Proceedings of the 2019 Winter Simulation Conference
N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, eds.

A NEW HIGH-VOLUME/LOW-MIX SIMULATION TESTBED FOR SEMICONDUCTOR MANUFACTURING

Michael Hassoun

Dept. of Industrial Engineering and Management
Ariel University
Milken Campus
Ariel, 40700, ISRAEL

Denny Kopp
Lars Mönch

Dept. of Mathematics and Computer Science
University of Hagen
Universitätsstraße 1
Hagen, 58097, GERMANY

Adar Kalir

Intel Corporation and Dept. of Industrial Engineering
Ben-Gurion University
Mailstop: LC2-3-ME
Qiriat-Gat, 82109, ISRAEL

ABSTRACT

As part of an effort to present to the semiconductor manufacturing community an updated wafer fab testbed, we provide the first of two simulation models, namely a High-Volume/Low-Mix (HV/LM) fab simulation model. The model is realistic in scale and level of complexity. A full description of the model features is provided, and its performance is studied based on an implementation using the AutoSched AP simulation tool. The simulation model is made publicly available online to allow researchers as well as practitioners to gain hands-on experience with it, and hopefully validate its features, or propose changes if needed. A final version of the testbed, including a low-volume/high-mix wafer fab simulation model will be presented within a year.

1. INTRODUCTION

The complexity associated with semiconductor manufacturing has driven researchers as well as companies to make extensive use of simulation (Fowler et al. 2015). Models used by practitioners usually tend to represent as faithful as possible the wafer fab at hand. Long-term phenomena can be studied using multiple scenarios and replications, while urgent tactical and operational decisions can be made using the current fab situation to initialize the simulation model. Various full factory simulation studies have been conducted by practitioners using specific simulation settings, e.g. on production planning, and on factory performance in a transient period (Klein and Kalir 2006). These have provided useful insight – but were limited in the ability to share, replicate, and compare results with others. On their side, researchers are bound to look for more generic models to study phenomena and policies at the different decision levels. But developing such models, each for the purposes of a unique research team has strong disadvantages. First, it bears the risk for the results to be strongly model-dependent. Moreover, it makes it difficult, sometimes even impossible, for other teams to reproduce the simulation results. Last but not least, the conclusions reached on the somehow simplified models used by academic research teams are often received with suspicion by practitioners, who fear, often rightly so, that a critical level of complexity is lost in them and the results are not relevant for

actual fab operation. For all these reasons, the need for a standardized simulation testbed exists, and fills two major needs:

1. It serves as a common testing platform for researchers.
2. It represent a bridge between academia and industry.

The first of these requirements is the easier to achieve, and all that is needed is a level of details and plasticity that fits the need of most researchers. The second, however, requires a great deal of attention. The model details, resolution, features, and scale need to satisfy practitioners if the conclusions reached through the simulation are to be accepted as relevant to the industry. In the last two and a half decades, among many wafer fab simulation models (see Subsection 2.2) the Measurement and Improvement of Manufacturing Capacity (MIMAC) dataset, see Fowler and Robinson (1995) has been highly successful in filling the bill. Many research efforts are based on this set of six wafer fab simulation models. These venerable models form, till this day and nearly three decades after their publication, the best available testbed for simulation-based semiconductor manufacturing research, and, with no better alternative, remain the benchmark of choice for researchers, e.g. Shin et al. (2019). Yet, some of their features are outdated, others are lacking. Formulating a new set of models in the same spirit, but with features that better represent a state-of-the-art wafer fab, seems in order. This is proposed by Hassoun and Kalir (2017) who outline the main features of such a testbed. The current paper is the first part of the realization of this testbed. The testbed, denoted SMT2020 (for Semiconductor Manufacturing Testbed and the expected year of completion and publication), is made available at http://p2schedgen.fernuni-hagen.de/index.php?id=296. A full report on both simulation models is also provided. We encourage both researchers and practitioners to use the model and provide the authors with feedback necessary to improve the final testbed. A full scale High-Volume/Low-Mix fab, denoted as dataset 1, is hereby presented which includes a great deal of the current technology intricacies.

The rest of the paper is organized as follows. In the following section we formulate requirements for a semiconductor manufacturing simulation testbed and discuss related work. In Section 3, we describe the dataset 1 features. Then, in Section 4, the model performance is studied. Conclusions are presented in Section 5.

2. PROBLEM ANALYSIS AND LITERATURE REVIEW

2.1 Requirements for a New Wafer Fab Testbed

For the sake of completeness, we summarize the main functional requirements for the SMT2020 testbed derived by Hassoun and Kalir (2017):

1. **Types of wafer fabs, products, and configurations:** The testbed must contain simulation models for low-volume/high-mix (LV/HM) and high-volume/low-mix (HV/LM) situations. In the former situation, ten products have to be included where due dates exist for all lots to support a make-to-order (MTO) environment. In the later situation, two products should be included, each of them at high volume to reflect the conditions of a make-to-stock (MTS) environment. The number of process steps per route should be more than 500 to reflect the complexity of modern wafer fabs. The number of tool groups should be more than 100 for the same reason.
2. **Availability modeling:** Both preventative maintenance (PM) actions and tool breakdowns must be included to model scheduled and unscheduled downtime.
3. **Machine rate modeling:** Batching tools and cascading tools have to be modeled to reflect process conditions in real-world wafer fabs. Unload and load times on the tools and transportation times between tool groups have to be represented in the models. Moreover, setups, especially sequence-dependent setups have to be represented in detail.

4. **Functional area specific complexities:** Important features of specific functional areas such as lot-to-lens dedication on steppers in the photolithography area, correct sampling rates on metrology tools, critical queue time constraints for the dry etch area, and correct modeling of PM and setup activities on implanters are important.
5. **Operational complexities:** Hot lots have to be included into the simulation models of the testbed. Moreover, since up to one third of all WIP lots can belong to engineering lots (Crist and Uzsoy 2011) it is important to represent them into the simulation models of the testbed. Critical Queue Times (CQT) have to be defined in the routes because they are typical for many modern wafer fabs (Klemmt and Mönch 2012).

Moreover, we identify the following two requirements that are related to the usage of the new testbed:

1. **Representation:** The simulation models of the testbed should be represented in such a way that users can easily use the datasets without relying on a specific simulation tool. ASCII or Excel files with a format similar to the MIMAC datasets are used to represent the models.
2. **Performance measure values:** The values of important performance measures including availability per functional area, tool utilization, major bottlenecks, and an operating curve have to be reported for users to check the correctness of their usage of the corresponding simulation model.

2.2 Discussion of Related Work

Discrete-event simulation is an important tool for decision-making in wafer fabs (cf. Mönch et al. 2013). It is well-known that building a full fab simulation model from scratch is time-consuming and error-prone (Fowler et al. 2015). Therefore, researchers use wafer fab simulation testbeds. The following four testbeds or simulation models are often applied by researchers:

1. The first one is the MIMAC testbed (Feigin et al. 1994; Fowler and Robinson 1995). It consists of six wafer fab models of different complexity. Up to 260 machines are included. In addition, up to 21 products with routes that have up to 340 process steps exist in these models. Various applications of the testbed datasets are known. For instance, Mönch and Zimmermann (2011) use the MIMAC 1 dataset to assess deterministic scheduling approaches in a rolling horizon setting. Multiple orders per job formation and release strategies for wafer fabs are investigated by Mönch et al. (2011) based on the MIMAC 1 dataset. The MIMAC 1 dataset is also used to assess the quality of release schedules for wafer fabs by executing them in the simulation model (Kacar et al. 2013; Kacar et al. 2016). Kopp and Mönch (2018) apply the MIMAC 1 dataset to evaluate qualification management decisions in a rolling horizon setting. Mittler and Schoemig (2000) and Zhou and Rose (2011) tackle dispatching problems using the MIMAC 6 model. The MIMAC 6 dataset is also used by Hildebrandt et al. (2014) and Kück et al. (2017) to discover dispatching rules based on genetic programming. Engineering lots are introduced into the MIMAC 1 dataset by Ziernetzky and Mönch (2016). Hassoun (2013) uses the MIMAC data sets to study methods for cycle time (CT) prediction in wafer fabs.
2. The second model is the so-called MiniFab model proposed by Spier and Kempf (1995). It is a low complexity simulation model that mimics the typical behavior of a wafer fab by containing reentrant process flows, batch processing, and significant sequence-dependent setup times. It is applied, for instance by Pfund et al. (2008) to derive problem instances to test a scheduling approach for wafer fabs. It is also used by Driessel and Mönch (2012) to assess an integrated scheduling and automated material handling approach in a rolling horizon setting.
3. The third model is a scaled-down representation of a real-world wafer fab at Harris Semiconductor described by Kayton et al. (1997). This model is applied, for instance, by Kacar and Uzsoy (2015) to estimate clearing functions using simulation-based optimization. Production planning

approaches for wafer fabs are also assessed in a rolling horizon setting based on this simulation model in (Ziernetzky et al. 2015; Ziernetzky et al. 2018).

4. The fourth model is a generic 300 mm model including automated material handing features proposed by SEMATECH (Campbell and Ammenheuser 2000). It is used, for instance, by Upasani et al. (2006) to generate problem instances for global fab-wide scheduling approaches. This model is less frequent applied compared to the models discussed before.

The main characteristics of existing testbeds and simulation models are summarized in Table 1. Note that the discussed testbeds and simulation models except the SEMATECH 300mm model are publicly available at http://p2schedgen.fernuni-hagen.de/index.php?id=296.

Table 1: Characterization of the testbeds/wafer fab simulation models.

Testbed/simulation model	# Machines/workcenters	# Products	#Process steps
MiniFab	5/3	2	6
Harris	12/11	3	up to 22
MIMAC datasets	up to 260/up to 85	up to 21	up to 280
SEMATECH 300mm model	275/103	1	364

In addition, a few efforts exist to provide testbeds for entire semiconductor supply chains. A semiconductor supply chain testbed is proposed by Ewen et al. (2017). Its front-end nodes are based on the MIMAC dataset 1. A generic data model for semiconductor supply chains is described by Laipple et al. (2018).

Overall, we see from this discussion that the already existing testbed and simulation models, despite the fact that they were extensively used in the past, do not provide the complexity that is typical for modern wafer fabs (Hassoun and Kalir 2017).

3. MODEL CHARACTERISTICS

In this section, we describe the SMT2020 dataset 1 fab simulation model in as much details as possible. The model will be provided by Excel sheets at http://p2schedgen.fernuni-hagen.de/index.php?id=296, and additional details are summarized in the file "Instructions.pdf" to allow readers a seamless transition from the article to using the model itself.

3.1. General Structure

The first important feature of the proposed model is its realistic scale. Thinking beyond the immediate need of an HV/LM fab (dataset 1), we have defined ten products whose route lengths are up to 632 operations. Three of the routes are longer and seven shorter. In the framework of dataset 1, only routes 3 and 4 are used, whose lengths are 632 and 374 steps, organized in 44 and 30 mask layers, respectively. The volume of both products is kept stable at 5100 wafer starts per week, from which 2.5% are hot lots. This translates to the introduction of a new regular lot every 50.68 minutes, and of a hot lot every 1976.47 minutes. A total of 105 toolsets are modeled, organized in the following areas:

- diffusion
- wet etch
- dry etch
- implant
- lithography
- dielectric
- planarization
- thin films

- lithography metrology
- thin films metrology
- defectivity metrology.

The three last areas perform quality control tasks along the process. The machine rates are defined using a processing time for a single wafer, a lot, or, in the diffusion area for a batch. A lot is an indivisible bundle of 25 wafers, while a batch is a group of lots that is processed at the same time on a single tool, usually a diffusion tool. A fixed duration for the lot or batch loading and unloading is added. During simulation, the processing time is sampled from a uniform distribution whose range is 10% of the nominal value (+- 5% around the nominal value). Transport of lots between toolsets is represented by a simple delay sampled from a uniform distribution on a range from 5 to 10 minutes.

All tools in wet etch, implant, planar and thin films, and about half of the tools in dielectric area allow for cascading. These tools have an inner transfer structure allowing different parts, i.e. wafers or lots, to work in parallel. The practical meaning of this is that for the first lot (or wafer in case of a single wafer processing tool) of a sequence, or "cascade" of lots/wafers (regardless of their operation), the time required to finish the process is longer than for the following lots/wafers in the cascade. Since processing goes on without interruption, it also means that the loading and unloading phases are happening in parallel with processing and are therefore accounted for only once for a "cascade" of lots/wafers.

Diffusion tools use batches. The maximum batch size (technological constraint) is between 100 and 150 wafers depending on operations. To ease the tool operation, an allowance of one lot is tolerated. For example, if the maximum batch size is 125 wafers (5 lots), a batch can be processed as soon as four lots are ready, waiting in the toolset queue.

Often, high volume manufacturing is associated with production to stock. Therefore, we do not represent any due date for the lots and adopt a simple First-Come First-Served (FCFS) dispatching rule along the process, with several exceptions that will be discussed in the following subsections.

3.2. Availability Modeling

One of the important novelties of SMT2020 is certainly the attention given to machine availability modeling. Both scheduled down time (SDT) which refers to preventive maintenance and unscheduled down time (UDT) or breakages, are modeled in detail. The balance between UDT and SDT is carefully chosen. In the lithography for example, where reliability is an issue, the target for UDT is set at 50% of the total down time, while in much more stable areas like metrology or planar, we aim at 90% of the total down time accounted by SDT and 10% by UDT.

3.2.1. SDT

The PMs are modeled on three levels of duration and frequency. For diffusion, lithography, and all the metrology, the maintenance is time-based. All tools undergo a medium duration monthly PM and a long quarterly PM. Tools in the lithography area undergo an additional short weekly PM. All remaining functional areas undergo counter-based PMs. The counters are set to trigger a short duration PM every week, a medium duration PM every month, and a long duration PM approximatively every quarter.

3.2.2. UDT

Breakages being of various types, they differ greatly in frequency and duration. We adopt an exponential distribution both for the Time Before Failure (TBF) with a mean (MTBF) of one week across the board and for the Time To Repair (TTR). The mean duration for the TTR (MTTR) is computed based on the assumptions on the target availability for the area and the balance between UDT and SDT.

3.3. Metrology

The sampling rate on metrology tools depends on the area and the operation location along the process, About half of the lithography metrology operations have a 100% rate (all lots are inspected), the rest have a measurement rate between 10% and 50%. Thin films metrology steps are characterized by measurement rate between the 30% and the 40% marks. In the first half of the process (front end), defects metrology has a sampling rate set between 50% and 60%. For the second half (back end), which is technologically less sensitive, a sampling rate of 10% is adopted.

Rework loops are defined for about one third of the lithography operations, with rework rates ranging from 0.5% to 2%.

We do not model scrapping of either full lots or single wafers since apart from slightly changing processing times for wafer-based process tools it does not seem to add any value in terms of managerial complexity.

3.4. Constraints

With batching and availability issues, setups are known to critically increase the WIP flow variability in the fab by forcing managers to run long trains of the same operation on a machine before switching to another. Setups are represented in three areas. In lithography, all tools have setups (reticle change), ranging from 5 to 20 minutes in duration regardless of the operation. Two dry-etch tools in the back end of the process have simple sequence-dependent setups (limited to two groups of operations). In the implant area, sequence-dependent setups are defined with three groups, one for each gas.

Three groups of lithography operations, on three different stepper types, are running under a lot-to-lens dedication (LTLD) scheme, meaning that once a lot is processed on a specific tool on the first of these operations, it must be processed on the same tool for the remaining operations in the group. This is done typically to lower the impact on quality associated with overlay discrepancies at critical steps of the process. CQTs are defined, but at this point no throttling mechanism at the entrance of the CQT segment.

3.5. Scheduling

As mentioned earlier, the baseline lot dispatching method is FCFS. For tools with setups, a setup avoidance scheme is used that exhausts any available WIP of a certain operation before converting the tool to another one. Hot lots are defined (2.5% of all lots) that take precedence on all other lots, and when required, trigger setups and break cascades.

4. MODEL PERFORMANCE

An overview of the characteristics and performance measures of each functional area of SMT2020 dataset 1 is presented in Table 2.

These are reviewed in more details in the report that will be provided online. The first two columns provide the number of toolsets and total number of tools in the area. Because we avoid modeling complex availability and machine rate schemes that allow partial machine functions or hybrid tools such as in dry etch and thin films, individual entities (chambers) are defined as tools. Also, a minimum number of two machines is required for each toolsets to ensure, as customary in wafer fabs, a minimum tool redundancy.

Table 2: Main characteristic and performance measures of SMT2020 – Dataset 1.

Area	Population		Availability		Utilization	
	Number of tool types	Number of tools	Average availability	Proportion of Scheduled DT	Average Utilization	Maximum Utilization
Dielectric	10	61	80.4%	71.2%	82.8%	89.6%
Diffusion	10	73	92.5%	80.3%	83.8%	89.7%
Dry Etch	21	366	88.5%	80.2%	84.1%	89.8%
Implant	9	36	79.6%	72.1%	55.7%	89.0%
Lithography	11	207	86.2%	50.9%	85.6%	96.3%
Planar	6	33	80.0%	89.9%	71.2%	89.4%
Thin Films	11	89	85.3%	70.7%	81.9%	89.6%
Wet Etch	14	113	89.0%	80.2%	78.2%	89.5%
Defectivity	7	17	96.5%	90.2%	49.5%	83.8%
Litho Metrology	4	55	96.5%	90.1%	87.3%	89.4%
Thin Films Metrology	2	4	96.5%	90.1%	46.0%	74.9%
total	105	1054				

4.1. Availability Utilization and Bottlenecks

Tool availability is the result of a set of decisions about the duration (or frequency) of PMs and the relative proportion of SDT, which is representative of the level of availability randomness and has an important impact on the variability of the tool's capacity. In Table 2, we provide the average availability in each area and the proportion of SDT. Metrology tools present the highest availability and are also more stable in the fab, with only about 10% of their downtime attributed to breakdowns. On the other end of the scale, tools in implant, dielectric, and planarization present availability figures around 80% on the average. Litho tools are the most unstable, with roughly half of the down time attributed to breakdowns.

By design, we try to make the lithography area be the fab constraint. Consequently, the seven highest toolset utilization figures, ranging from 93.3% to 98.5% are all found in lithography. The other extreme is found in toolsets where the sheer number of tools required is one, but because of the requirement for a redundancy, two tools are installed. This occasionally creates very low utilization levels. The ten lowest utilization toolsets all belong to this category.

4.2. Functional Curve

The behavior of the fab is represented by its operating curve. In order to unify results from two products, we normalize cycle time (CT) by each product's raw processing time (RPT), yielding what is known as Flow Factor (FF) (Mönch et al. 2013). As a preliminary step, we first run the model with lots starting at very long intervals, resulting in them progressing in an empty fab. Their CT is measured, and averaged to obtain the RPT. The source of variability in the resultant CT is that some lots experience delays due to PMs, or are sampled in metrology steps. The resultant RPTs for products 3 and 4 are 28.7 days and 17.1 days, respectively. The FFs for a wide range of loading scenarios are calculated based on these values, and averaged. To illustrate the distribution of FF values, we present in Figure 1 the various percentiles of the FF curves as a function of the wafer starts. The lowest curve, with FF very close to 1.00 represent the fastest amongst hot lots. We note that these hot lots are slightly affected by the fab loading, which may seem counter-intuitive. This is likely due to the fact that although their priority is the highest of all lots, they still have to wait if all tools are utilized upon their arrival. It may also be explained by multiple non-synchronized

PMs on tools with counter-based PMs. At full capacity (the baseline model), the median of FF is 2.63 while the maximum reaches a value of 3.42. 90% of all lots have a FF between 2.39 and 2.91, corresponding to the 5% and 95% percentiles, respectively.

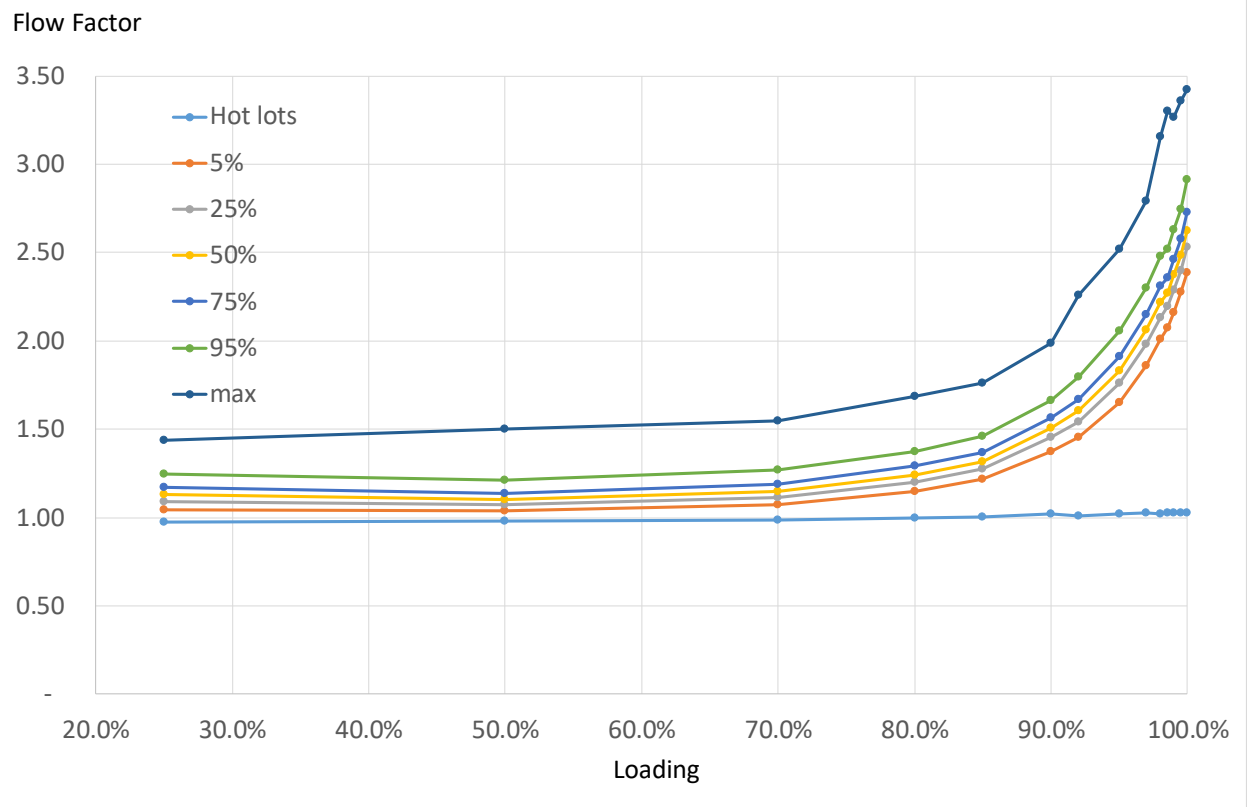

Figure 1: Functional curves: flow factor percentiles vs. wafer starts.

5. CONCLUSIONS

In an ongoing effort to provide the semiconductor community an up-to-date simulation testbed, we propose the first of two planned simulation models, namely, a HV/LM fab denoted SMT2020 dataset 1. The model, as well as a detailed report beyond what the space of this paper would allow will be made public at http://p2schedgen.fernuni-hagen.de/index.php?id=296. We consider this model as an advanced draft rather than a final model, and encourage members of the semiconductor manufacturing community to review the detailed report, use the model, and provide their insights to the authors. The next steps of this effort will be to consider and incorporate the received feedback, and finalize both models: the current dataset 1, and dataset 2 which simulates a LV/HM fab.

The latter will differ from dataset 1 in the number of products, i.e., ten products are defined rather than two. Note that their routes have commonalities with the two routes in dataset 1, and are already available online. The machines will be similar to those in dataset 1 but their number will be updated to be consistent with the new load requirements, and still reach reasonable utilization targets. Consistent with a MTO mode of operation, dataset 2 will run under a critical ratio dispatching rule. Due dates designed to match a target service level will be assigned to individual lots. As suggested by the name of this dataset, we hope to release and publish the final version of the SMT2020 testbed within a year.

ACKNOWLEDGMENTS

Parts of this research were carried out whilst the first author was visiting the Department of Mathematics and Computer Science of the University of Hagen, Germany, in March 2019. This research was partially supported by the iDev 4.0 project. This project is funded by the ECSEL Joint Undertaking (JU) under grant agreement No 783163. The second and fourth author gratefully acknowledge this financial support.

REFERENCES

Campbell, E. and J. Ammenheuser. 2000. "300mm Factory Layout and Material Handling Modeling: Phase II Report". SEMATECH Technical Transfer report #99113848BENG, Austin, Texas.

Crist, K. and R. Uzsoy. 2011. "Prioritizing Production and Engineering Lots in Wafer Fabrication Facilities: A Simulation Study". *International Journal of Production Research* 49(11):3105-3125.

Driessel, R. and L. Mönch. 2012. "An Integrated Scheduling and Material Handling Approach for Complex Job Shops: A Computational Study". *International Journal of Production Research* 50(20):5966-5985.

Ewen, H., L. Mönch, H. Ehm, T. Ponsignon, J. W. Fowler, and L. Forstner. 2017. "A Testbed for Simulating Semiconductor Supply Chains". *IEEE Transactions on Semiconductor Manufacturing* 30(3):293-305.

Feigin, G. E., J. W. Fowler, and R. C. Leachman. 1994. "Semiconductor Wafer Manufacturing Data Format Specification". SEMATECH Technical Report.

Fowler, J. W., and J. Robinson. 1995. "Measurement and improvement of manufacturing capacity (MIMAC) final report". Technology Transfer #95062861A-TR, SEMATECH.

Fowler, J. W., L. Mönch, and T. Ponsignon. 2015. "Discrete-event Simulation for Semiconductor Wafer Fabrication Facilities: A Tutorial". *International Journal of Industrial Engineering: Theory, Applications, and Practice* 22(5): 661-682.

Hassoun, M. 2013. "On Improving the Predictability of Cycle Time in an NVM Fab by Correct Segmentation of the Process". *IEEE Transactions on Semicinductor Manufacturing* 26(4):613 – 618.

Hassoun, M. and A. Kalir. 2017. "Towards a New Simulation Testbed for Semiconductor Manufacturing". In *Proceedings of the 2017 Winter Simulation Conference*, Edited by W.K.V. Chan, AD'ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer and E. Page, 3612-3623. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Hildebrandt, T., D. Goswami, and M. Freitag, 2014. "Large-scale Simulation-based Optimization of Semiconductor Dispatching Rules". In *Proceedings of the 2014 Winter Simulation Conference*. Edited by A. Tolk, S.Y. Diallo, I.O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 2580-2590. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Kacar, N. B., and R. Uzsoy. 2015. "Estimating Clearing Functions for Production Resources Using Simulation Optimization". *IEEE Transactions on Automation Science and Engineering* 12(2):539-552.

Kacar, N. B., L. Mönch, and R. Uzsoy. 2013. "Planning Wafer Starts using Nonlinear Clearing Functions: a Large-Scale Experiment". *IEEE Transactions on Semiconductor Manufacturing* 26(4):602-612.

Kacar, N. B., L. Mönch, and R. Uzsoy. 2016. "Modeling Cycle Times in Production Planning Models for Wafer Fabrication". *IEEE Transactions on Semiconductor Manufacturing* 29(2):153-167.

Kayton, D., T. Teyner, C. Schwartz, and R. Uzsoy. 1997. "Focusing Maintenance Improvement Efforts in a Wafer Fabrication Facility Operating under the Theory of Constraints". *Production and Inventory Management Journal* 38(4):51-77.

Klein, M. and A. Kalir. 2006. "A Full Factory Transient Simulation Model for the Analysis of Expected Performance in a Transition Period". In *Proceedings of the 2006 Winter Simulation Conference* . Edited by L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 1836-1839. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Klemmt, A. and L. Mönch. 2012. "Scheduling Jobs with Time Constraints between Consecutive Process Steps in Semiconductor Manufacturing". In *Proceedings of the 2012 Winter Simulation Conference*. Edited by C. Laroque, J. Himmelspach, R. Patsupathy, O. Rose, and A.M. Uhrmacher, 2173-2182. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Kopp, D. and L. Mönch. 2018. "A Hierarchical Approach to Qualification Management in Wafer Fabs". In *Proceedings of the 2018 Winter Simulation Conference*. Edited by M. Rabe, A.A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, 3514-3525. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Kück, M., E. Broda, M. Freitag, T. Hildebrandt, and E. M. Frazzon. 2017. "Towards Adaptive Simulation-based Optimization to Select Individual Dispatching Rules for Production Control". In *Proceedings of the 2017 Winter Simulation Conference*. Edited by W.K.V. Chan, AD'ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer and E. Page, 3852-3863. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Laipple, G., S. Dauzère-Pérès, T. Ponsignon, and P. Vialletelle. 2018. "Generic Data Model for Semiconductor Manufacturing Supply Chains". In *Proceedings of the 2018 Winter Simulation Conference*. Edited by M. Rabe, A.A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, 3615-3626. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Mittler, M. and A.K. Schoemig. 2000. "Comparison of dispatching rules for reducing the mean and variability of cycle times in semiconductor manufacturing". In *Operations Research Proceedings 1999*, 479-485. Springer, Berlin, Heidelberg.

Mönch, L. and J. Zimmermann. 2011. "A Computational Study of a Shifting Bottleneck Heuristic for Multi-Product Complex Job Shops". *Production Planning & Control* 22(1): 25-40.

Mönch, L., J. Zimmermann, S. Mason, and J. W. Fowler. 2011. "Multiple Orders per Job Formation and Release Strategies in Large Scale Wafer Fabs: A Simulation Study". *Journal of Simulation* 5(1): 25-43.

Mönch, L., J. W. Fowler, and S. J. Mason. 2013. *Production Planning and Control for Semiconductor Wafer Fabrication Facilities: Modeling, Analysis, and Systems.* New York: Springer.

Pfund, M. E., H. Balasubramanian, J. W. Fowler, S. J. Mason, and O. Rose, 2008. "A Multi-criteria Approach for Scheduling Semiconductor Wafer Fabrication Facilities". *Journal of Scheduling* 11: 29-47.

Shin, J., D. Grosbard,, J. R. Morrison, and A. Kalir. 2019. Decomposition without Aggregation for Performance Approximation in Queueing Network Models of Semiconductor Manufacturing. *International Journal of Production Research* (in press).

Spier, J. and K. G. Kempf. 1995. "Simulation of Emergent Behavior in Manufacturing Systems". In *Proceedings of the IEEE/SEMI Advanced Semiconductor Manufacturing Conference and Workshops*, 90-94. IEEE. Cambridge, MA.

Upasani, A. A., R. Uzsoy, and K. Sourirajan. 2006. "A Problem Reduction Approach for Scheduling Semiconductor Wafer Fabrication Facilities". *IEEE Transactions on Semiconductor Manufacturing* 19(2): 216-225.

Zhou, Z. and O. Rose. 2011. "A Composite Rule Combining Due Date Control and WIP Balance in a Wafer Fab". In *Proceedings of the 2011 Winter Simulation Conference*. Edited by S. Jain, R.R. Creasey, J. Himmelspach, K.P. White, and M. Fu, 2080-2087. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Ziarnetzky, T. and L. Mönch. 2016. "Incorporating Engineering Process Improvement Activities into Production Planning Formulations Using a Large-Scale Wafer Fab Model". *International Journal of Production Research* 54(21): 6416-6435.

Ziarnetzky, T., B. Kacar, L. Mönch, and R. Uzsoy. 2015. "Simulation-based Performance Assessment of Production Planning Formulations for Semiconductor Wafer Fabrication". In *Proceedings of the 2015 Winter Simulation Conference*. Edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 2884-2895. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Ziarnetzky, T., L. Mönch, and R. Uzsoy. 2018. "Rolling Horizon Production Planning for Wafer Fabs with Chance Constraints and Forecast Evolution". *International Journal of Production Research* 56(18): 6112-6134.

AUTHOR BIOGRAPHIES

MICHAEL HASSOUN is a lecturer at the Industrial Engineering Department of the Ariel University, Israel. His research interests focus on modeling and management of production systems, with a special interest in Semiconductor manufacturing. He earned his PhD and MSc in Industrial Engineering from Ben-Gurion University of the Negev, Israel, and his BSc in Mechanical Engineering from the Technion, Israel. He was a postdoctoral fellow at the University of Michigan in 2009. His email address is michaelh@ariel.ac.il.

DENNY KOPP is is a Ph.D. student at the Chair of Enterprise-wide Software Systems, University of Hagen. He received MS degree in Business Economics from the Otto-von-Guericke-University Magdeburg, Germany. His research interests include applied optimization and simulation-based production control. His email address is Denny.Kopp@fernuni-hagen.de.

ADAR KALIR received his B.S. and M.S. degrees in industrial engineering and management from Tel-Aviv University, Israel, and his Ph.D. degree in industrial and systems engineering from Virginia Tech. He is a Sr. Principal Engineer at the Fab/Sort Manufacturing network of Intel Corp., responsible for the application of operational optimization in high volume manufacturing across Intel's factories, driving improvements in WIP management, production capacity and cycle time, equipment and capital productivity. He is also an Adjunct Associate Professor at Ben-Gurion University, Israel and serves as a co-chair of the IEEE Technical Committee on Semiconductor Manufacturing Automation (TC-SMA). His email is kalira@post.bgu.ac.il.

LARS MÖNCH is Professor in the Department of Mathematics and Computer Science at the University of Hagen, Germany. He received a master's degree in applied mathematics and a Ph.D. in the same subject from the University of Göttingen, Germany. His current research interests are in simulation-based production control of semiconductor wafer fabrication facilities, applied optimization and artificial intelligence applications in manufacturing, logistics, and service operations. He is a member of GI (German Chapter of the ACM), GOR (German Operations Research Society), and INFORMS. He is an Associate Editor for the European Journal of Industrial Engineering, Journal of Simulation, Business & Information Systems Engineering, IEEE Robotics and Automation Letters, IEEE Transactions on Semiconductor Manufacturing, and IEEE Transactions on Automation Science and Engineering. His email address is lars.moench@fernuni-hagen.de.