorek T., 2007. *Positive 2D hybrid linear systems*. Bull. Pol. Ac. Techn. Sci., Vol. 55, No. 4: pp. 351-358.

Kaczorek T., 2008a. *Positive fractional 2D hybrid linear systems*. Bull. Pol. Ac. Techn. Sci., Vol. 56, No. 3: pp. 273-277.

Kaczorek T., 2008b. *Realization problem for positive 2D hybrid systems*. COMPEL, Vol. 2, No.3: pp. 613-623.

Kaczorek T., and Busłowicz M., 2009. *Pointwise completeness and pointwise degeneracy of linear continuous-time fractional order systems*. Journal of Automation, Mobile Robotics & Intelligent Systems, Vol. 3, No. 1: pp. 8-11.

Kaczorek T., Marchenko V., Sajewski Ł., 2008. *Solvability of 2D hybrid liner systems-Comparison of three different methods*, Acta Mechanica et Automatica. Vol. 2, No. 2: pp. 59-65.

Kurek J., 1985. *The general state-space model for a two-dimensional linear digital systems*. IEEE Trans. Autom. Contr. AC-30: pp. 600-602.

Olbrot A., 1972. *On degeneracy and related problems for linear constant time-lag systems*. Ricerche di Automatica, Vol. 3, No. 3: pp. 203-220.

Popov V. M., 1972. *Pointwise degeneracy of linear time-invariant delay-differential equations*. Journal of Diff. Equation, Vol. 11: pp. 541-561.

Roesser R. P., 1975. *A discrete state-space model for linear image processing*. IEEE Trans. On Automatic Control, AC-20, 1: pp. 1-10.

Trzasko W., Busłowicz M., Kaczorek T., 2007. *Pointwise completeness of discrete-time cone-systems with delays*. Proc. EUROCON Warsaw: pp. 606-611.

Weiss L., 1970. *Controllability for various linear and nonlinear systems models*, Lecture Notes in Mathematics, Vol. 144, Seminar on Differential Equations and Dynamic System II, Springer, Berlin: pp. 250-262.

AUTHORS BIOGRAPHY

Tadeusz Kaczorek, born 27.04.1932 in Poland, received the MSc., PhD and DSc degrees from Electrical Engineering of Warsaw University of Technology in 1956, 1962 and 1964, respectively. In the period 1968 - 69 he was the dean of Electrical Engineering Faculty and in the period 1970 - 73 he was the prorektor of Warsaw University of Technology. Since 1971 he has been professor and since 1974 full professor at Warsaw University of Technology. In 1986 he was elected a corresponding member and in 1996 full member of Polish Academy of Sciences. In the period 1988 - 1991 he was the director of the Research Centre of Polish Academy of Sciences in Rome. In June 1999 he was elected the full member of the Academy of Engineering in Poland. In May 2004 he was elected the honorary member of the Hungarian Academy of Sciences. He was awarded by the University of Zielona Gora (2002) by the title doctor honoris causa, the Technical University of Lublin (2004), the Technical University of Szczecin (2004), Warsaw University of Technology (2004), Bialystok Technical University (2008), Lodz Technical University (2009) and Opole Technical University (2009).

His research interests cover the theory of systems and the automatic control systems theory, specially, singular multidimensional systems, positive multidimensional systems and singular positive 1D and 2D systems. He has initiated the research in the field of singular 2D, positive 2D linear systems and positive fractional 1D and 2D systems. He has published 23 books (6 in English) and over 940 scientific papers.

He supervised 69 Ph.D. theses. He is editor-in-chief of Bulletin of the Polish Academy of Sciences, Techn. Sciences and editorial member of about ten international journals.

MODELING AND SIMULATION OF PROCESS INTERRUPTIONS WITH COPULAS

Ivan Ourdev[(a)], Simaan AbouRizk[(b)]

[(a,b)]Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Alberta, Canada

[(a)]iourdev@ualberta.ca, [(b)]abourizk@ualberta.ca

ABSTRACT
Construction projects are exposed to various unexpected interruptions, with equipment breakdowns being the most common. Proper modeling of these interruptions along with the associated uncertainty can significantly reduce risk and improve project management. We employ a new mathematical approach, the copula method, to model the field observations in a tunnel excavation project. We characterize the excavation process interruptions with their degree of severity and frequency, and develop a Student t copula model for the underlying dependence structure. The model is consecutively utilized in a Monte Carlo simulation that incorporates all available information to forecast the project completion time. The adaptive estimates then serve as a basis for project management decisions. Our approach allows to incorporate the uncertainty within an intuitive simulation framework and to accurately model the dependence of the different dimensions of the process interruptions.

Keywords: copulas, uncertainty, process interruptions, Monte Carlo simulation, project management

1. INTRODUCTION

Uncertainty permeates real-life project management: uncertain durations, uncertain cost, sudden weather changes, equipment breakdown, human resource problems, unexpected changes in project scope, etc. Uncertainty is rarely beneficial and takes the form of a risk that must be dealt with. It threatens the bottom line and, particularly, the project schedule. Many project activities are sequential, and alterations to the duration of some tasks have a ripple effect on the start times of all subsequent tasks down the activity chain. Although a certain amount of contingency time is normally built into all project schedules, changes in the schedule have to be managed in a timely fashion in order to ensure a relatively smooth flow of labor and materials. Thus, the forecasting of task execution times becomes an essential ingredient of successful project risk management.

The common approach to decisions made under uncertainty relies on probability theory, where the quantities of interest are considered random variables (r. v.) described by probability distributions. The mean of the probability distribution specifies the expected value of the modeled quantity, while the standard deviation quantifies our uncertainty about the 'true' value of this mean. The most common distribution used by researchers and practitioners alike is the normal, also called Gaussian, distribution, which takes the familiar bell shape. The popularity of this distribution is due to its convenient mathematical properties. It is analytically tractable, completely described by only two parameters: the mean and the standard deviation. A linear combination of normal distributions is also a normal distribution with parameters determined by the means and the covariance matrix of the original components. Also, according to the central limit theorem the distribution of the sum of many independent r. v. with finite variances approaches normal distribution. However, in real life, normal distributions are exceptional, rather than the rule.

The standard approach to modeling data generated by a vector-valued random process is to fit a multivariate probability distribution, using e.g. the maximum likelihood (ML) method. The drawbacks of this approach are the lack of control in the fitting process and imprecise physical interpretation of the components of the distribution.

Recently, new mathematical objects, called copulas, have become very popular for multivariate modeling of dependent variables and risk management (Nelsen 2003, Yan 2006, Frees and Valdez 1998). Copulas are multivariate functions with uniform marginals that allow the construction of the joint distribution from the constituent marginals capturing the dependency structure of the latter (see e.g. Nelsen 2006). The intuitive copula approach to building multivariate distributions is a two-step statistical procedure. In the first step, the empirical marginals are obtained by fitting univariate distributions to the data. In the second step, an appropriately chosen copula is used to combine the univariate marginals into a joint distribution. It has been pointed out (Mikosch 2006) that the use of this approach is not universally justified, but for our purposes it has definite advantages.

The first advantage of the two-step methodology is that it enables meaningful interpretations of the marginal distributions, and it applies to and compares with existing, well-researched models. The second advantage is that, by choosing a specific copula, we can tailor the fitting process to the relative importance of the dependence domain.

In this work, we model the unexpected process interruptions in a tunnel excavation project. The model was included as a separate component of a much larger distributed decision support and planning system, based on discrete-event simulation. The purpose of the model was to serve as basis for Monte Carlo simulations, where the randomly generated interruptions of the excavation are taken into account for an adaptive project schedule planning.

Using copulas allowed us to borrow the familiar intuition from actuarial science, where the unexpected losses are characterized by two random variables: severity and frequency (Klugman 2004). In our case, these become the marginal distributions of the severity of the interruptions and the time intervals between interruptions. The separate choice of the copula allows the importance of the relatively rare but severe breakdowns to be stressed.

The paper is organized as follows: Section 2 introduces the notion of copulas and presents the main definitions and important properties. It also includes some broad examples of copulas and gives accounts of the methodology for statistical inference and simulation. This paper considers a case with two random variables, so for the sake of simplicity and notational clarity we only use bivariate copulas, but all the results presented are also valid in higher dimensions (see e.g. Nelsen 2006 for a general treatment). Section 3 contains an overview of the tunnel excavation operations and the data collection, particularly of process interruptions. The copula model of the excavation process interruptions and the results from the simulations are presented in Section 4, which also contains a brief description of the software framework that encompasses the model. The conclusion, Section 5 contains an evaluation of the approach and some suggestions for future research. Some well-known probability concepts are included in Appendix A to serve as an easy reference for comparing the properties of two-dimensional probability distributions and copulas.

2. COPULAS

2.1. Definitions for bivariate copulas

The notion of mathematical copulas was introduced by Abe Sklar in 1959 as functions that link n--dimensional distributions to their one--dimensional margins (Sklar 1959). Copulas are, in general, distribution functions that have as arguments $\mathbb{R}^m$ – valued random vectors $\mathbf{X} = \left(X_1, \cdots, X_m\right)$; for simplicity, we restrict this presentation to the two-dimensional case, $\mathbb{R}^2$. Formally, in the case of two r. v., X and Y, the (bivariate) copula, C, is a function $[0,1]^2 \to [0,1]$ that has the following properties:

- It is a grounded function:

$$C(u,0) = C(0,v) = 0, \quad \forall u,v \in [0,1] \tag{1}$$

- consistent with its margins:

$$C(u,1) = u, \quad C(1,v) = v, \quad \forall u,v \in [0,1] \tag{2}$$

- It has a non-negative C – volume, i.e. for every $u_1, u_2, v_1, v_2 \in [0,1]$, such that $u_1 \le u_2$, and $v_1 \le v_2$ the following inequality holds:

$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \ge 0 \tag{3}$$

This property requires copulas to be "2-increasing" functions, which is the two-dimensional analog of a nondecreasing function of one variable (see Nelsen 2006 for details).

The basis for the theory of copulas is Sklar's theorem, which states that for a two-dimensional joint cdf $F_{X,Y}(x,y)$ with marginal distributions $F_X(x)$, $F_Y(y)$ there exists a unique 2-copula such that:

$$F_{X,Y}(x,y) = C\left(F_X(x), F_Y(y)\right) = C(u,v) \tag{4}$$

If the random variables X and Y are continuous, then Equation 4 is unique. Otherwise, the copula is uniquely determined on the range $Ran(X) \times Ran(Y)$. Conversely, if C is a bivariate copula and $F_X(x)$, $F_Y(y)$ are distribution functions, then the function $F_{X,Y}(x,y)$ defined by Equation 4 represents a joint cdf. Thus, copula "couples" the marginals to form a joint cdf.

One of the methods for copula construction is to use Sklar's theorem and invert the expression of Equation 4 as

$$C(u,v) = F_{X,Y}\left(F_X^{-1}(u), F_Y^{-1}(v)\right) \tag{5}$$

For continuous r. v., copulas, as every ordinary joint cdf, have their corresponding densities, *c*, analogously to Equation A2:

$$c(u,v) = \frac{\partial C(u,v)}{\partial u\, \partial v} \tag{6}$$

and the bivariate joint pdf has the following canonical representation:

$$f_{X,Y}(x,y) = c\left(F_X(x), F_Y(y)\right) f_X(x) f_Y(y) \tag{7}$$

This representation illustrates the decomposition of the multivariate probability density, $f_{X,Y}(x,y)$ into one-dimensional marginals, $f_X(x)$, $f_Y(y)$, and a dependence structure specified by the copula density, c.

2.2. Examples of copulas

Given that every multivariate distribution has a corresponding copula, the number of possible copulas is enormous. There are three copula families that have been found most useful: elliptical, Archimedean, and extreme-value copulas.

Elliptical copulas derive from elliptical distributions, with the two main representatives, Gaussian and Student's distributions. They are widely used for modeling financial time series, particularly in the context of factor models (Malevergne and Sornette 2005). The parameter that is needed for their specification is the correlation matrix in the multivariate case, or the correlation coefficient, ρ, in the bivariate. Figure 1 shows the copula densities for two elliptic copulas with the same correlation coefficient $\rho = .5$ a normal copula and a Student t copula with 3 degrees of freedom.

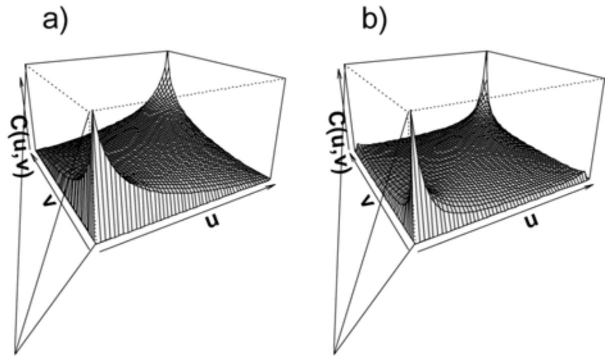


Figure 1: The Copula Densities for Two Elliptic Copulas with a Correlation Coefficient ρ = 0.5: (a) Normal Copula, and (b) Student t Copula with 3 Degrees of Freedom.

Student t copulas are particularly interesting because of their higher densities in the corners $(1,0)$, and $(0,1)$, as seen on Figure 1b. We illustrate the effect of the correlation coefficient on the density distribution of a Student copula that links two beta marginal distributions. Beta distribution is a flexible distribution with density

$$b(x;\alpha,\beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} \tag{8}$$

where α, and β are shape parameters and the normalization constant is the beta function, $B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt$.

Figure 2 shows an example of two marginal beta distributions with parameters $f_X(x;3,1.5)$, and $f_Y(y;2,5)$. The copula used to link these marginals is the Student t with two degrees of freedom. Figure 3 shows the contour plots for different values of the correlation coefficient, ρ.

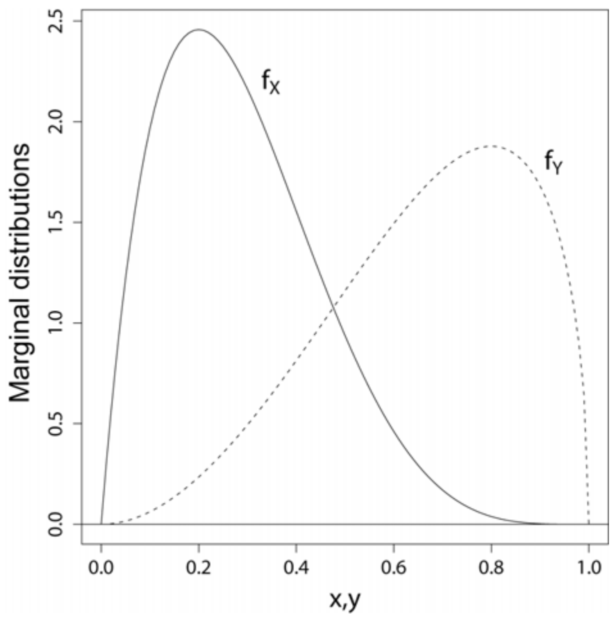


Figure 2: Two Marginal Beta Distributions with Parameters $f_X(x;3,1.5)$ and $f_Y(y;2,5)$

Archimedean copulas often arise in the context of the actuarial modeling of sources of risk (Frees and Valdez 1998). They are constructed without referring to the distribution functions. Instead, the construction is done using a continuous strictly decreasing function, $\varphi:[0,1]\to[0,\infty]$, called generator, such that $\varphi(1)=0$, and the formula:

$$C(u,v) = \varphi^{[-1]}\left(\varphi(u)+\varphi(v)\right) \tag{8}$$

Where $\varphi^{[-1]}$ is the pseudo-inverse of φ, defined as:

$$\varphi^{[-1]} = \begin{cases} \varphi(u), & \text{if} \quad 0 \le u \le \varphi(0), \\ 0, & \text{otherwise}, \end{cases} \tag{9}$$

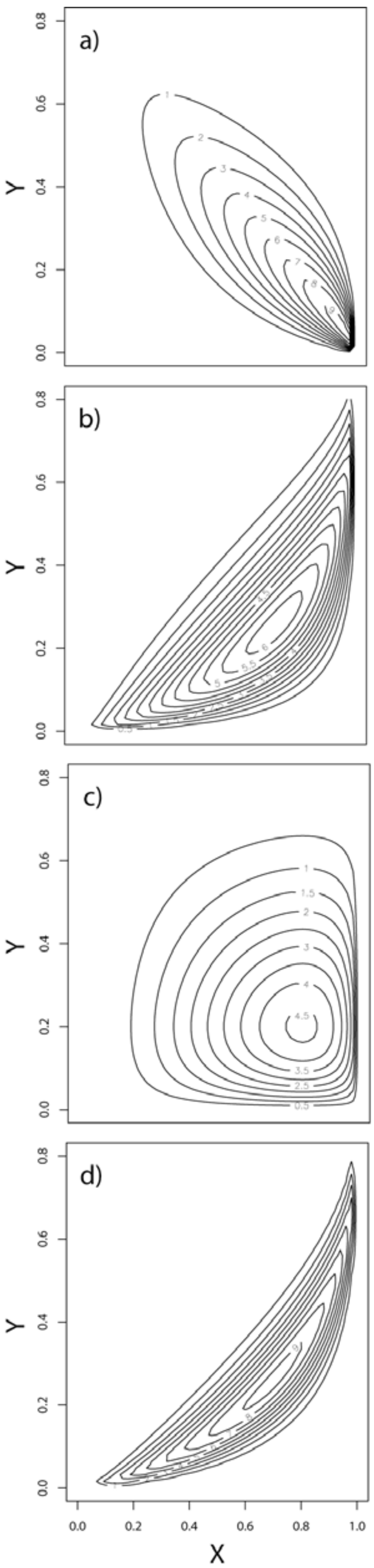


Figure 3: Contour Plot of Student t Copulas with Two Degrees of Freedom that Link the Beta Marginals from Figure 2 with Different Correlation Coeffieicents: a) $\rho = -.75$, b) $\rho = .75$, c) $\rho = 0$, and d) $\rho = .9$

Different generators give rise to different copulas. For example, the Clayton copula

$$C_\theta^{\mathrm{Cl}}(u,v) = \max\left(\left[u^{-\theta} + v^{-\theta} - 1\right]^{-1/\theta}, 0\right) \quad \forall\theta \in [-1,\infty), \tag{11}$$

is obtained from the generator:

$$\varphi_\theta^{\mathrm{Cl}}(u) = \frac{1}{\theta}\max\left(u^{-\theta} - 1\right) \tag{12}$$

Figure 4 shows the contour plots of the Clayton copula with parameter $\theta = 1$ and its density. It is clear that Clayton copula has a heavy lower tail, the $(0,0)$– corner. This copula is very important in the multivariate statistics of extremes, because it can be shown that it is the limiting copula for the class of the Archimedean copulas when the probability level of the quantiles approaches zero.

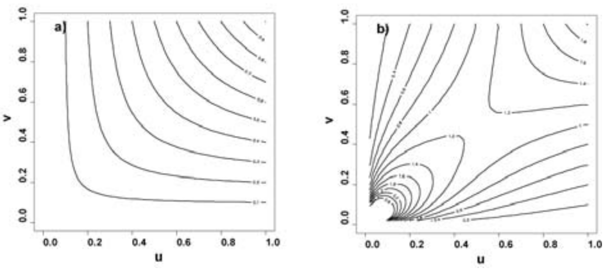


Figure 1: Contour Plots of the Clayton Copula with Parameter $\theta = 1$, Panel (a), and its Density, Panel (b).

Another popular example is the Gumbel copula

$$C_\theta^{\mathrm{G}}(u,v) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{-1/\theta}\right) \quad \forall\theta \in [1,\infty), \tag{13}$$

which is obtained from the generator:

$$\varphi_\theta^{\mathrm{G}}(u) = (-\ln u)^\theta \tag{14}$$

The Gumbel copula has a density that peaks at the $(1,1)$ corner and is, in a sense, complementary to Clayton copula which density is the highest at the $(0,0)$ corner. The Gumbel copula is also an example of extreme-value copulas, which are derived from generalized extreme value distributions.

2.3. Statistical inference and simulation

The majority of the copula estimation approaches rely on the maximum likelihood technique. ML can be applied either to the joint estimation of the parameters of the marginals and the copula, or the two can be treated separately. We follow the latter approach, because it gives a better control on the estimation process. The method, called inference function for margin (IFM) (Joe and Xu 1996), is a two-step procedure: first, the marginals are fitted to the data, and then the copula is estimated conditionally on the fitted marginals. Both the fitting of the marginals and the copula involve a choice of the appropriate distributions. Fitting univariate distributions to data is a well-studied problem (see e. g. Joe and Xu 1996). The identification of the appropriate copula is still largely empirical. Currently, there is a non-parametric identification methodology only for the Archimedean copula class (Genest and Rivest 1993, Wang and Wells 2000).

The best fit for the marginals was found to be the gamma distributions. The probability density function of the gamma distribution, as parameterized by the shape parameter, α, and the rate parameter, λ, is given by

$$g(x;\alpha,\lambda) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \tag{15}$$

where the normalization constant is the gamma function, $\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-\alpha} dt$. The expression $X \sim Gamma(\alpha_X, \lambda_X)$ is used to signify that the random variable X has a gamma distribution with the corresponding shape and rate parameters.

The fitting of the marginal distributions is done in two steps. First, using the method of the moments we obtain rough estimates for the shape parameter, $\hat{\alpha}$, and the rate parameter, $\hat{\lambda}$. Then we use the moment estimates as a starting point for the maximum-likelihood estimation step.

The method of moments is a well-known technique for obtaining parameter estimates. The construction is done by matching the sample moments, m_i, with the corresponding distribution moments μ_i, and solving for the latter. In the case of the gamma distribution only the first two sample moments are needed, $m_1 = \bar{x}$, and $m_2 = (n-1)s^2/2$, where $\bar{x}$ is the usual sample mean, s^2 is the sample variance and n is the sample size. The corresponding (central) distribution moments are defined as $\mu_i = E\left(X^i\right)$ and for the gamma density, Equation 15, the integrations give a mean $\mu_1 = \mu = \alpha / \lambda$, and a variance $\mu_2 = \sigma^2 = \alpha(\alpha+1)/\lambda^2$. The matching step yields the following starting estimates:

$$\hat{\lambda}_0 = \frac{n\bar{x}}{(n-1)s^2 - n\bar{x}^2}, \qquad \hat{\alpha}_0 = \frac{n\bar{x}^2}{(n-1)s^2 - n\bar{x}^2}. \tag{16}$$

These initial estimates are used as starting points for the ML step, which maximizes the sample likelihood.

The IFM method for a two-parameter marginal distribution functions, as in our case, involves a clear separation of the marginal parameters $\boldsymbol{\beta} = (\alpha, \lambda)$ from the association parameters $\boldsymbol{\theta}$. The likelihood function for n independent observations $\{x_k : k = 1, \cdots, n\}$ and a density distribution g is defined as

$$L(x_1, \cdots, x_n; \boldsymbol{\beta}) = \prod_{k=1}^{n} g(x_k; \boldsymbol{\beta}) \tag{17}$$

Substitution of the gamma density, Equation 15, in this expression yields the following form of the log-likelihood function:

$$l(\boldsymbol{\beta}) = n\alpha \log \lambda - n \log \Gamma(\alpha) + (\alpha+1) \sum_{k=1}^{n} \log x_k - \lambda \sum_{k=1}^{n} x_k \tag{18}$$

For our bivariate case, there are n pairs of observations of interruptions $\{(x_k, y_k) : k = 1, \cdots, n\}$ with severity, x_k, occurring at intervals, y_k. We also need to introduce additional index, M, that enumerates the two marginals, $M = X, Y$. Thus the IFM estimates for the parameters $\hat{\boldsymbol{\beta}}_M$ of two gamma distributions with densities, g_X, and g_Y, become

$$\hat{\boldsymbol{\beta}}_M = \arg\max_{\boldsymbol{\beta}_M} \sum_{k=1}^{n} \log g_M\left(m_k; \boldsymbol{\beta}_M\right), \quad M = X, Y \tag{19}$$

The second step finds the IFM estimate of the association parameter, $\boldsymbol{\theta}$, using the copula density, c, as

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \sum_{k=1}^{n} \log c\left(F_X\left(x_k; \boldsymbol{\beta}_X\right), F_Y\left(y_k; \boldsymbol{\beta}_Y\right)\right) \tag{20}$$

The ML estimates, obtained above form the foundation of Monte Carlo simulations (Fishman 2003, McLeish 2005). The simulation of process interruptions modeled by a specific copula uses Sklar's theorem. The approach relies on some algorithm (e.g. Malevergne and Sornette 2005) for generation of two uniform random numbers u_X, and u_Y, on the interval $(0,1)$ with a dependence structure given by the copula, C. In order to generate two random variables, X, and Y, from the proper joint distribution $F_{X,Y}(x,y) = C\left(F_X^{-1}(x), F_Y^{-1}(y)\right)$, only the application of the generalized inverse is needed: $x = F_X^{-1}(u_X)$, and $y = F_Y^{-1}(u_Y)$.

We apply the steps for modeling and simulation described above to the data for excavation process interruptions.

3. DATA

The data consists of the durations of the delays and interruptions in the stage SW3 of the South Edmonton Sanitary Sewer (SESS) tunneling project in the City of Edmonton, Canada. The project involves the excavation of a 3.5 km long sanitary sewer tunnel using a tunnel boring machine (TBM). It started in February 2006 and was completed in August 2007. The tunneling operations are constantly monitored and the relevant data is recorded and collected by a decision support system, called COSYE.

The information about the daily operation of the TBM comes from two sources: one is an engineering survey system called TACS (tunnel advance control system), and the other is the report prepared at the end of the day. The daily report contains information about the number of work shifts per day, the length of the shifts in hours, and the source and the duration of the project delays and interruptions.

3.1. TBM operations

Tunnel construction by means of tunnel boring machines is considered to be a state-of-the-art technology. The main TBM element is a cylindrical rotating cutterhead with a diameter approximately equal to that of the tunnel that bores in the earth strata. The support for the forward press is provided by gripper shoes that engage outwardly with the tunnel wall. The support for the wall of the tunnel is provided by one-meter long cement rings that are placed as the tunnel is being dug. Each ring consist of two semi-circular segments, called liners. The front part of the tunnel, where the actual excavation takes place is called the tunnel face.

The SW3 tunnel has a relatively small diameter, 2.34 m, which to a large extent determines the tunneling operations. It has a single-track railway for most of the tunnel length, which becomes a double-track only in the area close to the entrance shaft. Still, in order to save time on loading and unloading operations, two trains carry loads between the face of the tunnel and the entrance shaft.

The excavation is a batch process with activities naturally partitioned into cycles. The beginning of a cycle is marked by the unloading of the liners from the train. The unloaded train is positioned behind the TBM and excavation begins. The carts of the train collect the dirt from the excavation. After the one-meter length is excavated, the train, loaded with dirt, starts traveling back towards the entrance shaft, while the TBM begins the installation of the liner blocks. The loaded train dumps the dirt into a sump pocket, while the first train, already loaded with liner blocks, starts traveling towards the face of the tunnel. The crane hoists the dirt from the sump pocket to the surface, where it is stockpiled. Afterward, the crane lowers down the liners blocks for the next segment of the tunnel. This completes one cycle of tunnel operations.

3.2. Excavation interruptions

Tunnel construction is a process of several interdependent activities, placing the equipment under a significant strain. Machine breakdowns and system malfunctions are common and often result in interruptions of the whole chain of operations. In our approach, we disregard the specific source of interruption. No distinction between a breakdown in the excavation process and an interruption in some of the support operations is made. The system of tunneling operations is modeled as a whole. The main reason for this approach is that we do not have enough data to model the elements separately. Another reason is the high degree of coupling (correlation) between the system elements. The general system approach solves both problems.

Thus, the only assumption we make is that the characteristics of the system will remain practically constant for the duration of the project until completion. For example, no new equipment will be introduced, or the load on the existing equipment will remain the same. The soil composition profile of the site, which is one of the main determinants of the load on the TBM, has little variation as inferred on the basis of the exploratory borehole samples. Also, experience indicates little effect due to seasonal changes.

Figure 5 shows the interruptions that occurred between September 14, 2006 and May 10, 2007. The x-axis represents the time line measured in work shift operating hours. The breakdown's severity is measured in terms of the work shift time it takes to fix the problem. The excavation operations take place during work-shifts. Normally there is one 10-hour shift per day, with weekends off, but depending on the overall progress of the project or external events, the project manager can decide on splitting the work into unusual 8-hour shifts, one or two per work-day. Only the time during which the equipment is in operation contributes to the probability of breakdowns, thus only work shift time is taken into account. The frequency of the interruptions is quantified by the interval between two successive breakdowns. The severity of the break is measured in terms of the work shift time it takes to fix the problems and restart excavation.

4. MODEL

4.1. The COSYE system

The model of the process interruptions was embedded in the general simulation and decision support system COSYE (Construction Synthetic Environment). Details of the system are given elsewhere (AbouRizk and Mohammed 2000). Here we only include a very brief general description.

The COSYE simulation environment is a .NET implementation of the HLA (High Level Architecture) IEEE standard for modeling and simulation (SISC 2000). The HLA architecture is a general framework for creating complex distributed simulations from relatively independent simulation units called federates. It has two main elements: the federate interface specification (FIS), and the object model template (OMT). FIS specifies the communication interface for combining the individual simulation components and maintaining the interoperability between them, while OMT describes the exchanged data. The model execution is provided by a run time infrastructure (RTI) server.

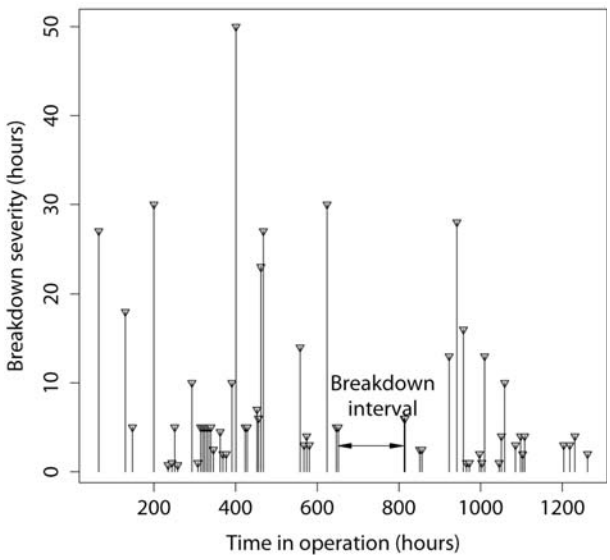


Figure 5: Excavation Interruptions Occurring between September 14, 2006 and May 10, 2007

The COSYE architecture for the simulation of the tunnel boring operations is comprised of several federates. One federate simulates the operations at the face of the tunnel, which include the excavation and the installation of the liners; another federate simulates the creation of tunnel sections; a third one handles the motion of the trains and the crane operations, etc. (see Ourdev et al. 2007) for details). The unexpected interruptions due to equipment failures and breakdowns are included in the breakdown federate, which implements the model described below. Figure 6 shows a plot of the breakdown characteristics, tracking the severity of the breakdown by the number of hours needed for repair and the time interval between breakdowns. The top histogram represents the marginal distribution of the severity, and the histogram on the right represents the marginal of the time interval.

4.2. Copula interruptions simulation

As pointed out in Section 2, the inference function for margin method involves two steps: first, fitting the marginals to the data, and then estimating the copula conditionally on the fitted marginals. The first step for finding the best fit for the marginals is to calculate the starting point for the numerical procedure. The data consist of $n = 58$ observations. The sample mean of the breakdown severity is $\overline{x}_S = 8.8$ hours and its variance is $\sigma_S^2 = 93.84$ hours. Substitution of these values into Equation 4 yields the following starting estimates for the parameters of the marginal distribution of the severity of the breakdowns: shape parameter, $\hat{\lambda}_{S0} = 2.419$, and rate parameter $\hat{\alpha}_{S0} = .299$. Similarly, we calculate the sample mean of the breakdown interval is $\overline{x}_F = 21.76$ hours and its variance is $\sigma_F^2 = 800.85$ hours. The substitution of these values into Equation 4 yields the following starting estimates for the parameters of the marginal distribution of the interval between the breakdowns: shape parameter, $\hat{\lambda}_{F0} = 1.510$, and rate parameter $\hat{\alpha}_{F0} = 0.069$.

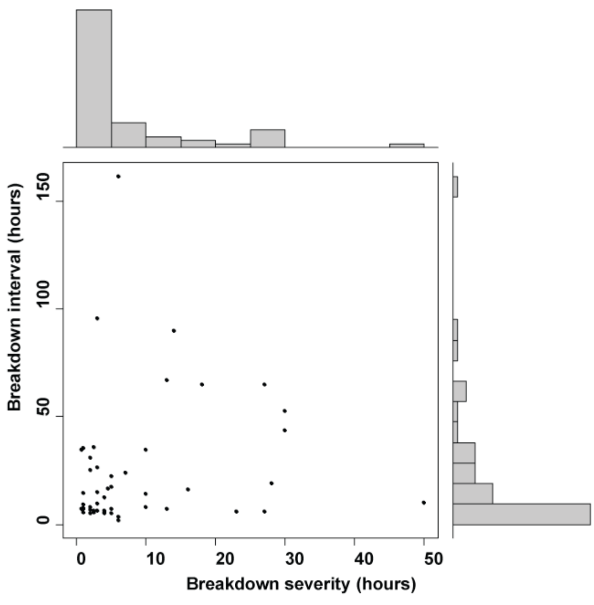


Figure 6: Plot of Breakdown Characteristics

Using the above initial estimates as starting points for the MLE procedure, we find that the best fit for the severity is given by a gamma distribution with parameters $S \sim Gamma(0.987, 0.153)$. The standard errors of the estimated parameters are: $SE(\hat{\alpha}_S) = 0.177$, and $SE(\hat{\lambda}_S) = 0.028$. For the frequency of the breakdowns we find another Gamma distribution with parameters $F \sim Gamma(1.1, 0.056)$. The standard error of the estimated parameters of this distribution are: $SE(\hat{\alpha}_F) = 0.177$, and $SE(\hat{\lambda}_F) = 0.010$. Figure 7 shows the histograms of the breakdown severity, and the intervals between breakdowns, with the corresponding probability densities of the best fits. Similarly, Figure 8 shows the empirical cumulative distribution functions (ecdf) of the breakdown severity, and the intervals between breakdowns, with the corresponding cumulative distributions of the best fits.

We used the Kolmogorov-Smirnov (KS) test to ascertain formally the goodness of fit of the above distributions. KS test quantifies the difference between the ecdf and the theoretical cdf, as shown in Figure 8 to formulate a hypothesis testing. The H_0 hypothesis is that the data comes from the specified distribution, versus the alternative, H_a, that the data is not from that distribution. The test statistics are formulated as the greatest difference between the ecdf, $F_{emp}(x)$, and the hypothesized theoretical cdf, $F_h(x)$

$$t = \sup_X \left| F_{emp}(x) - F_h(x) \right| \qquad (21)$$

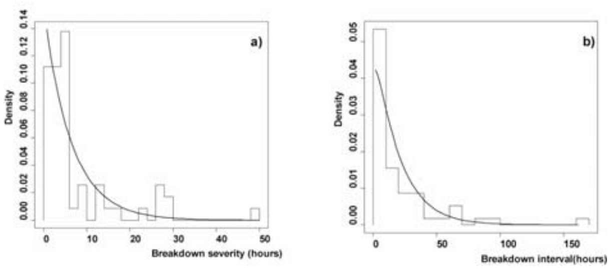


Figure 7: Histograms of the Breakdown Severity, Panel (a), and the Intervals Between Breakdowns, Panel (b), with the Corresponding Probability Density of the Best Fits

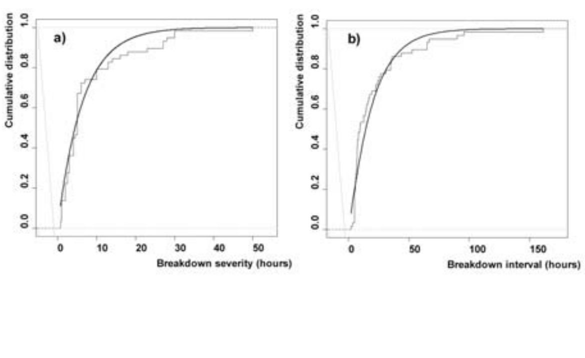


Figure 8: Empirical Cumulative Distributions of the Breakdown Severity, Panel (a), and the Intervals Between Breakdowns, panel (b), with the Corresponding Cumulative Distribution of Best Fits

The null hypothesis is rejected if the test statistics is greater than some critical value, or, alternatively, if the p-value is below the significance level.

The calculation for the KS test yields a p-value of $p_S = 0.268$ for the severity of the breakdowns, and a p-value of $p_F = 0.071$ for their frequency, so for both cases we cannot reject the null hypothesis at the significance level of $\alpha = 5\%$.

Having the best fits for the marginals, next, we proceed to find the best fit copula. Based on the observed two-dimensional data distributions, Figure 6, and the general considerations outlined above, we searched for a copula from the Student t copula class. We use the estimates of the shape parameters for the severity and frequency of the interruptions, $\hat{\alpha}_S$, and $\hat{\alpha}_F$, and the corresponding rate parameters, $\hat{\lambda}_S$, and $\hat{\lambda}_F$, obtained as described above as starting points of the maximum likelihood method.

The result of the MLE procedure brings a slight modification for the parameter of the marginals. The gamma distribution parameters for the severity become $S \sim Gamma(1.078, 0.132)$ with standard errors of $SE(\hat{\alpha}_S) = 0.176$, and $SE(\hat{\lambda}_S) = 0.028$. For the gamma distribution parameters for the frequency we find $F \sim Gamma(1.075, 0.048)$ with standard errors of $SE(\hat{\alpha}_F) = 0.176$, and $SE(\hat{\lambda}_F) = 0.010$. The correlation is estimated as $\rho = 0.267$ with a standard error of $SE(\hat{\rho}) = 0.135$. The contour plot for the resulting multivariate distribution is presented in Figure 9.

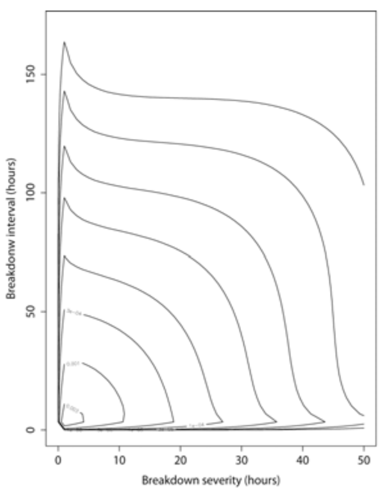


Figure 9: Contour Plot for the Multivariate Distribution Function Obtained as a Best Fit

With these estimated parameters, we can draw random samples from the interruptions distribution, to be used for the Monte Carlo simulation. The dots on Figure 10 represent 350 simulated values for process interruptions with severity and frequency obtained from the best fit to the observed interruptions as described above. The triangles on the figure visualize the actual observed to that moment interruptions and serve as additional check for the goodness of fit.
The procedure outlined above, involving the steps of fitting to the data and the Monte Carlo simulation, can be repeated every time a new interruption is registered by the system. Thus, the quality of the statistical fit will improve with the increase of the available data point. Such an online algorithm also allows the model to adapt to changes in the environment, such as equipment wear or variation among excavated strata.

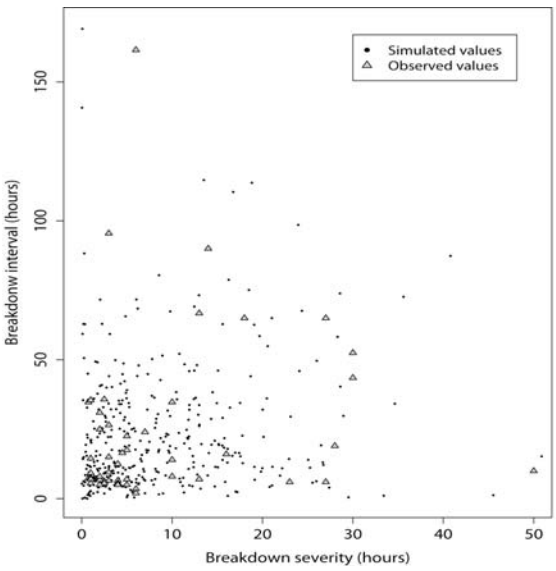


Figure 10: Simulated Values for Process Interrruptions with Severity and Frequency Obtained from the Best Fit to the Observed Interruption

5. CONCLUSIONS

In this paper we presented the first, to our knowledge, application of copulas to a construction project. We modeled the process interruptions that occur during a tunnel excavation by applying a system approach. In order to retain enough data for a meaningful statistical inference, we considered the operations comprising the excavation process as a system with the interruptions as one of the characteristic variables.

Although copulas are inherently multivariate objects, we restricted ourselves to the two-dimensional case. This allowed us to preserve the interpretation of one of the marginal distributions as the severity of the interruptions and the other as the frequency of interruptions (respectively: the interval between the interruptions). We carefully fitted the marginal distributions and the copula to the available data. The resulting two-dimensional probability distribution was used to generate random samples for a Monte Carlo simulation. The simulation results were used to estimate the changes in the project schedule and for more accurate and adaptive project management.

Our results showed the power of the copula approach for modeling and simulating uncertain dependent variables. The ease with which the copulas fit into the framework of Monte Carlo indicate a much broader application area, which would include more adequate risk modeling and risk management that does not rely on the assumption of normality.

APPENDIX A

The purpose of this appendix is to serve as an easy reference for comparison between some properties of the two-dimensional probability distributions and the copulas. The probabilistic approach to modeling uncertain quantities is to treat them as *random variables* (r. v.). Random variables, generally speaking, consist of two parts: the expected (most often occurring) value of the variable, and a measure of how uncertain we are about this expected value. The natural representation of a r. v., X , discrete or continuous, is its *cumulative distribution function* (cdf), defined as:

$$F_X(x) = P(X \le x), \forall x \in \mathbb{R} \qquad \text{(A1)}$$

For continuous random variable there is an alternative presentation, the *probability density function* (pdf), defined as the first derivative of cdf, i.e.

$$f_X(x) = F'_X(x) \qquad \text{(A2)}$$

For a given pdf, $f_X(x)$ $f_X(x)$, the fundamental theorem of calculus allows calculating the corresponding cdf:

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du \qquad \text{(A3)}$$

It is very useful to introduce also the *quantile* function, which is the generalized inverse of the cdf, defined as:

$$F_i^{-1}(u) = \inf \left\{ x \mid F_i^{-1}(u) \ge x \right\}, \quad \forall u \in (0,1) \qquad \text{(A4)}$$

For strictly increasing F_i, the quantile function F_i^{-1} becomes the ordinary inverse.

For a *pair* r. v., (X,Y), the dependence is completely described by their *joint cdf* defined as:

$$F_{X,Y}(x,y) = P(X \le x, Y \le y), \quad \forall (x,y) \in \mathbb{R}^2 \tag{A5}$$

The equivalent probability model for the r. v. X and Y is given by their *joint pdf*, $f_{X,Y}(x,y)$ defined as:

$$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)\,du\,dv \tag{A6}$$

The relation between the joint cdf and the joint pdf, corresponding to Equation 12 is given by:

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x\,\partial y} \tag{A7}$$

Integration over one of the r. v. yields the *marginal distribution* of the other, e.g.

$$f_X(x) = \int_{-\infty}^{y} f_{X,Y}(u,v)\,dv \tag{A8}$$

Two random variables, *X* and *Y*, are *independent* if and only if

$$f_{X,Y}(x,y) = f_X(x) f_Y(y), \quad \forall (x,y) \in \mathbb{R}^2 \tag{A9}$$

The same proposition also holds in terms of cdfs, i.e.

$$F_{X,Y}(x,y) = F_X(x) F_Y(y), \quad \forall (x,y) \in \mathbb{R}^2 \tag{A10}$$

REFERENCES

AbouRizk, S. M. and Mohamed, Y. 2000. *Simphony*: an integrated environment for construction simulation. *Proc. 2000 Winter Simulation Conf.*, IEEE, Orlando, Fla., 1907-1914.

Fishman, G. 2003. *Monte Carlo*, 4th ed. Springer, Cambridge, Mass.

Frees, W. E. and Valdez, E. A. 1998. Understanding relationship using copulas. *N. Amer. Actuarial J.*, 2, 1-15.

Genest, C. and Rivest, L.-P. 1993. Statistical interference procedures for bivariate archimedean copulas. *J. Amer. Stat. Assoc.*, 88423, 1034-1043.

Joe, H. and Xu, J. J. 1996 The estimation method of inference functions for margins for multivariate models. *Technical Report 166.* University of British Columbia, Vancouver, B.C.

Klugman, S. A., Panjer, H. H., Willmot, G. E. 2004. *Loss Models: From Data to Decisions*, 2nd ed. Wiley-Interscience, Toronto, Ont.

Malevergne, Y. and Sornette, D. 2005. Notions of copulas. *Extreme Financial Risks: From Dependence to Risk Management.* Springer, Cambridge, Mass.

McLeish, D. L. 2005. *Monte Carlo Simulation and Finance*. Wiley, Toronto, Ont.

Mikosch, T. 2006. Copulas: Tales and facts. *Extremes*, 9, 3-20.

Nelsen, R. B. 2003. Properties and applications of copulas: a brief survey. *Proc. 1st Braz. Conf. on Stat. Modeling in Insurance and Finance*, University of São Paulo, São Paulo, Brazil, 10-28.

Nelsen, R. B. 2006. *An Introduction to Copulas*, 2nd ed., Springer, Cambridge, Mass.

Ourdev, I., AbouRizk, S. M., and Al-Bataineh, M. 2007. Simulation and uncertainty modeling of project schedules estimates. *Proc. 2007 Winter Simulation Conf.*, IEEE, Washington, D. C.

SISC Simulation Interoperability Standards Committee of the IEEE Computer Society. 2000. *IEEE Standard for Modeling and Simulation M & S High Level Architecture HLA: Framework and Rules*. *Std. 1516-2000*.

Sklar, A. 1959. Fonctions de repartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, 8, 229-231.

Wang, W. and Wells, M. 2000. Model selection and semiparametric inference for bivariate failure-time data. *J. Amer. Stat. Assoc.*, 95449, 62-72.

Yan, J. 2006. Multivariate modeling with copulas and engineering applications. *Handbook of Engineering Statistics*, H. Pham, ed., Springer, Cambridge, Mass., 973-990.

A study of computational modeling of abdominal aortic aneurysm

S.M. Kim[a], M.P Lee[b], H.C Kang[c] , J.H Lee[d] , J.H Jeong[e] , S.G Roh[f]

[a]Department of Medical Bio Engineering, College of Biosystem, Dongguk University-Seoul
[b] Department of Research Institute of Biotechnology, Dongguk University-Seoul
[c] Department of Medical Bio Engineering, College of Biosystem, Dongguk University-Seoul
[d] Department of Medical Bio Engineering, College of Biosystem, Dongguk University-Seoul
[e] Department of Medical Bio Engineering, College of Biosystem, Dongguk University-Seoul
[f] Department of Electronic Engineering, Konkuk University

[a]sungmin2009@gmail.com, [b]nonsame@gmail.com, [c] kang572k@gmail.com , [d] tupac1969@gmail.com,
[e]bfsl.jeong@gmail.com, [f]gyu1106@naver.com

ABSTRACT

Abdominal Aortic Aneurysm is caused by changes to the blood flow pattern due to the change of the elasticity and shape of the blood vessel. This type of aneurysm had been placed at the 10th highest cause of deaths of white males aged between 65 and 74 in the U.S.A, in the year of 2000.(National Center for Health Statics)

Abdominal Aortic Aneurysm is generated due to the unbalance of the enzyme which controls the aortic wall. The continuing expansion of the aortic wall drastically increases due to the mechanical quality change of the vessel. And ruptures occur, ultimately.

The cause of aneurysm has not been known exactly yet, but smoking is being given attention as the largest cause. And other causes are hypertension, family history and gender in statistically.

This study analyzes blood flow patterns and observed the movement changes according to the diameter of the abdominal aortic aneurysm. It was confirmed that as the expansion part gets larger, the blood flow focuses on the top of the expansion part. And the flow velocity at the bifurcation point which connects to the common iliac artery increases.

In this study, we tried to specify the process element of the abdominal aortic aneurysm based on the analysis results, that the increase of the expansion part due to the aneurysm was viewed as change of the blood flow.

Keywords: Velocity Profile, Blood Flow Patterns, Abdominal Aortic Aneurysm (AAA), Computational Fluid Dynamics

1. INTRODUCTION

The Aorta is the largest artery of human body, an important organ which carries the necessary elements such as oxygen and nutrients throughout the body in the cardiovascular system from the heart to the peripheral vessels.

The abdominal aorta is the last part of the aorta, following the vertebra in the chasm of the diaphragm to the front of 4th lumbar vertebra and branches to the iliac arteries. (*Medifocus Guidebook*, Abdominal Aortic Aneurysm, Medifocus.com, Inc., 2008).

Abdominal Aortic Aneurysm, which is one of major blood vessel diseases as a cause of death in the cardiovascular system, is a disease which is generated by the expansion of part of the abdominal aorta. If the vessel is ruptured, it can lead to death. According to the statistics published by U.S. National Center of Health Statistics in 2000, deaths by abdominal aortic aneurysms and artery incisions are 15,810, and the cause of the deaths was ranked at the 10th among the people aged between 65 and 74 (Anderson RN, 2002).

In general, the diameter of the abdominal aorta is 1.7-2.4cm, depending upon genders and ages. The abdominal aortic aneurysm is when a lesion is 50% larger than the normal diameter. Generally, it is abdominal aortic aneurysm of expansion size is 3cm or larger.

This study analyzes the blood flow patterns according to the change of the diameter of the abdominal aortic aneurysm, thus observing the flow in the expansion part through a symmetrical shape. And we obtain the flow trend which was depends on the diameters based on the result.

2. MATERIALS AND METHODS

2.1. Abdominal Aortic Aneurysm Size

Aneurysm can be formed anyplace in blood vessels, but mainly in the aorta. An aneurysm is an abnormal expansion of an artery with an increase of greater than 1.5 times the normal diameter. The aneurysm which is formed in the lower part of the renal artery is called the abdominal aortic aneurysm (Fig.1) (William H Pearce 2007).

The cause of the abdominal aortic aneurysm has not been exactly known, but various factors are suspected.

The strongest risk factor is smoking. About 90% of the patients with abdominal aortic aneurysm had smoked or were smoking. Smoking results in the increase of the process rate of aneurysm, and the rate is higher among males than females. Family history of abdominal aortic aneurysm is asked in the diagnosis, and determines its connection with the disease process (*Medifocus Guidebook*, Abdominal Aortic Aneurysm, Medifocus.com, Inc., 2008)(Frank A 1997). Other causes of the AAA are hypertension,

Table 1: Relationship of expansion size to rupture in 473 non-resected abdominal aortic aneurysm

Size(cm)	Ruptured	Unruptured	Total	% Ruptured
$\leq$ 5.0	34	231	265	12.8
> 5.0	78	116	194	40.0
No size recorded	6	8	14	43.0
Total	118	355	473	24.9

previous vascular surgery, chronic obstructive pulmonary disease, etc.

The size of the expansion is the priority factor for the determination of surgery. It is judged that as the expansion diameter increase, the possibility of rupture increase (*Medifocus Guidebook*, Abdominal Aortic Aneurysm, Medifocus.com, Inc., 2008).

The rupture risk based on the diameter of 5cm showed that 12.8% for up to 5cm, 40.0% for 5cm or larger, about a three times higher (Table 1) (Darling, R.C. 1977). The rupture risk is different depending upon situations, but it is confirmed that if the size is 5cm or larger, the risk increases rapidly. The relationship between the size of AAA and rupture is reported every year based on clinical cases, and there are various engineering approaches to it.

In this study, we have gradually increased maximum diameter of the abdominal aortic aneurysm, from 17mm to 64mm, based on standard model (Fig. 2) Blood vessel model was designed as a cylindrical form, and incremental ratio were applied as a maximum diameter of each model (Fig. 3)

2.2. Blood Flow in Abdominal Aorta

Aneurysm can be formed anyplace in blood vessels, It was assumed that the blood is an incompressible newtonian fluid, and the blood density which changes according to changes of pressure.

Figure 1: Abdominal Aortic Aneurysm CT image

.

Table 2: Parameters of the abdominal aortic aneurysm

Parameters	Abdominal Aorta	Iliac Artery
Top Diameter	1.7cm	1.7cm
Bottom Diameter	1.7cm	1.1cm
Length	11.0cm	5.5 / 5.7cm (R / L)
Bifurcation Angle	-	52.45°
Center of Curvature	-	2.9 / 3.9cm (R / L)

The viscosity (μ) was determined as 0.003696 kg/m·s, and the normal state was assumed to solve the Navier-Stokes Equations.

The blood flow in the aorta goes in a certain shape, and is circulated to the iliac artery in a time-dependent similar shape. A certain wave form of blood flow is formed due to the difference between the pressures of the contraction and the release period. A physiologic wave form was composed for the analysis of the flows (Mette S. Olufsen 2000). To measure the mass flow by using MRI and ultrasound waves, studies assumed the relationship between section area and flow speed (Eq. 1).

$$Q_i(t) = A_i(t) \cdot V_i(t) \qquad \text{(Eq. 1)}$$

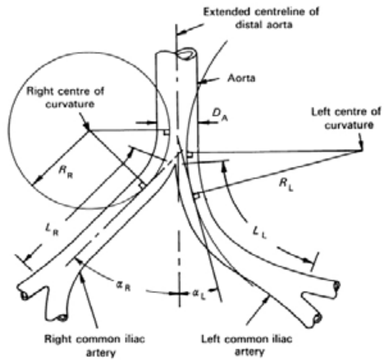


Figure 2: Standard model of abdominal aortic aneurysm

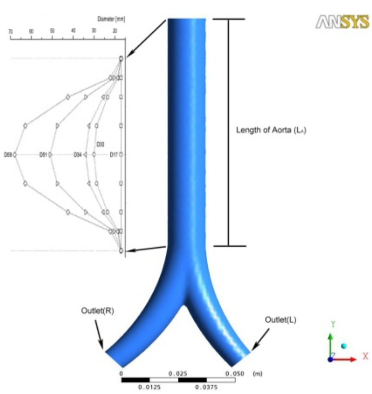


Figure 3: Diameter of each models

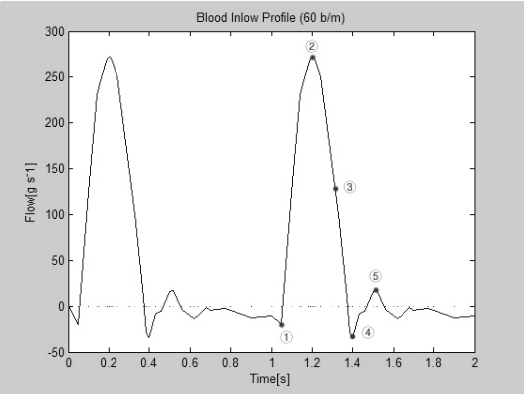


Figure 4: Adapted Blood Inflow Profile

①:1.05s(50ms), ②:1.20s(200ms), ③:1.30s(300ms)

④:1.40s(400ms), ⑤:1.52s(520ms)

The Mass Flow Profile, obtained from the model of the blood vessel tree structure, was applied to the movement model (Mette S. Olufsen 2000) (Fig. 4).

Fig. 5 shows variance of the blood flow dependent to time on abdominal aortic aneurysm, and the similar results were derived from the blood inflow profile as shown in Fig. 4 (Vieli A. 2000)

The average flow was 78.169 g/s, the period was 1.0s, and the heart rate was 60.

To suppose the continuing period of the input condition in simulation, two wave forms were applied, and the data of the only second one were used in the result.

2.3. Simulation of AAA

The geometry of the abdominal aorta based the values measured from 100 cadavers as the normal model (P.M.Shah 1978)(Irving Geller 1961), and each parameters are shown in Table 2.

In the model of the normal abdominal aorta without an expansion part, five models of from 1.7 up to 6.4cm were designed with symmetrical expansions

The diameter of expansion was determined as 3.0, 4.2, 5.3, 6.4cm. The initial 3.0cm is the ordinary size of an aneurysm. An aneurysm with a diameter of 6.4cm, about 3.8 times larger than the normal vessel model, was designed as the model of the greatest rupture risk.

3. RESULTS

3.1. Sectional Flow of the Abdominal Aortic Aneurysms

In this study, the geometry presented a symmetric model. To observe the flow movement, each sections were divided (Fig. 6-10). The upper part and lower part of Line 2 are divided as follows.

At 200ms, the flow draws a stream line according to the shape of the abdominal aorta. At the greatest diameter of expansion increases, the gaps of the steam

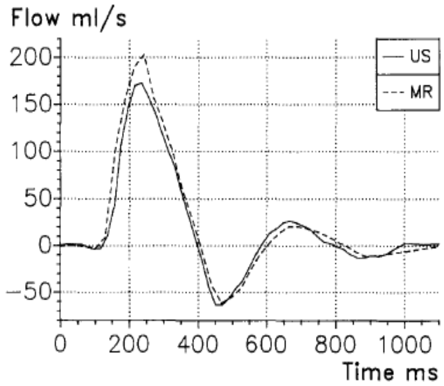


Figure 5: Flow curve (Vieli A. 2000)

tracking lines change. But other vortices are not observable in the aneurysms.

At 400ms, vortices are divided into the upper and lower part of Line 2. The greatest diameter of the expansion becomes 42mm or larger, the upper vortices were concentrated around the upper part of aneurysm. And as the diameter increases, the concentration of vortices becomes severer. The lower vortices can be formed due to bifurcation that is branches of iliac artery.

The bifurcation angle of both iliac arteries shows a little difference between the right and left, and the vortices generated thereby occur relatively more in the left iliac artery. Vortices, if severe, ascend to Line 2.

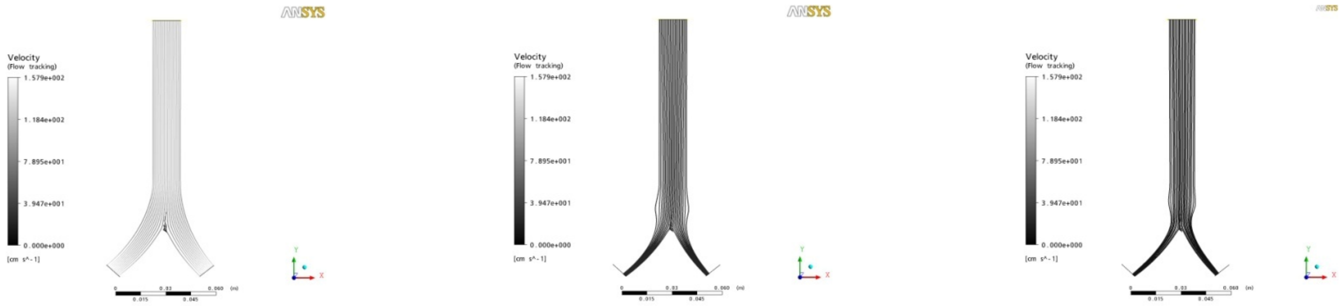
In each figure, at 200ms (a), the blood flows running the shape of the abdominal aorta, and at 400ms (b), vortices occur in the upper and lower parts. As the diameter increases, the vortices concentrate to the upper part.

In each figure, the lower vortices are disappeared at 520ms (c), and flows concentrate around the center of the bifurcation. The upper vortices show a concentrating more.

3.2. Sectional Flow Upper and Lower Line 2

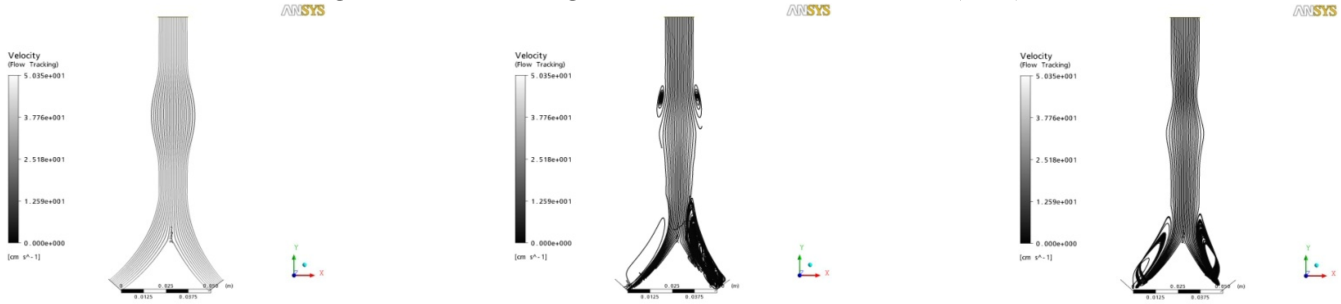
The flow in the lower part of Line 2 changes very fast as time passes, concentrated more to the right common iliac artery than the left. The flow at the bifurcation point runs in a different stream line as dispersing in both directions.

In the size of the expansion part increases, the flow under line 2 relatively shows a higher velocity in the left iliac artery than in the right. The shape of the flow at the bifurcation point affects the blood flow velocity entering into the iliac artery according to the size of the expansion of the abdominal aortic aneurysm. As the expansion size increases, the maximum flow velocity in the contraction period in the lower part of Line 2 increases (Fig. 11) (A. Vieli 1989).



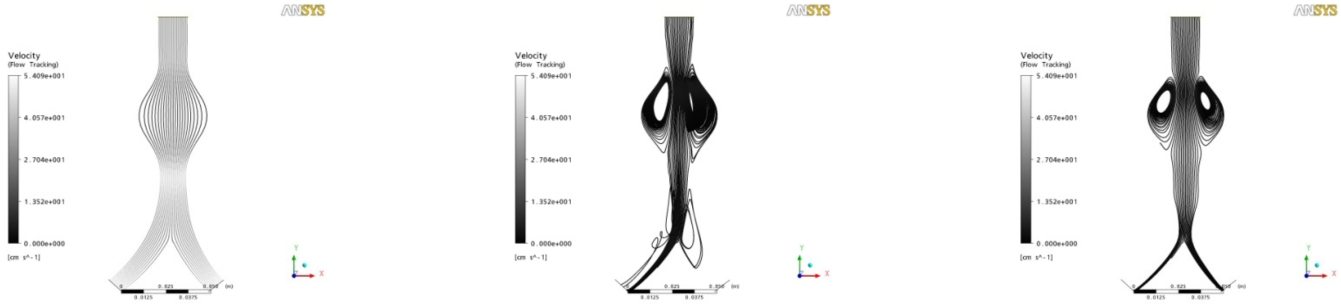
(a) 200ms (b) 400ms (c) 520ms

Figure 6: Flow Tracking Line of Normal Abdominal Aorta (D:17)



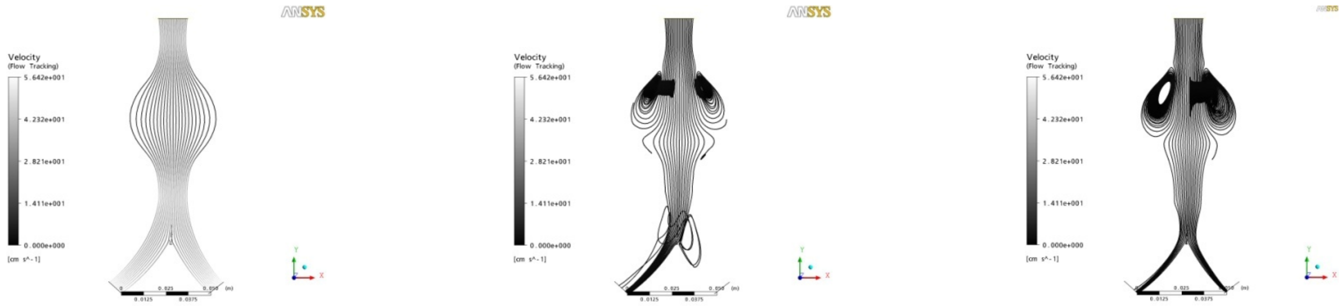
(a) 200ms (b) 400ms (c) 520ms

Figure 7: Flow Tracking Line of Normal Abdominal Aorta (D:30)



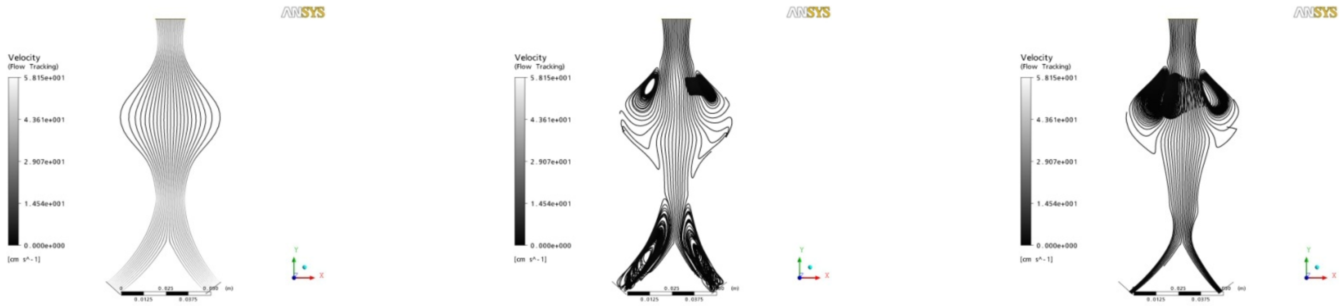
(a) 200ms (b) 400ms (c) 520ms

Figure 8: Flow Tracking Line of Normal Abdominal Aorta (D:42)



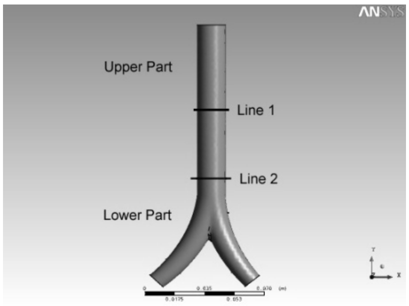
(a) 200ms (b) 400ms (c) 520ms

Figure 9: Flow Tracking Line of Normal Abdominal Aorta (D:53)

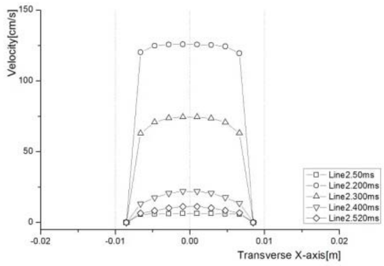


(a) 200ms (b) 400ms (c) 520ms

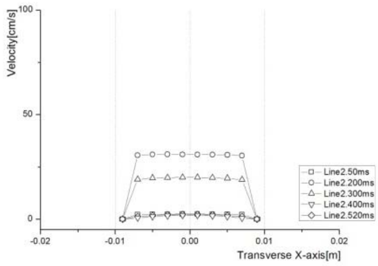
Figure 10: Flow Tracking Line of Normal Abdominal Aorta (D:64)



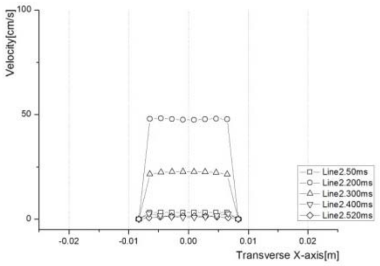
(a) Line 2 for midline velocity profile



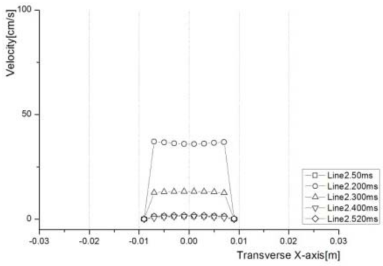
(b) at diameter 17mm



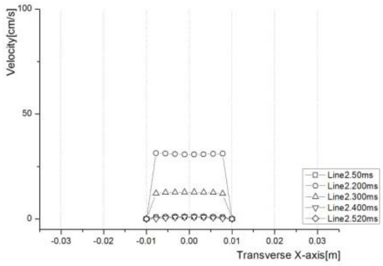
(c) at diameter 30mm



(d) at diameter 42mm



(e) at diameter 53mm



(f) at diameter 64mm

Figure 11: Velocity profile at line 2

4. CONCLUSION

When the blood flows in a vessel, the flow gets loss due to friction with the vessel wall, the change due to the size, shape, and direction of the section, and also due to changes of the entrance, exit and internal shape.

This study did not take into account thrombosis, plaque, vessel elasticity, vessel vibration by waves, etc., but observed only the flow as time passes in each section at expansions of different sizes.

The flow depended on time, the velocity and stream line changed in this study, are important factors for observing energy losses.

The flow stream in the upper and lower parts of Line 2 has great correlations with the expansion size and vessel section area of the abdominal aorta. The velocity change at a certain point becomes smaller as the vessel section area becomes larger.

This study formed symmetrical models to analyze the flow pattern of the abdominal aorta. And we observed the velocity change. As a result, the diameter change of the abdominal aortic aneurysm has great effect on flow, and as the diameter increases, vortices concentrate to the upper part. It is judged that a velocity change by initial vortices of an aneurysm can cause the expansion of the aneurysm.

5. DISCUSSION

In this study, we looked for the factors that related the process of an aneurysm by the flow of the upper expansion parts. The phenomenon that the velocity changes and vortices by diameter concentrated in the upper part is one of the factors which generate the process of an aneurysm in the symmetric geometry.

To realize a flow pattern in the aneurismal space, we used symmetric models. But the actual abdominal aortic aneurysm processes in an anterior asymmetrical shape due to the spinal cord. However, according to previous studies (David A. Vorp 2007)(Zhonghua Li 2007), flow concentrated in the upper part in the asymmetrical model. It is similar to us.

The change to the neck angel and thrombosis were found to be another factors to the process of the abdominal aortic aneurysm. In this study, the flow pattern by diameter change can predict the directional characteristic of the abdominal aortic aneurysm which processes through the abnormal shape.

For actual analysis, based on the analysis of the flow pattern by using symmetrical models, we are conducting an analysis of the flow patterns of the abdominal aortic aneurysm realized by CT, thereby specifying the process factor of the abdominal aortic aneurysm based on the shape like this.

ACKNOWLEDGMENTS

This study was supported by Dongguk university-Seoul 2009 research fund

REFERENCES

National Center for Health Statistics[http://www.cdc.gov], " Mortality Tables"

Medifocus Guidebook, Abdominal Aortic Aneurysm, Medifocus.com, Inc., 2008

Anderson RN, 2002. Death: leading causes for 2000. *National Vital Statistics Reports(NSVR)*, 50, 1-85.

William H Pearce, M.D., Brian G Peterson, M.D., 2007. Abdominal Aortic Aneurysm, *emedicine from WebMD*

Frank A. Lederle, Gary R. Johnson, Samuel E. Wilson, Edmund P. Chute, Fred N. Littooy, Dennis Bandyk, illam C. Krupski, Gary W. Barone, Charles W. Acher and David J. Ballard, 1997. Prevalence and Associations of Abdominal Aortic Aneurysm Detected through Screening. *Annals of Internal Medicine*, 126(6), 441-449.

Darling, R.C., Messina, C.R., Brewster, D.C., Ottinger, L.W., 1977. Autopsy study of unoperated abdominal aortic aneurysms. *Circulation,* 56(suppl.3), 161-164.

Mette S. Olufsen, Charles S. Peskin, Won Yong Kim, Erik M. Pedersen, Ali Nadim and Jesper Larsen, 2000. Numerical Simulation and Experimental Validation of Blood Flow in Arteries with Structured-Tree Outflow Conditions. *Annals of Biomedical Engineering*, 28(11), 1281-1299.

Vieli A., Moser U., Maier S., Meier D., Boesiger P., 2000. Velocity Profile in the Normal Human Abdominal Aorta: A comparison Between Ultrasound and Magnetic Resonance Data. *Utrasound in Medicine & Biology*, 15(2), 119-133

P.M.Shah, H.A.Scarton and M.J.Tsapogas, 1978. Geometric Anatomy of the Aortic-Common Iliac Bifurcation. *Journal of Anatomy*, 126(3), 451-458.

Irving Geller, Russell T. Woodburne,1961. Surgical Anatomy of the Abdominal Aorta. *Annals of Surgery,* 154(suppl.6), 239-252.

A. Vieli, U. Moser, S. Maier, D. Meier and P. Boesiger,1989. Velocity Profile in the Normal Human Abdominal Aorta: A Comparison Between Ultrasound and Magnetic Resonance Data. *Ultrasound in Medicine & Biology*, 15(2), 133-119.

David A. Vorp, 2007. Biomechanics of Abdominal Aortic Aneurysm. *Journal of Biomechanics*, *40*(9), 1887-1902.

Zhonghua Li, Clement Kleinstreuer, 2007. A Comparison between Different Asymmetric Abdominal Aortic Aneurysm Morphologies Employing Computational Fluid-Structure Interaction Analysis. *European Journal of Mechanics B/Fluid*, 26(5), 615-631.

Designing irregular gears via splines

Piotr Krawiec [(a)], Adam Marlewski [(b)]

[(a)] Chair of Basics of Machine Design, Poznań University of Technology, 60-965 Poznań, Poland
[(b)] Institute of Mathematics, Poznań University of Technology, Poland, 60-965 Poznań, Poland

[(a)] piotr.krawiec@put.poznan.pl, [(b)] adam.marlewski@put.poznan.pl

ABSTRACT. To construct a new machine, to improve the existing one or to reconstruct the worn up, damaged machine/tool, we need its theoretical model, in both functional and material issues, and the means to produce its final form. Contemporarily, there are used computers in all the stages of the production process, as well as in so-called inverse tasks, when we want to reconstruct (the parts of) machines. It demands still more and more precise mathematical description. In this paper we discuss such description via spline curves (composed of third degree polynomial arcs represented parametrically) and we apply it to gears/pulleys of non-typical irregular profiles.

Keywords: spline, noncircular wheel, reverse engineering, computer aided design

1. MOTIVATION AND INTRODUCTION

The usage of noncircular gears and pulleys makes possible to have better characteristics of the transmission, e.g., the changeable kinematical features in the gear ratio and the velocity. The required degree of speed variability is obtained by the use of pulleys constructed with wheel rims having shapes of ellipses, ovals or other non-circular disks. There is already well recognized the usage of chains in such drives. On the other hand, the usage of toothed belts in such transmission systems is not described in details. That is due to the different kinematics and coupling characteristics of toothed belts and pulleys as compared with chain drives. In a chain drive the driven strand can be slack, whereas in case of a toothed belt drive it must be tight. It makes that there must be met two conditions. The first condition concerns the length of the belt: it must be equal to the length of the envelope. In order to ensure the correct operation of a variable-speed transmission system, the active and passive sections of the belt must be tightened by an appropriate constant force. The second condition of the correct work of a variable-speed transmission system is to ensure the cyclycity of its movement. This cyclicity is absolutely required during the circular motion of machine elements. The circumference of every wheel has to be an integer multiple of the pitch (as always, a metric pitch is the distance between neighboring belt teeth based on millimeters). Thus, one is able to determine the average transmission ratio of the system as the relation of circumferences of wheels or the number of their teeth. The cyclicity of drives can be guaranteed only by toothed belts having following property: during the operation the plastic strain varies slightly. At the same time the belts must be initially pre-tightened in order to avoid the slip of the belt as well as the skip on the teeth of the wheels.

If both conditions, a constant tension of the belt and the cyclicity condition, are met simultaneously, one is able to search for a design of a variable-speed transmission system as that shown in Fig.1.

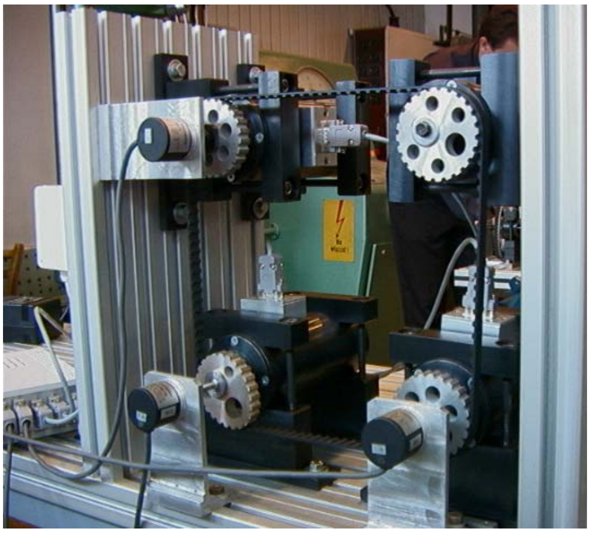


Fig. 1. Belt and pulley transmission systems installed in LEDM (Laboratory for Experimental Design and Manufacturing at Faculty of Machines and Transport, Poznań University of Technology); the system has two circular wheels, one noncircular wheel and one eccentrically mounted wheel

The problem at hand is widely studied in the case of regular noncircularity, e.g., when there are applied elliptical or trochoidal wheels, both types being well-described mathematically. In this study we deal with irregular gears in case when their profiles are not covered by well-known mathematical equations. The design of such gears involves more advanced techniques in both technological and mathematical aspects, in the last case it is often needed a numerical treatment.

The design of noncircular gears and pulleys working in tooth transmissions is widely described in literature, see, e.g., Laczik (2003), Danieli et al. (2005), Ming-Feng Tsay and Zhang-Hua Fong (2005), Bair et al. (2007), JianGang Li et al. (2007), Bair (2009). In most cases there are considered regular noncircular wheels, i.e., the gears of the elliptical and cycloidal profile; only such plane curves are discussed in the book by Litvin and Fuentes (2004), a bible for gear designers. In last years there are undertaken practical experiments and theoretical considerations concerning non-typical irregularly shaped elements of belt/chain drivers. As far as we know, the literature dedicated to such non-typical toothing is rather modest, it is treated in, e.g., Li Xin Cao et al. (2002), Krawiec (2005). It is also discussed by Gajda, Krawiec and Marlewski (2008), where Bézier curves are applied to describe such profiles.

A mathematical description of the profile is necessary when there are used modern machines applying CNC (computer numerical control). These machines manufacture, for instance by laser devices or compressed water streams, elements of machines and they cut a desired profile via moving their cutting tools along the trajectory which has to be defined mathematically.

2. SPLINE CURVE INTERPOLATING A PROFILE TO BE RECOVERED

In more details, we present the way at which we obtained the spline description of the closed curve passing through given m points $P_j = (x_j, y_j)$, ($j = 1, 2, ..., m$), sitting on the profile of a non-typical gear, as that seen in Fig.2 below.

Fig.2. Belt pulley manufactured in LEDM

The coordinates of thousands of such points are provided, as pairs of two numbers, x_j and y_j, by a CMM (coordinate measuring machine, see Fig.3) and a designer/constructor decides which ones of them are taken in aim to get a model; in the case reported here there were taken 24 extreme points of the outside

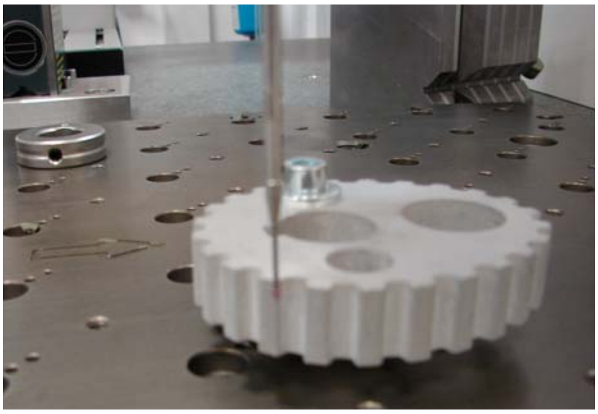


Fig.3. Measuring geometrical parameters of a non-circular pulley on the CMM Contura G2 from Carl Zeiss

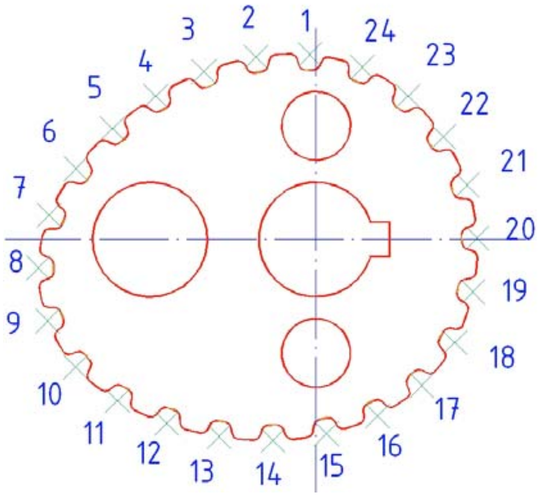


Fig.4. A sample profile and points P_j ($j = 1, 2, ..., m$; $m = 24$), sitting on its extreme envelope

envelope of the profile; you can see these points in Fig.4 (and we refer to them as measured points). These coordinates form two one-column vectors, namely

$$X = [\, x_1, x_2, ..., x_m]^T, \;\; Y = [\, y_1, y_2, ..., y_m]^T;$$

X and Y store the abscissas and the ordinates, respectively.

Since the profile is closed, it is natural to augment both vectors X and Y by the element equal to x_1 and y_1, respectively. This way we have so-called **abscissa-vector** X and **ordinate-vector** Y,

$$X = [\, x_1, x_2, ..., x_m, x_{m+1}]^T, \; Y = [\, y_1, y_2, ..., y_m, y_{m+1}]^T,$$

where $x_{m+1} = x_1$ and $y_{m+1} = y_m$. Obviously, we can say that there are given no m points P_j, but there are given m+1 points

$$P_j = (x_j, y_j), \; (j = 1, 2, ..., m, m+1),$$

where $P_{m+1} = P_1$.

The augmenting we did simplifies the presentation of the method below. This method provides the equation of the curve interpolating the sequence ($P_1, P_2, ..., P_m, P_{m+1}$) of measured points.

Obviously, the profile at hand, as well as the sequence ($P_1, P_2, ..., P_m, P_{m+1}$) of measured points, can not be entirely covered by the equation of the form $y = f(x)$. Fortunately, it can be describe parametrically; we can find the functions $x = x(t)$ and $y = y(t)$, both in the variable t, such that $x = x(t)$ interpolates the abscissa-vector X and $y = y(t)$ interpolates the ordinate-vector Y. Then the entire curve is governed by the equation

$$s(t) = [\, x(t), y(t)\,],$$

where t runs an appropriate interval and $s(t_j) = P_j$ for appropriately chosen values t_j of the variable t. These values, t_j, are called knots, or nodes, of interpolation. Below it will appear clear that it is very convenient to deal with the knots $t_j = j$. These knots determine the intervals $<t_j, t_{j+1}> = <j, j+1>$, $j = 1, 2, ..., m$; every one of them is called an elementary, or basic, interval (for the method we apply below).

Since we want to have an interpolating spline curve, we look for splines $x = x(t)$ and $y = y(t)$ satisfying collocation conditions

$$x(t_j) = x_j \text{ and } y(t_j) = y_j \text{ for } j = 1, 2, ..., m.$$

As in numerous applications, we will find the **third degree spline interpolation**. So we look for polynomials of third degree, f_j and g_j ($j = 1, 2, ..., m$), to interpolate the abscissa-vector X and ordinate-vector Y, respectively. For instance, on the j-th basic interval $<t_j, t_{j+1}>$ the spline $x = x(t)$ can be taken in the form

$$x(t) = a_j + b_j\cdot(t - t_j) + c_j\cdot(t - t_j)^2 + d_j\cdot(t - t_j)^3$$

and it is clear that it is defined by its coefficients, a_j, b_j, c_j, d_j.

Analogously, the j-th part of the spline $y = y(t)$ is determined by other quadruple, α_j, β_j, γ_j, δ_j, forming the coefficients of the linear combination of the same basic functions as above, i.e.,

$$t \to (t - t_j)^{k-1}\ , k = 1, 2, 3, 4.$$

The set of these four functions can be called a **Herriot-Reinch basis**. This basis is a key point to find all desired coefficients in **HeRA**, a **Herriot-Reinsch algorithm** (see Herriot and Reinsch 1973 and, e.g., Krawiec and Marlewski 2011).

In this algorithm we first calculate multipliers of the second power, $(t–t_j)^2$, next we use some combinations of them to get the multipliers of $(t - t_j)^3$ and $(t - t_j)^1$. At last, by the collocation condition the free terms are produced at once: it is $a_j = x_j$ for the spline $x = x(t)$ and, as easily as here, $\alpha_j = y_j$ for the spline $y = y(t)$.

The multipliers c_j of $(t - t_j)^2$ standing in the spline $x = x(t)$ are solutions of the system of linear algebraic equations

$$A\cdot c = \xi,$$

where A is the matrix of the system at hand, c comprises the coefficients to be calculated, ξ is the vector of right sides, so

$$c = [\, c_1, c_2, ..., c_m\,]^T,$$
$$\xi = [\, \xi_1, \xi_2, ..., \xi_m\,]^T,$$

$$A = \frac{1}{3}\cdot\begin{bmatrix} 4 & 1 & 0 & 0 & ... & 0 & 0 & 1 \\ 1 & 4 & 1 & 0 & ... & 0 & 0 & 0 \\ 0 & 1 & 4 & 1 & ... & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 & ... & 0 & 0 & 0 \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & .0 & ... & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & ... & 1 & 4 & 1 \\ 1 & 0 & 0 & 0 & ... & 0 & 1 & 4 \end{bmatrix},$$

The matrix A of order m and the vector ξ are built according to the definition of the spline with the additional requirement saying that this spline has to be periodic. This periodicity makes that the respective spline curve is closed and, at the same time, it puts both 1's in the left down and right upper corners (and there is the only difference with respect to the natural spline interpolation, where instead of these 1's we have 0's).

The elements of the right-side vector ξ are

$$\xi_1 = \rho_1 - \rho_m,$$
$$\xi j = \rho_j - \rho_{j-1} \quad \text{for } j = 2, 3, ..., m,$$

where $\rho_j = x_{j+1} - x_j$ for $j = 1, 2, ..., m$.

Solution c calculated, we get the vectors $d = [\, d_j\,]$ and $b = [\, b_j\,]$ by the formulas

$$dj = \gamma_{j+1} - \gamma_j \quad \text{for } j = 1, 2, ..., m–1,$$
$$d_m = \gamma_1 - \gamma_m\ ,$$
$$bj = r_j - 2\gamma_j - \gamma_{j+1} \text{ for } j = 1, 2, ..., m–1,$$
$$d_m = r_m - 2\gamma_m - \gamma_1,$$

where $\gamma_j := c_j/3$.

If necessary, go to Krinzeβa (2006) to see how the system $A\cdot c = \xi$ is built, and to Carnahan et al. (1969) to state that A is symmetric, irreducible and positively determined. The last fact was proved by de Boor and DeVore (1985) for arbitrarily spaced knots, but in our case we have a regular mesh and the determinants for m = 3, 4, 5, 6, 7, .., 24, 25 are equal 18, 64, 242, 900, 3362, 12544, ..., 17767236614400, 66308229755042 (it seems that there is still unknown a general formula for the value of m-th determinant).

In view of above remarks, the system $A\cdot c = \xi$ can be solved iteratively by Jacobi method. It can be also solved directly, $c = A^{-1}\cdot\xi$, by SOSes (symbolically oriented systems, as *Derive* from Texas Instruments, Inc., *Mathematica* from Wolfram Research, Inc.) without any roundings. Let's give an example: with m = 5 there is

$$A^{-1} = \frac{1}{8}\cdot\begin{bmatrix} 7 & -2 & 1 & -2 \\ -2 & 7 & -2 & 1 \\ 1 & -2 & 7 & -2 \\ -2 & 1 & -2 & 7 \end{bmatrix}.$$

Fortunately, the system $A\cdot c = \xi$ can be also solved by direct methods (including the basic one: Gauss elimination method) in NOSes (numerically oriented systems, e.g. Pascal, C++), because its matrix

$A = [\ a_{j,k}\ ]_{j,k=1,2,...,m}$ is well-conditioned; for instance, its condition number generated by maximal norm, $\| A \| = \max \{\ |\ a_{j,k}\ |\ :\ j, k = 1, 2, ..., m\ \}$, does not exceed 1.16 when $m \le 51$.

The construction of the equation $y = y(t)$ concerning the ordinate-vector Y is identical as that presented above, and this way it is reduced to solve the linear system $A{\cdot}\gamma = \eta$ with the same matrix A. Therefore we can treat these two tasks at once, namely instead of solving two systems of linear algebraic equations we solve the system

$$A{\cdot}W = F,$$

where both $W = [\ c \mid \gamma\ ]$, $F = [\ \xi \mid \eta\ ]$ are $m{\times}2$-matrices, their columns are vectors c and γ, and the vectors ξ and η are determined by the abscissa-vector X and the ordinate-vector Y, respectively,

A is as above, so it is worthy to get the inverse matrix A^{-1} or to apply any other method simultaneously to the pairs (c, ξ) and (γ, η) composed of columns involved in matrices W and F.

The solution W of this system, via its columns c and γ, yields the other coefficients (b_j, d_j etc.) of desired functions

$$x = x(t),\ y = y(t)$$

covering the positioning of abscissas and ordinates related to the equidistantly distributed knots $t_j = j$ related to the variable t, $j = 1, 2, ..., m+1$. This way we finally get the desired spline

$$s(t) = [\ x(t), y(t)\ ].$$

where t is the parameter; for $t \in \langle j, j+1\rangle$ we have the j-th fragment of the final spline curve, $j = 1, 2, ..., m$,

$x = x(t)$ is the function in the variable t; it covers the behavior of the spline at hand along the horizontal axis Ox,

$y= y(t)$ is the function in t; it describes the changes along the vertical axis Oy.

3. ANGLES AT WHICH A TOOL CUTS A PLATE

Obviously, to cut a desired pulley, or gear, off a steel (an aluminium alloy etc.) plate, the cutting tool controlled by a computer has to have not only the equation of the outside profile, but also the depth of toothing and the direction at which the a laser (a water jet etc.) moves. The depth, in millimeters, is the same for every tooth and it is simply passed in through the control panel. The directions are defined by angles, in grades, at which the tool is oriented with respect to the zero-direction (it is set just when the cutting machine starts its work, see Fig.5). The angles vary from tooth to tooth and they are passed to the cutting machine controller through a specialized programme as the sequence of numbers accompanying the coordinates of nodes P_j.

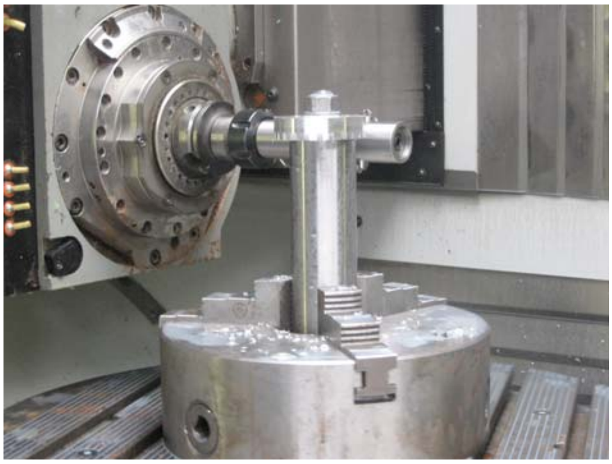


Fig.5. Forming the non-regular noncircular belt pulley on CNC milling machine Deckel Maho from DMG

Let

$$x(t) = a_j + b_j{\cdot}(t-j) + c_j{\cdot}(t-j)^2 + d_j{\cdot}(t-j)^3,$$
$$y(t) = \alpha_j + \beta_j{\cdot}(t-j) + \gamma_j{\cdot}(t-j)^2 + \delta_j{\cdot}(t-j)^3,$$

be the parametric equation of a j-th part of the spline curve $[\ x, y\ ] = s_j(t)$ obtained above; it says that this equation covers this curve when j runs from the point P_j to the point P_{j+1}. Since for every t where the derivative $x'(t)$ does not vanish there holds true

$$\frac{dy}{dx} = \frac{\dfrac{dy}{dt}}{\dfrac{dx}{dt}} = \frac{\beta_j + 2\gamma_j \cdot (t-j) + 3\delta_j \cdot (t-j)^2}{b_j + 2c_j \cdot (t-j) + 3d_j \cdot (t-j)^2},$$

so at P_j the tangent line has the slope β_j / b_j and, in consequence,

$$\varphi_j = \operatorname{arc\,cot} \frac{-b_j}{\beta_j}$$

is the angle under which, at the point P_j, the normal line to the curve $s = s_j(t)$ is inclined to the zero-direction of the cutting machine.

Obviously, the same straight line is perpendicular to the adjacent spline fragment $s = s_{j-1}(t)$ (for $j = 1$ we identify $s_0 = s_{m+1}$). If $b_j = 0$, then the tangent line is parallel do the axis Ox and φ_j is 90° or 270° depending on the sign of the coefficient β_j.

In Fig.6 there are traced segments of perpendicular lines to the spline curve. For instructive purpose, a segment of the normal at the knot P_2 is traced longer than other segments. Since here the zero-direction of the cutting machine coincides with the direction of the Oy axis and the direction at which the object in forming is clockwise, we have $\varphi_2 = 170.489°$ (the normal at P_2 is inclined to the axis Oy upon the angle 170°29’) and, e.g. $\varphi_1 = 4.8545° = 4°51’$, $\varphi_{24} = 22.8724° = 22°52’$, $\varphi_{20} = 86.9719° = 86°58’$.

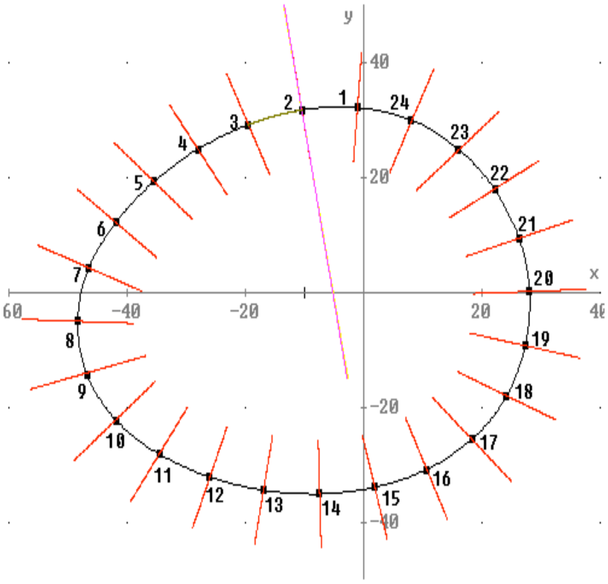


Fig.6. Nodes P_j (marked by values of their index j: 1, 2, 3, ..., 24) of the interpolatory spline, the spline curve itself and the normal lines to it passing through the nodes (figure produced in Derive 5 for Windows)

5. MANUFACTURING IRREGULAR PULLEYS AND GEARS

Belt pulleys, as well as gears, are manufactured by the profiling (a.k.a. shaping) method or by the envelope method; they both are classified as matching techniques. The disadvantage of the profiling technique is that it needs to use few cutting tools and, practically always, an additional mechanical polishing has to be performed. Moreover, manufactured gears have so-called pitch error, which results from nature of process. By contrast, the envelope technique, a.k.a. the direct generation of noncircular gears, requires the design of non-typical manufacturing process and application of numerical controlled slotting machine; as several years earlier, the last one treatment is still not very common, see Kujawski (1992).

Noncircular belt pulleys can be also shaped by abbrasive water jets, a.k.a. watersaws, which are tools capable of slicing into metal, or other material, using a jet of water (usually enriched with an abrasive substance) at high velocity and pressure.

Another way to manufacture gears and pulleys is the laser cutting (in particular by lasers where CO_2 is the lasing material), but here only relatively thin gears can be obtained, see, e.g., Krawiec (2009).

As we read in Krawiec (2010), a good alternative method to all techniques mentioned above is the application of universal CNC milling machine with set of end mills, as well as rapid prototyping and rapid manufacturing methods (e.g., 3D-printing, FDM, SLS) to get gear wheels. A relatively new idea of gears forming is the usage of numerical controlled milling machine, where there is installed special two cutting edges tool. In accordance with this idea the manufacturing process of gear wheel is composed of the following movements: tool rotation in relation to spindle axis, rotation and displacement of numerical controlled table, where initially shaped blank is positioned. In the referred method the basic task in the elaboration of the controlling program is the proper correlation between the movement of the table and the rotation of the cutting head. This relation can be doubtlessly stated by mathematical formulas and the spline description presented in this paper provides it.

4. FINAL COMMENTS

Taking into account the regimes obligatory for the manufacturing belt pulleys and gears we derived the equation of spline curve passing through given points chosen, by a designer, from points provided by any scanning machine. The accuracy at which these data are gathered is fairly better than the ranges at which the desired gear/pulley has to be manufactured; the respective errors are even 0.001 mm in CMM and 0.1 mm in cutting process if the diameter ranges between 50 and 100 mm. That's why we did not smooth the scanned data and we did not smooth the spline determined on given points. The second procedure, aiming in the smoothing of a spline generated by given sequence of points P_j, is discussed, a.o., in Reinsch (1967) and, moreover, Hutchinson and de Hoog (1985).

Since invented, the spline interpolation is widely applied not only in civil engineering and mechanics, but also in such areas as statistics, geometry and (rail)way planning – see Kamenschykova (2008), Kranjc (2009) and Moreb (2009). When applied to the problem considered in this paper, it also provides a fast way to determine the angles of the cut and that's why it is more friendly to the practice than Bézier approximation (although the formulas involved in the spline description are more complex that the simply, elegant dependency taking place in Bézier approach).

REFERENCES

Bair, B.W. 2009, Tooth profile generation and analysis of crowned elliptical gears, *ASME J. Mech. Des.* 131, 6, 074503-1-6.

Bair, B.W., Chen, C., Chen, S., Chou, C., 2007, Mathematical model and characteristic analysis of elliptical gears manufactured by circular-arc shaper cutters, *ASME J. Mech. Des* 129, 2, 210-217

Burden, R.L., Faires, J.D., 1985, *Numerical analysis*, PWS Publishers, Boston.

Carnahan, B., Luther, H.A., Wilkes J.O., 1969, *Applied numerical methods*, John Wiley & Sons Inc.

Chapra, S.C., Canale, R.P., 1990, *Numerical methods for engineers*, McGraw Hill Int. Editions

Danieli, G.A., Mundo, D., 2005, New developments in variable radius gears using constant pressure angle teeth, *Mech. Mach. Theory* 40, 203-217

de Boor, C., DeVore, R., 1985, A geometric proof of total posivity for spline interpolation, *Math. of Computation*, Vol.45, No. 172, 497-504

Gajda, K., Krawiec, P., Marlewski, A., 2008, Bézier fitting to almost oval gear devices, *Proc. 7th Int. Workshop on Modelling & Applied Simulation*, Campora S.Giovanni, Italia, 210-216

Herriot, J.G. , Reinsch, C.H., 1973, Algorithm 472: procedures for natural spline interpolation, *Communications of the Association for Computing Machinery*, 16, 12, 763-768

Hutchinson, M.F., de Hoog, F.R., 1985, Smoothing noisy data with spline functions, *Numerische Mathematik* 47, 99-106

JianGang Li, Xutang Wu, ShiMin Mao, 2007, Numerical computing method of noncircular gear tooth profiles generated by shaper cutters, *Int .J. Adv. Manuf. Technology*, 33, 1098-1105

Kamenschykova, O., 2008, Approximation of random processes by cubic splines, *Theory Stoch. Processes* 14, no.3-4, 53–66

Krajnc, M., 2009, Geometric Hermite interpolation by cubic G1 splines, *Nonlinear Analysis*, 70, No.7, 2614—2626

Laczik, B., 2003, Involute profile of pon-circular gears, *Proc. of 11th World Congress in Mech. & Machin. Sciences*, Tianjin, China; Available from: http://next-1b.manuf.bme.hu/Laczik/Non circular. pdf [accessed 13 July 2010]

Li Xin Cao, Jian Liu, 2002, Research on the mapping models and designing methods of variable-ratio chain/belt drives, *Mech. Mach. Theory*, 37, 9, 955-970

Litvin, F.L., Fuentes, A., 2004, *Gear geometry and applied theory*, Cambridge University Press

Krawiec, P., 2005, A multi-wheel toothed belt transmission system with a periodically variable transmission ratio, *Archive of Mechanical Technology and Automation*, 25, 2, 125-132

Krawiec, P., 2009, Kształtowanie nieokrągłych kół pasowych ciętych laserem oraz strumieniem wody ze ścierniwem [Forming of non-circular belt pulleys with gas laser cutting and water stream with abrasive material], *Przegląd Spawalnictwa* 7-8, 56-60

Krawiec P., 2010, Kształtowanie wieńców nieokrągłych kół pasowych na uniwersalnych frezarkach sterowanych numerycznie [Noncircular cogbelt pulleys forming with universal NC milling machines], *Mechanik* 1, 36-39

Krawiec, P., Marlewski, A., 2011, Spline description of not typical gears for string transmissions, to be submitted to *Journal of Theoretical and Applied Mechanics*

Krinzeβa, F., 2006, *Fast construction and evaluation of interpolatory periodic spline curves*; Available from: http://www.krinzessa.de/Splines.pdf [accessed 13 July 2010]

Kujawski, M., 1992, *Mechanizmy obiegowe z nieokrągłymi kołami zębatymi* [Planetary mechanisms with noncircular gears], Wyd. Politechniki Pozn. Poznań

Ming-Feng Tsay, Zhang-Hua Fong, 2005, Study on the generalized mathematical model of noncircular gears, *Mathematical and Computer Modelling* 41, 555-569

Moreb, A., 2009, Spline technique for modeling roadway profile to minimize earthwork cost, *J. Ind. Manag. Optim.* 5, No.2, p. 275–283

Reinsch, C.H., 1967, Smoothing by spline functions, *Numerische Mathematik*, 10, 177-183

AUTHORS' BIOGRAPHIES

Piotr Krawiec is employed as tutor in the Chair of Basic Machine Design at Poznań University of Technology. His scientific interests are designing, manufacturing and measuring problems as well as experimental research of string drives. He is the manager of laboratory of Computer Aided Design and Engineering Record. He gives lectures on Machine Design, Fundamentals, CAD, CAM, CAE. He is a member of the technical committee at the Polish Committee for Standardization.

Adam Marlewski works for the Institute of Mathematics at Poznań University of Technology. He got his Ph.D. in the theory of approximation and since years is involved in studies in maths (number theory, geometry and topology) and its application (in the electricity, the flows along various profiles, the chemical equilibrium conditions, the remote sensing). He authored first Polish book on a computer algebra system (*Derive* from Soft Warehouse, in 1992), he published textbooks in numerical methods and (matrix and abstract) algebra. In 1990 he co-founded the popular series of books, *Microcomputer User's Library*, and the journal *Pro Dialog*. featuring articles on information processing system design, computer software and practical applications.

Authors Index