*Simulation in Produktion
und Logistik 2019*
Matthias Putz & Andreas Schlegel (Hrsg.)
Wissenschaftliche Scripten, Auerbach 2019

Usage of Simulation Models to Compare and Increase Accuracy of Statistical Forecast of Logistical KPI in Manufacturing Environments

Die Nutzung von Simulationsmodellen für die Verbesserung und Vergleich von statistischen Vorhersagemethoden logistischer KPI in Fertigungsumgebungen

Mike Gißrau, Lutz Muche, Michael Ertelt, SYSTEMA GmbH, Dresden (Germany),
mike.gissrau@systema.com, lutz.muche@systema.com,
michael.ertelt@systema.com

Claudio Hartmann, TU Dresden, Dresden (Germany),
claudio.hartmann@tu-dresden.de

Abstract: Today simulation is one of the still growing techniques to identify and analyse complex production environments, but also requires a lot of manual and complex input to achieve reasonable results. In this paper, statistical methods for forecasting production key performance indicators (KPI) in manufacturing environments are presented to achieve a real-time capable and sufficient prediction result. The results are compared by different scenarios based on simulated and real manufacturing data. The idea of improving the accuracy of statistical forecasts by simulation models is discussed and evaluated. The results show inconsistency in prediction quality, where machine learning approaches with additional simulation model usage have the highest potential for valuable results.

1 Introduction and Motivation

Today simulation is one of the still growing techniques to identify and analyse complex production environments, e.g. predict what will happen in future. This could be the future "behaviour" of a machine or a whole manufacturing environment in order to predict upcoming KPI trends or problems (e.g. upstream to bottleneck machines). Banks (1997) viewed simulation to be "consistently one of the top three methodologies used by industrial engineers, management scientists, and operations research".

In manufacturing environments, simulation becomes more and more a major element on short-term and long-term planning and improvement of wafer fabrication. Typical fields of investigation are production planning, ramp-up scenarios or the improvement of dispatching and scheduling approaches (Noack

2008). Nevertheless, simulation is often too slow or resource intensive to fulfil different requirements or use cases. There are some attempts to provide or support simulation models also with some real-time capability e.g. by using an online-model approach (Noack 2010). In case of performance prediction, future problems should be detected in almost real-time, that appropriate counter measures can be taken.

Today's manufacturing environments are characterised by a lot of automation projects in order to increase production capacity and efficiency, which can be summarised with typical key words like *Digitalisation* or *Industry 4.0*. Within these developments, also the automation of data plays an important role. Prediction capabilities for future events or behaviours will allow additional optimisation potential, e.g. for maintenance activities.

Real-time capability is one major part to allow fast and efficient reactions on what currently happens at the facility floor. Manufacturing KPIs forecast of what will happen in the future even in case of production interruptions. Various KPIs are available to evaluate effectiveness and speed, e.g. work in process (WIP), cycle time or throughput (Hopp and Spearmen 2001).

In this paper, we focus on the WIP as one major representation of typical KPI on manufacturing level. Different perspectives are investigated, e.g. tool or area level. Important statistical algorithms are evaluated for their correctness regarding future KPI developments. Simulation data is used from two perspectives, as input for our evaluation as well as additional input to improve the forecast results in case of long-term applications.

2 Previous Work

The current paper continues the work by Gißrau (2017), where simulation models are generated based on a predefined data model and structure. For this a defined event and state model is applied in both, simulation and real environment. Based on the event and the real-time aggregated data, different master elements for the simulation model are produced in real-time. This includes for example the production flow, equipment schedules or equipment down statistics. This data is used to generate valid long-term simulation models for new or unknown products or scenarios in order to increase accuracy of short-term and long-term statistical prediction. Figure 1 shows the general workflow of the model generation process.

This method is used for two application fields, the generation of valid data for the evaluation of the statistical forecast methods and for investigation of simulation support for machine learning approaches.

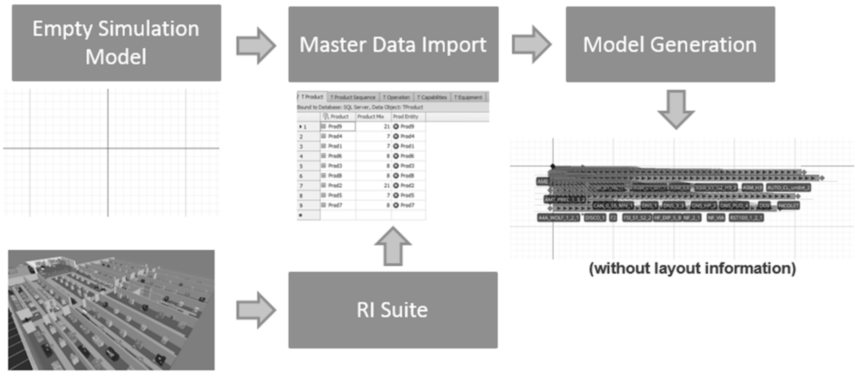

Figure 1: *General workflow of automated simulation model generation*

3 Statistical Methods for Forecast

The statistical forecast of manufacturing KPI is still a difficult and complex task. Different impacts of various manufacturing disturbances cause high variability even in semiconductor fabrication. These disturbances could be unplanned machine breakdowns or lot flow disturbances. Even with today's increasing hardware capabilities, simulation models are still too complex to produce real-time forecast results. Therefore, statistical forecasts seem to be favourable in production.

In manufacturing environments, two general classes of performance indicators are available, continuous and discrete KPI. For continuous KPI, at each time, a value is measureable (like an analogue signal in electricity). For the WIP, a typical continues KPI, at each time a value is available. For discrete KPI, only at defined points of time, a value is available (like a digital signal). Typical KPIs are cycle time or on-time delivery, where new values are only generated by defined events like a lot-finished event.

In this paper, we focus on the WIP as a central indicator for various problems in a manufacturing area, e.g. detection of bottleneck situations. Furthermore, planning activities like tool maintenance scheduling can be supported by reasonably WIP predictions. The topic of the forecast can be further differentiated in short-term and long-term periods. For short-term periods, the real-time capability as well as fast results are necessary to support the manufacturing control and data automation process. Typical use cases are maintenance period planning for the current day or bottleneck detection of moving bottlenecks. Long-term predictions in the period of weeks or months are used for detecting general trends or problems, e.g. on-time delivery.

In the next sections, we discuss different statistical approaches for short-term (8 to 48 hours) prediction of WIP for tools and areas within manufacturing environments. The test data set used for the evaluation of the prediction results originates from real mid-size, high-mix, low-volume semiconductor facilities as well as from a simulated model facility with the following properties:

- 9 main products with 234 to 355 process steps per product
- 189 machines

3.1 Simple Propagation Models

Simple propagation models are often used to generate a simple forecast result for future behaviours. In our case, we applied a simple propagation model where the WIP of the operation in a period depends on three main elements: WIP which stays in the operation during the period, WIP which leaves the operation during the period and WIP which arrives at the operation during the period

Figure 2 shows a simplified scheme for this idea. The operation WIP of period T is calculated as follows:

$$WIP_{Op}(T) = WIP_{Stay}(T) + \sum_{Operation} WIP_{Arrive}(T) - \sum_{Operation} WIP_{Leave}(T) \qquad (1)$$

This simple propagation model requires detailed information about execution times of lots, the work schedules of each lot as well as other variability sources like tool downs or maintenance activities. In our case, we simplified the arrival and the following processing via throughput rates based on historical evaluations for each operation.

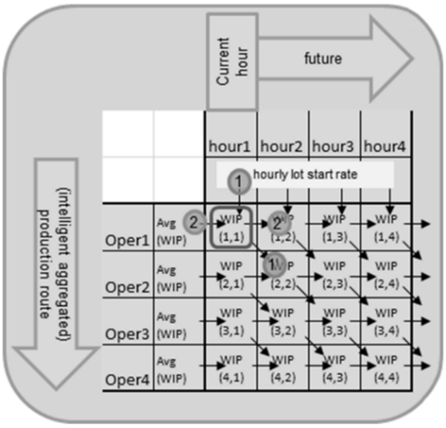

Figure 2: *Simple WIP propagation model*

Because of the simplified approach, the results are not sufficient for a valid short term forecast in complex environments, but will work in simplified environments which no product mix and re-entrancy. Figure 3 shows an example where the throughput model is applied and used as a dashboard reporting.

	29.07.2017																	
	29.07.2017																	
	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00
1410	23,47	12,11		23,51	12,06		23,49	12,08		23,51	12,03		23,50	12,11		23,51	12,11	
1436		11,38			11,34			11,34			11,33			11,34			11,36	
2130		0,51	10,73		0,60	10,67		0,58	10,71		0,64	10,63		0,55	10,70		0,53	10,74
2281			11,27			11,28			11,25			11,29			11,27			11,27
3725	24,00	8,94	2,00	24,00	9,01	2,05	24,00	8,97	2,04	24,00	8,82	2,07	24,00	8,95	2,03	24,00	8,98	1,99
3800		7,28			7,27			7,31			7,60			7,27			7,27	
3820		5,11			5,11			5,13			5,09			5,07			5,10	
3827		2,67	2,44		2,60	2,52		2,59	2,56		2,49	2,62		2,71	2,43		2,65	2,44
3992	9,84		21,56	9,95		21,48	22,70		21,44	9,92		21,38	10,16		21,57	10,00		21,56
4040	6,69			7,03			1,30	5,71		6,71			7,03			7,01		
4041	6,19			6,17				6,18		6,17			6,19			6,15		
4070	1,29	4,98		0,85	5,40			6,27		1,20	5,05		0,62	5,65		0,84	5,42	

Figure 3: *Simple propagation model in reporting dashboard*

3.2 Independent Linear Models

At the first step, independent linear models are used to predict the WIP as time series. For this, the hourly average WIP is calculated based on a fractal scheme visualised in Figure 4.

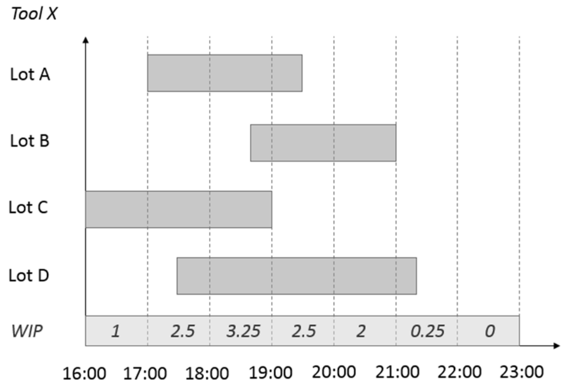

Figure 4: *Hourly WIP based on Fractal-Calculation*

Each lot which is located e.g. on a machine or in an area is added to the average WIP based on its time portion it spent at this position. Simple independent linear model approaches like regression generate models based on the following equation:

$$WIP_{Op}(t) = c + \sum_{i=1}^{T} a_i * WIP_{Op}(t-i) \tag{2}$$

Figure 5 illustrates an example based on simulated model data.

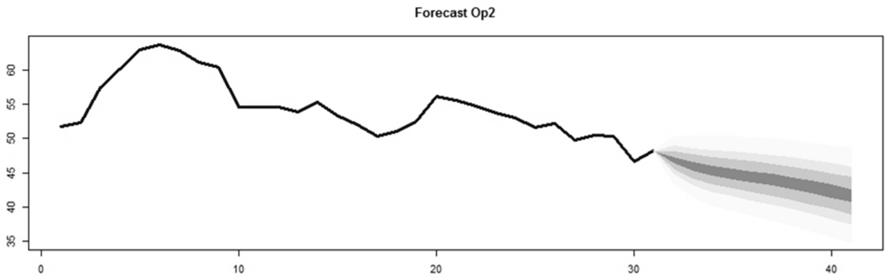

Figure 5: *Linear model calculation with confidence belt*

The practical evaluation of this approach also does not produce sufficient results, because of known and unknown correlations between the operations and its behaviour. One example for the correlation is visualised in Figure 6, where additional influences are indicated by the auto correlation (ACF).

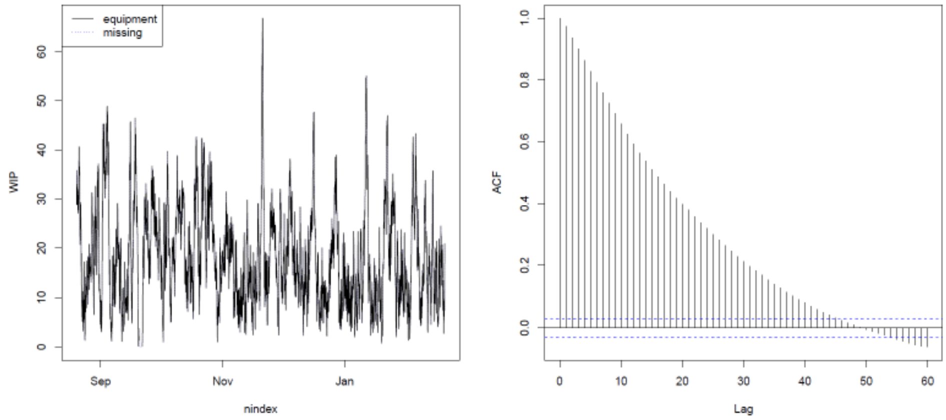

Figure 6: *WIP evolution of one tool and its ACF*

Several different linear model assumptions do not improve the forecast quality in contrast to the naive forecast based on the historical average of the WIP (e.g. in Figure 7).

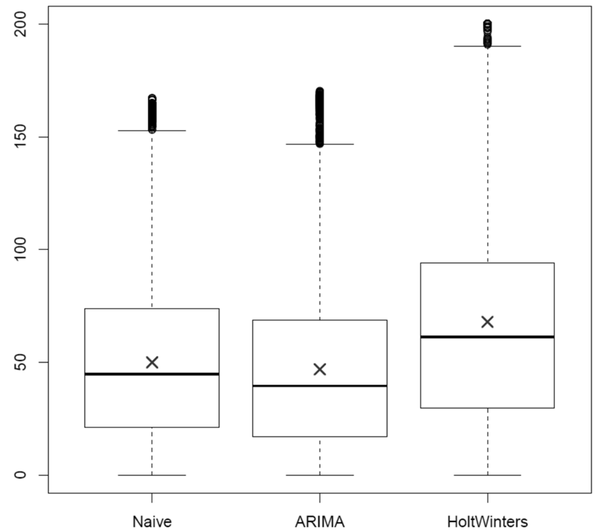

Figure 7: *Naive forecast error in contrast to ARIMA and Holt Winters*

3.3 Dependent Linear Models

As a second step, some mathematical-statistical methods for extrapolating multiple time series are examined. The basic idea is to combine several historical time series (e.g. consecutive operations and their WIP) for generating a forecast for an operation or machine. With the R-programming package, three different algorithm types are investigated:

- Multiple time series analysis with vars or MTS
- Multivariate adaptive regression with hinge functions with Earth/Mars (en-

hanced adaptive regression through hinges/multivariate adaptive regression splines)
- Regression with a NN approach with Caret (classification and regression training).

These three methods work with hourly fractals of the WIP as in the linear model section described. The calculations run very slow with large numbers of process steps (several hundred operations or equipment). They therefore require a targeted selection of predecessor processes that are included in the calculation, or the knowledge of a given sequence of process steps. Some R libraries offer functions for detecting relationships between curves (causal relationships, measures of similarity, degree of interaction).

The general tendency of the functional patterns shows an insufficient forecast quality and too large deviations from reality. An extension of the linear model by the use of several dependent linear quantities as well as various options for data preprocessing and training control did not bring any significant improvements (see Figure 8).

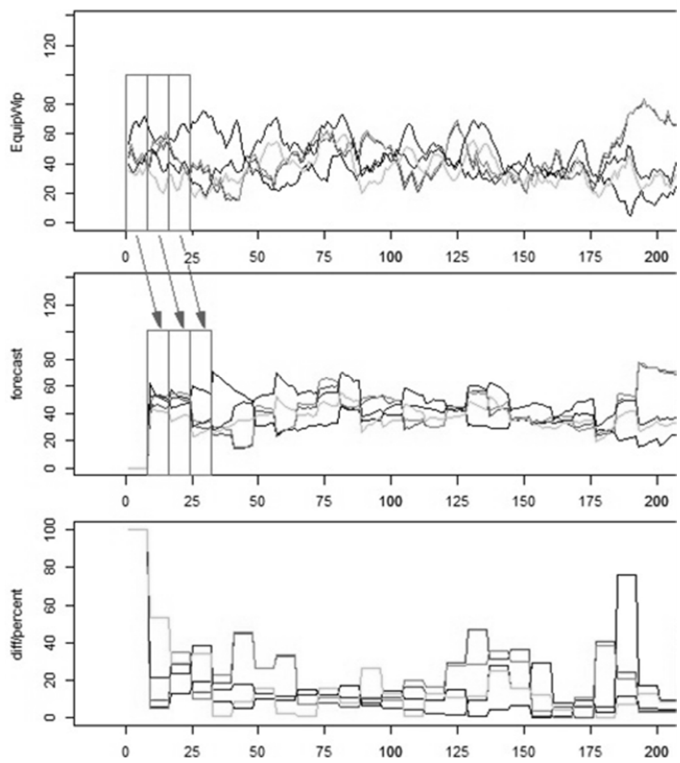

Figure 8: *VARS model prediction result with 8 hours forecast horizon*

3.4 Machine Learning Approaches

Today, a wide range of different approaches and solutions are also examined via machine learning approaches. The basic idea of a self-learning approach via a preprocessed set of input data seems to have high potential when applied to forecast use cases.

The basic algorithm depends on two steps. The data preprocessing consists of an graphical ordering or sorting by Kohonen (1995). Kohonen nets enable an unsupervised or supervised learning. The algorithm allows training of self-organising maps (SOMs). The training process can be controlled by many different parameters, such as the shape of the grid, the variable learning rate, the variable neighborhood radius, the length and number of codebook vectors and the number of iterations.

A codebook vector is a vector with the most frequently occurring features, in the case of multiple time series, sets of typical curve segments describe processes that are frequently repeated. Data segments can be assigned to a codebook vector that describes a frequently occurring operating state. It can be used to classify fabrication stages into groups. In order to be able to make predictions about the future (in a second step), model extensions in form of hidden layer are necessary to generate a forecast.

Transition probabilities for changing from one state to the next can be determined. This forms the basis for a forecast of the further WIP processes. The Kohonen Net itself is an input to a hidden-layer structure to predict the forecast for one or several operations. Figure 9 shows an example of the forecast result in comparison to the real WIP evolution.

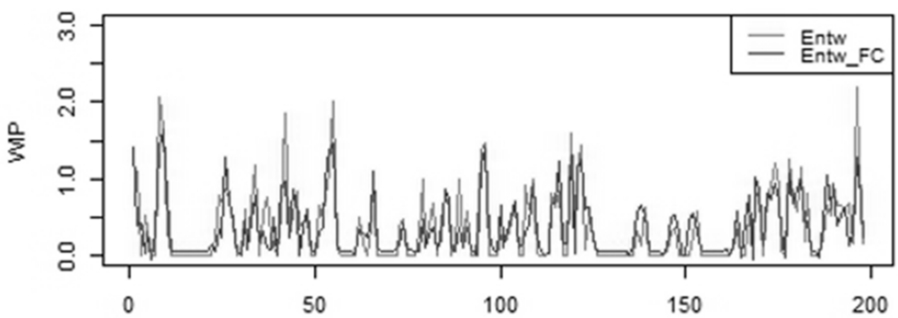

Figure 9: *Forecast (red) vs. real by usage of Kohonen Nets and Neural Networks*

In general, machine-learning approaches indicate a high potential for usage in forecast situations. Our test data error margins grew up to 20%, which is acceptable in most cases.

4 Supportive Simulation Models for Statistical Forecast (long-term)

Besides the short-term prediction of typical KPI representing the current and near future state of a manufacturing environment, long-term predictions (weeks or month) have different aspects. Typical use cases are for example prediction of on-time delivery for a new order or a new product. Even when having a complete new product without any historical experience, simulation models can support the long-term forecast by producing simulated statistical data. This may also be used in statistical short-term forecast.

For this, long-term simulation models are generated (see section 2) for this use case. First investigations show a small improvement of forecast accuracy in case of short-term prediction with new unknown products. Further investigations are currently undertaken.

5 Conclusion and Outlook

In this paper, we compare several statistical methods for a typical continuous KPI in the manufacturing performance evaluation process. With the WIP, the machine learning approach seems to have the best potential of delivering sufficient results. Our use cases show error rates up to 20% to the real value with is quite acceptable. Of course, the forecast quality depends mainly on the available data as well as the manufacturing properties at each site. For long-term forecast use cases like on-time delivery prediction, simulation models allow more flexible experimental settings without changing the input data stream. Additionally simulation can support the short-term statistical prediction without the need of online-connection of the simulation model to the factory environment. Further investigations of long-term prediction currently undertaken.

6 Acknowledgement

A part of the work has been performed in the project Forecast funded by the European Regional Development Fund (ERDF) under co-financing by the Free State of Saxony (100320127, 100320128).

Literature

Andersson, M.; Ng, A.; Grimm, H.: Simulation optimization for industrial scheduling using hybrid genetic representation. In Mason, S. J.; Hill, R. R.; Mönch, L.; Rose, O; Jefferson, T; Fowler, J. W.(editors), Proceedings of the 2008 Winter Simulation Conference (WSC), Miami (USA), 7.-10. December 2008, pp. 2004–2011.

Afolabi, M.; Olude, O.: Predicting Stock Prices Using a Hybrid Kohonen Self Organizing Map (SOM), In Hansen, Sean; Berente, Nicholas; Lyytinen, Kalle (editors), 40th Annual Hawaii International Conference on System Sciences, 3.-6. January 2007, Hawaii (USA) DOI:10.1109/HICSS.2007.441.

Banks, J.: Discrete event system simulation. Prentice Hall, 3. ed., 2001. New York (USA)

Banks, J.: Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice. John Wiley & Sons, Inc., 2007. New York (USA)

Chick, M.; Muhamad, M.: The Influence of Industrial Engineering in Semiconductor Fabrication Factory Optimization, International Conference on Industrial Engineering and Operations Management, March 8-10, 2016, Kuala Lumpur (Malaysia). Pp. 2845-2853

Gißrau, M, Muche, L: Statistical KPI Forecast in Semiconductor Manufacturing, APCM Conference Dresden, 16.-18. April 2018, Dresden (Germany).

Gißrau, M, Gereke, T: Application of real-time reporting data for online simulation models, In Wenzel, S; Peter, T (editor) ASIM Conference Kassel, 20.-22. September 2017, Kassel (Germany), pp. 393-402

Gißrau, M.: Development and Introduction of Semiconductor Production Control Enhancements for Customer Oriented Manufacturing, PhD Thesis, Universität der Bundeswehr München, Faculty Computer Science, 2013

Hopp, W.J.; Spearman, M.L.: Factory physics: foundations of manufacturing management. Irwin McGraw-Hill, 2. ed. edition, 2001. New York (USA)

Hunt, J.W.; McIlroy, M.D.: "An Algorithm for Differential File Comparison". Computing Science Technical Report, Bell Laboratories. 1976, New Jersey (USA)

Luhn, G.; Orbanz, J.: “Deep” Production Complexity Management. Real-Time Manufacturing in Fractal Dimensions, Scientific Presentation at 16th European Advanced Process Control and Manufacturing Conference (apclm), April 11-13, 2016, Reutlingen (Germany)

Luhn, G.; Bartl, K.; Habich, D.; Zinner, M.; Postel, J.: Real-Time Information Base as key enabler for Manufacturing Intelligence and Industrie 4.0 - Building the bridge between “real” Real-time Exploitation and Big Data via mathematically grounded Information Fundamentals; 26th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), Saratoga Springs (USA) 3.-6. May 2015

Luhn, G.; Bartl, K.: Real-Time Information System as key enabler for “Industrie 4.0” inspired Semiconductor Manufacturing – TURNING DATA INTO FUNDAMENTAL INFORMATION, scientific talk at 15th European Advanced Process Control and Manufacturing Conference (apclm), 13.-15. April 2015, Freising (Germany)

Luhn, G.; Zinner, M, Ertelt, M: Real-Time Information Systems and Methodology based on Continuous Homomorphic Processing in Linear Information Spaces; patent application pending, placed as US Provisional PA, 7.3.2014

Noack, D.; Rose, O.: A simulation based optimization algorithm for slack reduction and workforce scheduling. In Mason, S. J.; Hill, R. R.; Mönch, L.; Rose, O; Jefferson, T; Fowler, J. W.(editors), Proceedings of the 2008 Winter Simulation Conference (WSC), Miami (USA), 7.-10. December 2008, pp. 1989–1994

Noack, D. et. al. Towards Realization of a high-fidelity simulation model for short-term horizon forecasting in wafer fabrication facilities. In Johansson, B.; Jain, S.; Montoya-Torres, J.; Hugan, J.; Yücesan, E. (editors), Proceedings of the 2010 Winter Simulation Conference, Baltimore (USA), 5.-8.December 2010 pp. 2563 – 2574

Noack, D.: Online Simulation in Semiconductor Manufacturing, PhD Thesis, Universität der Bundeswehr München, Faculty Computer Science, 2012

Son, Y.; Wysk, R.A.: Automatic simulation model generation for simulation-based, real-time shop floor control. Journal Computers in Industry. Volume 45 Issue 3, July 2001. Pages 291 – 308

Kohonen, T: Self-Organizing Maps. Springer, Berlin (Germany) 1995.