*Proceedings of the 2012 Winter Simulation Conference*
*C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, eds.*

# SINGLE TOOLSET MODELING APPROACHES IN SEMICONDUCTOR MANUFACTURING

Kamil Erkan Kabak

Beykent University
Sisli-Ayazaga Campus
34396 Istanbul, TURKEY

Cathal Heavey

Enterprise Research Centre
University of Limerick
Limerick, IRELAND

Brian Kernan

Enterprise Research Centre
University of Limerick
Limerick, IRELAND

## ABSTRACT

Traditional industrial engineering techniques including mathematical models are not sufficient to examine sophisticated manufacturing systems such as semiconductor manufacturing. As such, simulation modeling is used extensively in the design and analysis of semiconductor manufacturing operations. This study explores the use of simulation modeling of single semiconductor toolsets. In the literature a number of modeling approaches for single toolset analysis can be identified. The purpose of this study is to review and evaluate these approaches.

## 1 INTRODUCTION

This study evaluates the use of simulation modeling of single semiconductor toolsets, which in the abstract, can be analyzed using a full fab simulation model. However, this is generally not feasible due to the scale and complexity of semiconductor manufacturing especially if detailed engineering analysis of the toolset is required. This is because a full fab model will not have the engineering detail captured at the toolset level. For example, it may not include details of a wetbench toolset, such as the robot control or the number of carriers in this toolset. For detailed analysis of a single toolset, typically a simulation model is built using the required level of details of the toolset being analyzed and that simplifies the other elements of the fab. This simplification is not straightforward due to the reentrant nature of semiconductor manufacturing. In the literature a number of modeling approaches for single toolset analysis can be identified. The purpose of this paper is to review and evaluate these approaches.

This paper is structured as follows, different simulation modeling approaches used in semiconductor manufacturing are surveyed in the literature in Section 2. In Sections 3 and 4, these three approaches are evaluated using a photolithography toolset as a testbed built from ASIC (Application-Specific Integrated Circuts) fab data and simulation results are discussed. Finally, the paper ends with concluding remarks in Section 5.

## 2 LITERATURE RESEARCH

With regard to toolset modeling, two main approaches can be identified from the literature: 1) Reentrant approach and 2) Non-reentrant approach. The reentrant approach is one of the most widely applied simplification methods for DES (Discrete Event Simulation) models in the literature (see Brooks and Tobias (2000), Jacobs et al. (2003), Jacobs et al. (2006), Veeger et al. (2011), Zeigler (1976) for other

simplification approaches). In the reentrant approach, the models introduce delays to simplify the re-entrant visits into the toolset being analyzed, for example the photolithography toolset. Each visit into the photolithography area is referred to as a different pass or stage, also called layer or level. The simulation modeling structure in this approach is based on three types of delays (or stages) as illustrated in Figure 1 (Kabak, Heavey, and Corbett 2008). These stages are: arrival, first delay, photolithography area, second delay and third delay (Kabak, Heavey, and Corbett 2008). Note that there may be several instances of the second delay.

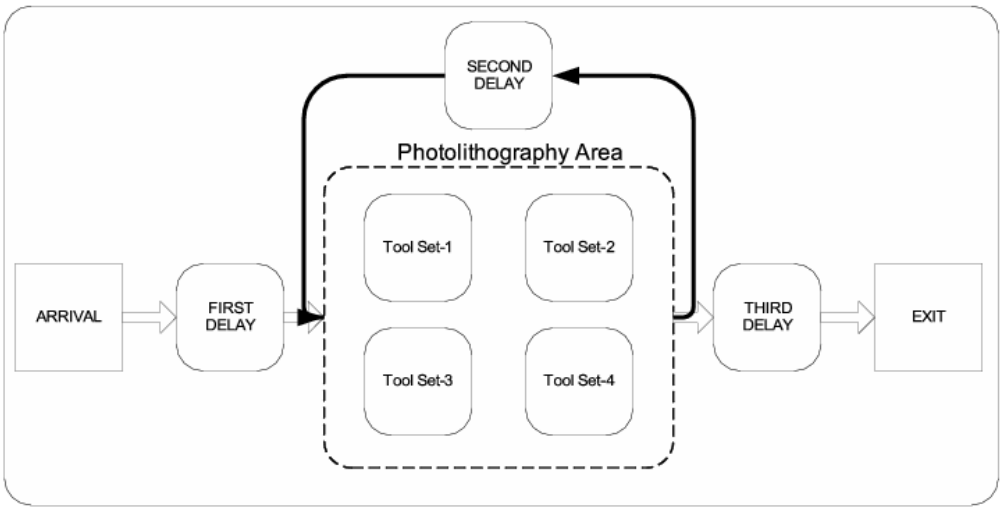



Figure 1: Conceptual model of reentrant approach.

Reentrant models can be further categorized into two sub groups labeled reentrant simple models and reentrant advanced models. An example of a reentrant simple models is Peikert, Thoma, and Brown (1998) who analyzed a photolithography toolset modeling the delays as dummy work stations with an infinite number of servers. They model each visit to the photolithography area using a single dummy stage. While they do not label the dummy stations as the first delay, second delay and third delays, their modeling approach is the same as the approach in Figure 1. Each dummy stage has a process time from a triangular distribution and a queue delay time which is calculated using a cycle time factor. The cycle time factors are estimated from historical fab data for different product groups.

Rose (1998) uses a simplified fab model similar to the approach in Figure 1 in order to analyze WIP (Work-In-Progress) and cycle time evolution after a catastrophic failure at a bottleneck workcentre. However, rather than categorize the delays, his conceptual model includes a delay unit and a control unit after the bottleneck workcentre as shown in Figure 2. Rose (1998) analyzed WIP and cycle time evolution under three types of delay patterns; constant delays, shifted exponential delays and Erlang-5 distributed delays. In a similar model, Rose (1999) explicitly models three different delay patterns as given and applied the shifted Erlang-k distributions for the delay units to compare CONLOAD (Constant load), CONWIP (Constant WIP) and CONWORK (Constant Work).

Drawbacks of the reentrant simple models have been reported in the literature. Rose (2000) explains the shortcomings of the reentrant simple models. In short, the main shortcomings are correlation, overtaking and deviations in cycle time estimation. These shortcomings are mainly the result of the lack of lot dependences which exist in real fabs (Rose 2000). Due to these shortcomings, Rose (2000) and Rose (2007) proposes improvements to the reentrant simple models. One proposal Rose (2007) uses is load-dependent distributions rather than fixed distributions. He also compares cycle time distributions with the improved bottleneck based approach with cycle time distributions of a full fab model. He observes that cycle time distributions with the reentrant advanced model have a larger range than the cycle time distributions of the full fab

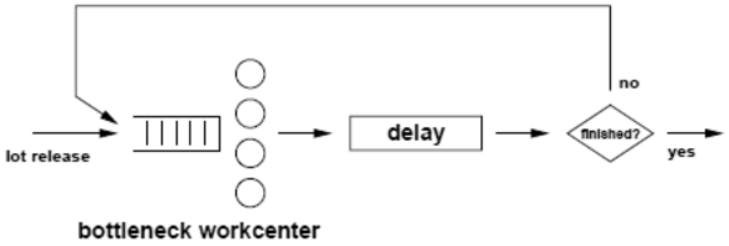



Figure 2: Simplified wafer fab model.

model. The main reason is attributed to the overtaking of the lots in the reentrant simple models. To prevent overtaking of the lots, Rose (2007) suggests using the service time distribution for the delays by computing the interarrival times of the lots at the bottleneck. Three approaches are proposed to compute the interarrival times of the lots. The simulation models built according to these approaches are called reentrant advanced models. The first approach is to apply a single distribution for all lots. The second approach is to apply individual distributions for each product type. The third approach is to apply individual distributions for all products (Rose 2007).

The third approach found in the literature used for toolset modeling is labeled the non-reentrant approach. In this approach, the modeling structure takes into account only the internal design of a process area and it does not include the delays as in the previous modeling approaches described. For this reason, non-reentrant models do not have the shortcomings of the reentrant simple models such as overtaking. As an example of the non-reentrant approach, Spence and Welter (1987) examine the photolithography area with cycle time-throughput trade-off curves. They assume Poisson arrivals with multiple input streams together with exponential machine failures and repair times. Also, they use the triangular distribution for process times with standardized lot sizes of 20 wafers. In another study of the photolithography area, Mönch, Prause, and Schmalfass (2001) apply historical fab data for machine failures and preventive maintenance and consider lot priorities from an ERP (Enterprise Resource Planning) system. They use WIP distributions obtained from the MES (Manufacturing Execution System) to initialize the simulation model. A key point on the models that use the non-reentrant modeling approach is the generation of WIP build-ups before the toolset being analyzed to reflect the arrival pattern into the process area correctly. Byrne, Heavey, and Kabak (2007) apply a detailed generation of the WIP profile for lot arrivals in a highly detailed simulation model of the photolithography area. In this approach, the conceptual model includes four modules, these are: Arrival, Dispatching, Toolset and Exit modules as shown in Figure 3.

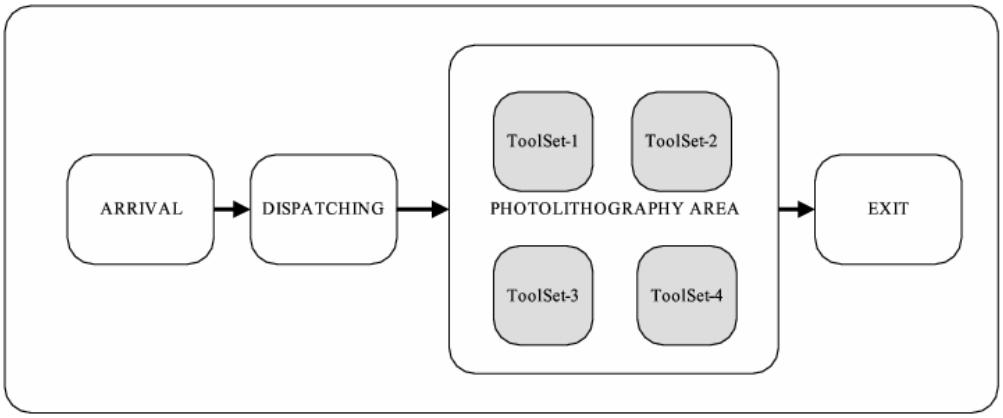



Figure 3: Conceptual model of non-reentrant approach.

The Arrival module creates a WIP profile based on historical data of the fab. The WIP profile is created by taking into account the process mix, lot priorities, recipes, number of passes for each process and the number of alignments per day in the photolithography area from historical fab data. In the second module, the lots are dispatched according to lot priorities, queueing discipline and the availability of the tools and auxiliary resources, e.g., reticles for the case of the photolithography area. The reticle sets are specific to each product type and layer. It is assumed that only one reticle set is available for each product type. The photolithography area consists of a number of toolsets that have different processing capabilities. In other words, toolsets process different sets of recipes. In the Exit module, performance measures of interest are measured and recorded.

## 3 TOOLSET MODELING APPROACHES WITH HISTORICAL FAB DATA

To compare the three different modeling approaches, a photolithography toolset from an ASIC fab is used as a testbed. This toolset together with data input common to the all three simulation models is described in Subsection 3.1. The dataset having 19 different process flows is obtained from an ASIC fab. The three simulation models were developed using Tecnomatix Plant Simulation 9.0.1 software (Siemens. 2010) using a existing ASIC fab dataset.

### 3.1 Photolithography Process

The photolithography process consists of four main steps, which are coat, expose, develop and post-photolithography analytical operations. In coat, the wafer is coated with a photoresist material, which is a type of light-sensitive polymer. Then in the expose step, the wafer is exposed with ultraviolet light (UV) in order to print the circuit pattern onto the wafer. This is done using a reticle, which is a chrome patterned glass that defines the circuit pattern. This pattern tends to be unique for each layer. In develop, the exposed sections are dissolved away in order to remove the photoresist material on the exposed sections of the wafer. Finally in post-photolithography analytical operations, wafers are inspected manually, which is dependent on both the product and layer. The photolithography process modelled has a total of nine toolsets. Further, these toolsets are grouped into two different groups according to two different mask sizes (i.e., type-1 and type-2). The first toolset group which has a mask size of type-1 consists of six toolsets. The second toolset group which has a mask size of type-2 consists of three toolsets. Also, each toolset has a different set of recipes.

The simulation inputs for the simulation models include part input mix, process flows, distributions for daily wafer quantity, lot priority and lot size, process times of toolsets, failures of toolsets and capabilities of toolsets. A representative of a three-month part mix was obtained from an existing ASIC fab. The part mix includes a total of 61 different products with 19 different process flows. The number of layers for each process flow is between 15 and 45, and the average number of layers is 27. The first six process flows (i.e., P1-P6) have a mask size of type-2 on the photolithography process. These processes constitute a total of 24.56 % of the part input mix. The rest of the process flows require a mask size of type-1 and they constitute 75.44 % of product mix. The daily amount of wafers released is determined by the pilot simulation runs for the reentrant simple and reentrant advanced models. Accordingly, an average of daily start volume of 170 wafers per day is determined by the pilot simulation runs. The fab from which the photolithography area data is obtained releases 5000 wafers a month. This is equivalent to an average daily start volume of 178 wafers per day. The daily amount of wafers released per day is determined from an empirical distribution. Five different lot priority levels exist in the wafer fab. These lot priority levels are numerated from 0 to 4 with decreasing precedence. In other words, the lots with priority level of 0 have the highest importance over all lots, the lots with priority level of 1 have the precedence over the lots with priority levels of 2, 3 and 4. Within each of the 5 priority classes, the lots are dispatched according to First-In First Out (FIFO) rule. An empirical distribution is used for lot priorities with lot frequencies.

An empirical distribution is used for lot sizes in the fab. In the fab, lot sizes are between 1 wafer and 25 wafers. The process times for photolithography toolsets are modelled by a gamma distribution with parameters of $\alpha = 4240$ and $\beta = 0.846$ (Mean of the dataset is 525.4 mins and std. dev is 540.5). The second toolset group processes the first six processes(i.e., P1 to P6). The rest of the processes are processed by the first toolset group. Also, the toolsets in the photolithography area have different capabilities according to the recipe operations. Two separate failure sets are defined for the first and second toolsets respectively. The simulation length is 420 days and the warm-up determined according to Welch's procedure is 60 days for the simulation models.

## 4 RESULTS

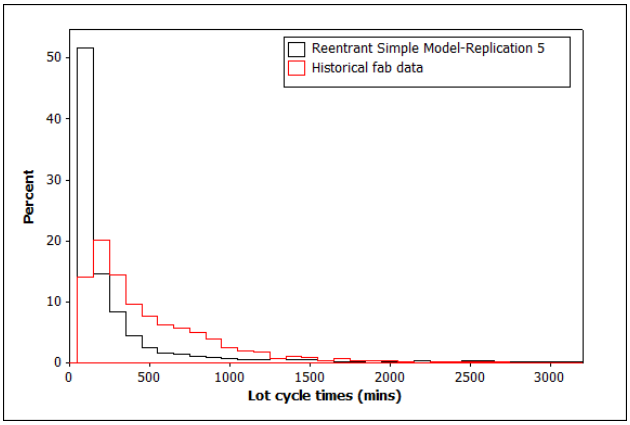
With regard to lot cycle times, Figure 4(a) compares the lot cycle time pattern obtained from the reentrant simple model based on a single representative replication with the lot cycle time pattern obtained from historical fab data. Figure 4(a) shows that the reentrant simple model shows a higher percentage of small lot cycle times than the historical fab data. It is noted that a similar conclusion is reported by Rose (2007). Rose (2007) compares flow factor and cycle time densities under three improved versions of the simple fab models with the full fab model. He concludes that the improved simple models are superior to the simple fab models since they show lower cycle time deviations. Figure 4(b) illustrates a comparison of lot cycle time patterns for the reentrant simple model and reentrant advanced model with historical fab data based on a single representative replication. It is observed that the graph of lot cycle times in the reentrant advanced model gives closer results to the historical fab data since the patterns of lot cycle times are closer to the pattern of the fab data. Figure 4(c) illustrates a comparison of the lot cycle times under the non-reentrant model and the historical fab data based on a single representative replication. It is shown that non-reentrant model gives close results to historical fab data in comparison to the other models. Similar lot cycle time patterns are obtained under all replications.

The main issues (see Section 2) with simple fab models are (1) correlation between the different cycles (i.e., stages or passes) for the same product; (2) lot overtaking. In Figure 5(a) and Figure 5(b), the y-axis shows the correlation for the observations and the x-axis shows the lags from 1 until 24. When compared to the reentrant simple model, less autocorrelation values are observed (see Figure 5(b)). The main reason for the correlation is the high percentage of lots overtaking at delay-II observed in the reentrant simple model. A similar conclusion is also reported by Rose (2000) (See also results of Q-statistics for autocorrelation tests in Table 2 in Appendices 6).

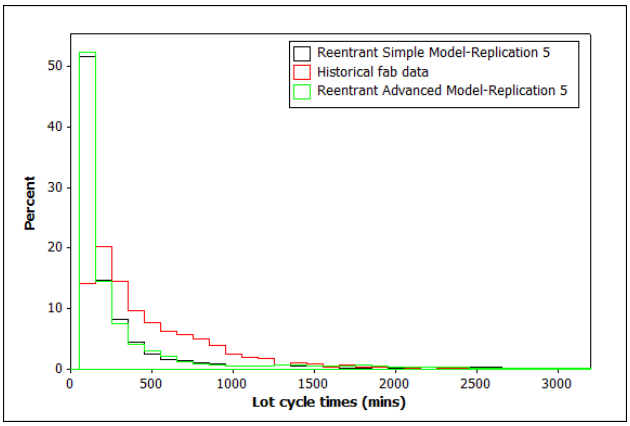
Table 1 gives a comparison of number of lot overtakings in the reentrant simple model and the reentrant advanced model for process P3. According to Table 1, the number of lot overtakings is significantly reduced in the reentrant advanced model. In Table 1, the cycle number shows the stage or pass number. The average difference for the number of overtakings is 131 and the standard deviation is 30,8. Accordingly, the lot overtaking in the reentrant advanced model is more representative of the real system. Insignificant number of overtakings was observed in the fab dataset.

Figure 6 gives a comparison of average cycle times of toolsets under the reentrant simple model, reentrant advanced model, non-reentrant model and historical fab data.

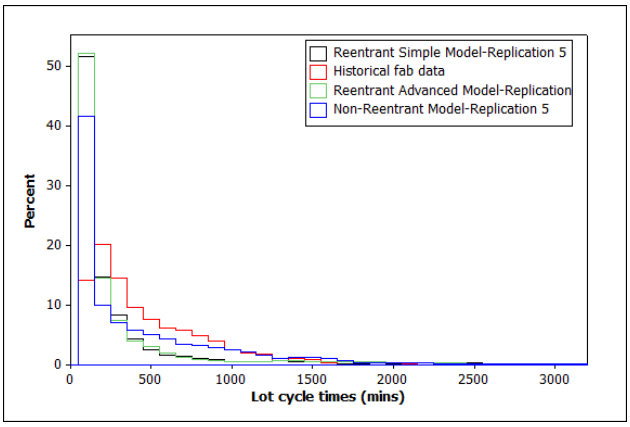
Figure 6 shows that the reentrant simple and reentrant advanced models give similar average toolset cycle times, with the average, the average toolset cycle times in both models showing high differences in comparison to the average toolset cycle times in the historical fab data, particularly on toolsets 5 and 8. In contrast, the differences between the average toolset cycle times in the non-reentrant model and historical fab data are not high and the non-reentrant model represents the photolithography cycle times better than the reentrant simple and reentrant advanced models as shown in Figure 6.




(a) Comparison of lot cycle times based on single replication from the reentrant simple model with the fab data.

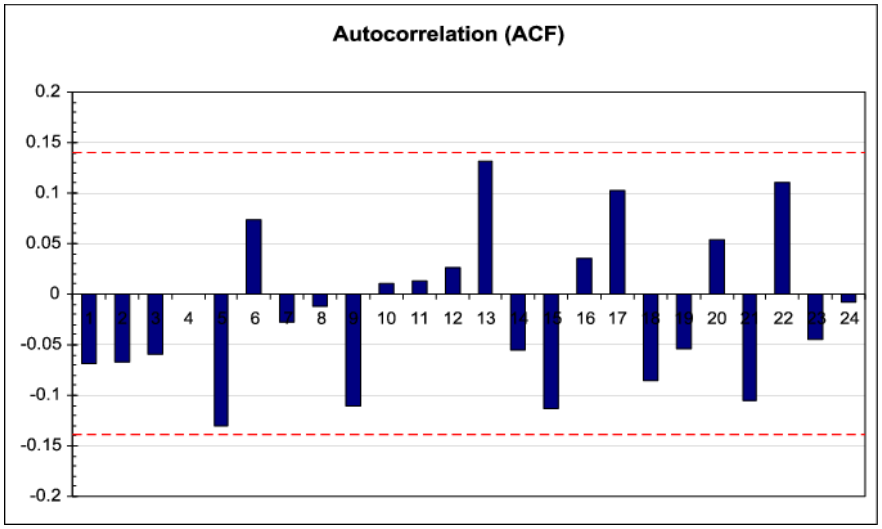



(b) Comparison of lot cycle times in the reentrant advanced model based on single replication with lot cycle times in the reentrant simple model and the historical fab data.

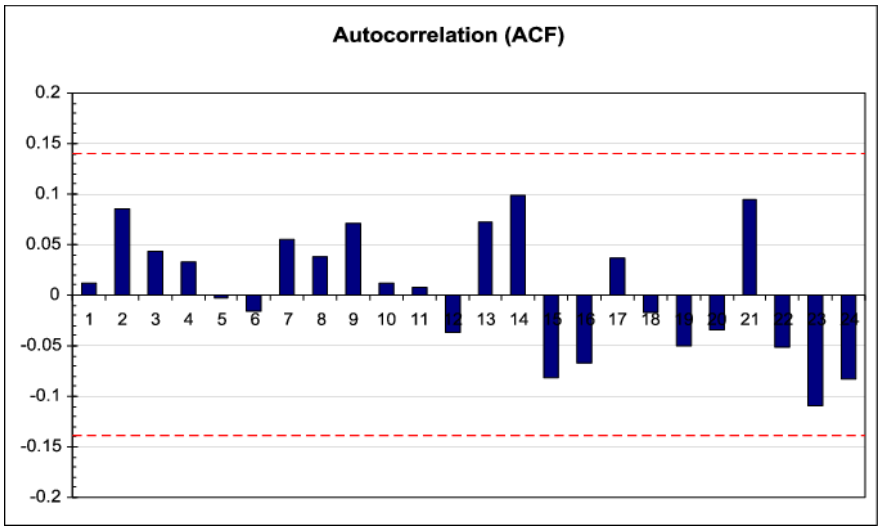



(c) Comparison of lot cycle times in the non-reentrant model based on single replication with lot cycle times in the the historical fab data.

Figure 4: Comparison of the simulation models.




(a) Autocorrelation analysis of the part in the second cycle based on single replication in the reentrant simple model.




(b) Autocorrelation analysis of the part in the second cycle based on single replication in the reentrant advanced model.

Figure 5: Autocorrelation analysis.

Table 1: Number of lot overtakings in the reentrant simple model and the reentrant advanced model.

| Cycle no | Simple model | Advanced model | Cycle no | Simple model | Advanced model |
|---|---|---|---|---|---|
| #2 | 41 | 40 | #19 | 174 | 21 |
| #3 | 199 | 23 | #20 | 167 | 14 |
| #4 | 178 | 17 | #21 | 127 | 8 |
| #5 | 189 | 21 | #22 | 150 | 18 |
| #6 | 171 | 3 | #23 | 123 | 6 |
| #7 | 151 | 16 | #24 | 149 | 9 |
| #8 | 166 | 8 | #25 | 114 | 18 |
| #9 | 195 | 20 | #26 | 119 | 9 |
| #10 | 165 | 18 | #27 | 131 | 17 |
| #11 | 170 | 19 | #28 | 129 | 22 |
| #12 | 176 | 5 | #29 | 119 | 4 |
| #13 | 140 | 13 | #30 | 128 | 104 |
| #14 | 160 | 14 | #31 | 135 | 26 |
| #15 | 120 | 17 | #32 | 118 | 14 |
| #16 | 143 | 13 | #33 | 116 | 16 |
| #17 | 155 | 29 | #34 | 125 | 6 |
| #18 | 166 | 20 | | | |

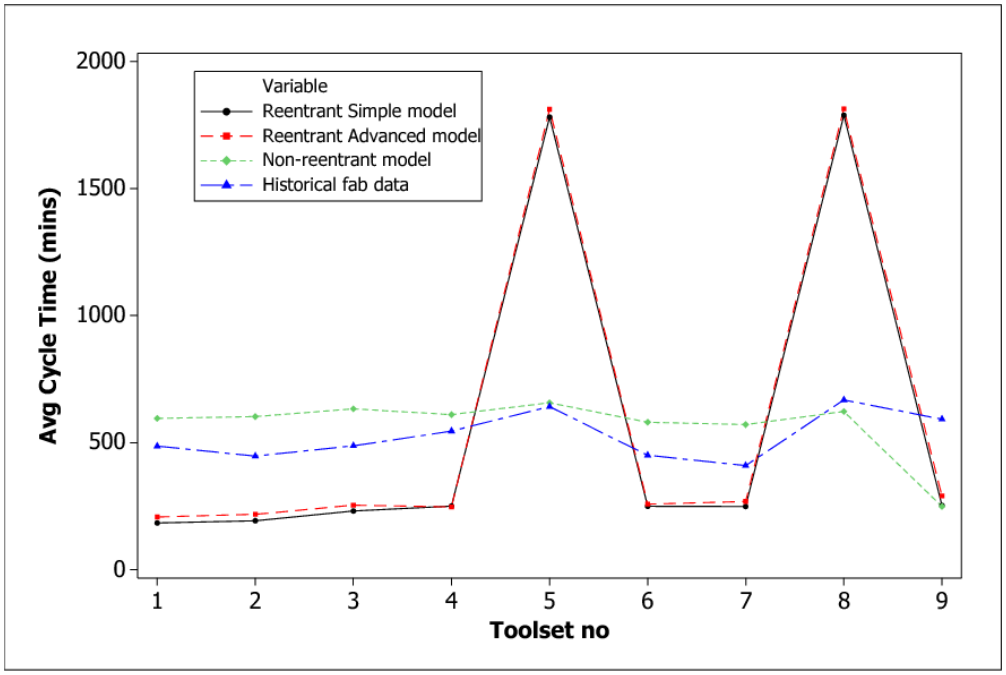



Figure 6: Comparison of average toolset cycle times based on average of five replications.

## 5 CONCLUSIONS

In this study, different approaches for simulation modelling of single toolsets in semiconductor manufacturing are first discussed systematically and compared and classified as reentrant and non-reentrant approaches. Reentrant approaches in the literature are further classified as reentrant simple and reentrant advanced models. The simulation models were compared using input derived from historical ASIC fab data at a given fab loading level. The results show that the reentrant simple model has shortcomings with regard to correlation and lot overtaking. These shortcomings are reduced in the reentrant advanced model and nonreentrant models, however, there are differences between the reentrant advanced model and nonreentrant models when the average toolset cycle times are compared with historical data.

In comparison to the reentrant models, the non-reentrant model does not represent explicitly the cyclic behavior of lots. Instead, it creates a WIP profile for lot arrival into the photolithography area. It has model development and experimentation advantages over the reentrant models as there is only one source of lot arrival generated from the WIP profile. As a result, it is easier to control the inputs to the toolset in particular with regard to process loading. While the results in this paper are limited, the results presented show that the nonreentrant model provides better or as good results when compared to the advanced nonreentrant approach, meriting its consideration by researchers and practitioners for single toolset studies.

## 6 APPENDICES

Table 2: Results of Q-Statistics (i.e., The Ljung-Box-Statistics).

| Lag | Q-Stat1 | Q-Stat2 |
|---|---|---|
| 1 | 0,92 | 0,03 |
| 2 | 1,81 | 1,49 |
| 3 | 2,52 | 1,86 |
| 4 | 2,52 | 2,07 |
| 5 | 6,00 | 2,07 |
| 6 | 7,11 | 2,12 |
| 7 | 7,26 | 2,75 |
| 8 | 7,29 | 3,06 |
| 9 | 9,82 | 4,09 |
| 10 | 9,84 | 4,12 |
| 11 | 9,88 | 4,13 |
| 12 | 10,03 | 4,42 |
| 13 | 13,67 | 5,54 |
| 14 | 14,31 | 7,64 |
| 15 | 17,03 | 9,07 |
| 16 | 17,29 | 10,05 |
| 17 | 19,57 | 10,36 |
| 18 | 21,17 | 10,42 |
| 19 | 21,82 | 10,97 |
| 20 | 22,47 | 11,24 |
| 21 | 24,90 | 13,21 |
| 22 | 27,62 | 13,79 |
| 23 | 28,08 | 16,48 |
| 24 | 28,09 | 18,03 |

## REFERENCES

Brooks, R. J., and A. M. Tobias. 2000. “Simplification in the simulation of manufacturing systems”. *International Journal of Production Research* 38 (5): 1009–1027.

Byrne, P. J., C. Heavey, and K. E. Kabak. 2007, December. “An analysis of tool capabilities in the photolithography area of an ASIC fab”. In *Proceedings of the 2007 Winter Simulation Conference*, edited by S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Jacobs, J., L. Etman, E. van Campen, and J. Rooda. 2003. "Characterization of operational time variability using effective process times". *IEEE Transactions on Semiconductor Manufacturing* 16 (3): 511–520.

Jacobs, J., P. van Bakel, L. Etman, and J. Rooda. 2006. "Quantifying variability of batching equipment using effective process times". *IEEE Transactions on Semiconductor Manufacturing,* 19 (2): 269–275.

Kabak, K. E., C. Heavey, and V. Corbett. 2008, December. "Analysis of multiple process flows in an ASIC fab with a detailed photolithography area model". In *Proceedings of the 2008 Winter Simulation Conference*, edited by S. J. Mason, R. R. Hill, L. Moench, O. Rose, T. Jefferson, and J. W. Fowler, 2185–2193. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Mönch, L., M. Prause, and V. Schmalfass. 2001, December. "Simulation-based solution of load-balancing problems in the photolithography area of a semiconductor wafer fabrication facility". In *Proceedings of the 2001 Winter Simulation Conference*, edited by B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, Volume 2, 1170–1177. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Peikert, A., J. Thoma, and S. Brown. 1998, December. "A rapid modeling technique for measurable improvements in factory performance". In *Proceedings of the 1998 Winter Simulation Conference*, edited by D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan, Volume 2, 1011–1015. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Rose, O. 1998, December. "WIP evolution of a semiconductor factory after a bottleneck workcenter breakdown". In *Proceedings of the 1998 Winter Simulation Conference*, edited by D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan, 1481–1490. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Rose, O. 1999, December. "CONLOAD-a new lot release rule for semiconductor wafer fabs". In *Proceedings of the 1999 Winter Simulation Conference*, edited by P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. Evans, Volume 1, 850–855. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Rose, O. 2000, December. "Why do simple wafer fab models fail in certain scenarios?". In *Proceedings of the 2000 Winter Simulation Conference*, edited by J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, 1481–1490. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Rose, O. 2007, December. "Improved simple simulation models for semiconductor wafer factories". In *Proceedings of the 2007 Winter Simulation Conference*, edited by S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 1708–1712. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Siemens. 2010. "Tecnomatix Plant Simulation 9.0.1". Siemens Product Lifecycle Management Software Inc. Accessed October 1, 2010. http://www.plm.automation.siemens.com/en_us/products/tecnomatix/plant_design/plant_simulation.shtml.

Spence, A., and D. Welter. 1987. "Capacity planning of a photolithography work cell in a wafer manufacturing line". In *IEEE International Conference on Robotics and Automation*, Volume 4, 702–708.

Veeger, C., L. Etman, E. Lefeber, I. Adan, J. van Herk, and J. Rooda. 2011. "Predicting Cycle Time Distributions for Integrated Processing Workstations: An Aggregate Modeling Approach". *IEEE Transactions on Semiconductor Manufacturing* 24 (2): 223–236.

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

## AUTHOR BIOGRAPHIES

**KAMIL ERKAN KABAK** is an assistant professor in the Department of Industrial Engineering at the Beykent University, Turkey. He received his bachelor degree in the Department of Environmental Engineering from Middle East Technical University, Turkey. Later, he obtained his master degree from the Department of Industrial Engineering, Dokuz Eylul University, Turkey. His research interests include instability of production plans, simulation modeling along with its applications to analysis and control methods for semiconductor manufacturing. He is a member of INFORMS and Operational Research Society of

Turkey. He can be reached via email at erkankabak@gmail.com.

**CATHAL HEAVEY** is a Senior Lecturer of Operations Management in the Department of Manufacturing and Operations Engineering at the University of Limerick. He is an Industrial Engineering graduate of the National University of Ireland (University College Galway) and holds a M. Eng.Sc. and Ph.D. from the same University. He has published in the areas of queuing and simulation modeling. Research interests are: Simulation Modeling of Discrete Event Systems; Modeling and Analysis of Supply Chains and Manufacturing Systems; Process modeling; Component- based simulation; Decision support systems. His email address is cathal.heavey@ul.ie.

**BRIAN KERNAN** is a researcher and software developer at Entellexi at the National Technology Park, Ireland. He received his PhD from the University of Limerick in January 2012, with his doctoral research focusing on enterprise simulation models and human performance modeling. Additionally, he has several years' experience as a lecturer in manufacturing technology at the University of Limerick, and as a private consultant and software developer for Irish industry. His email address is brian.kernan@gmail.com.