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.

USING DELAYS FOR PROCESS FLOW SIMPLIFICATION

Igor Stogniy

Technische Universität Dresden
Institute of Applied Computer Science
Dresden, 01062, GERMANY

Wolfgang Scholl

Infineon Technologies Dresden
Königsbrücker Straße 180
Dresden, 01099, GERMANY

ABSTRACT

Infineon Technologies Dresden uses discrete event simulation to forecast key performance indicators. The simulation has also been used to perform experiments to improve production planning. It is important to reduce the efforts required for the creation and maintenance of the simulation models. Especially for the simulation studies, less detailed models can be utilized where components could be omitted. We considered a simplification of the process flows through operation substitution for constant delays. Different levels of model complexity were investigated. For each level, different tool sets were determined which were substituted for delays. First In First Out and Critical Ratio dispatching rules were used. Lot cycle time distributions were utilized in order to compare simplified models with a detailed model.

1 INTRODUCTION

Fab simulation is already used as a standard practice in the production planning of the semiconductor industry. Today's simulations require considerable effort for maintenance. Data inconsistency and lack of data availability (heterogeneous databases and files), multiplied by the quantity of data, lead to significant simulation efforts. To carry out simulation experiments, there is a need to construct simplified models to reduce the run time of the model and the efforts for manual model calibration. By default, it is assumed that a simplified model should be accurate enough for the simulation purposes.

In the semiconductor industry, simplified models are often considered in the framework of supply chain simulation (Jain et al. 2000). In this case the main criterion is the run time of the model, because it is necessary to make numerous simulation runs to achieve an optimized decision. In fab simulation, the maintainability of the model becomes more important, because a common use of the simulation model is to prove different production scenarios, which is implemented with the help of manual changes to the model. A simplified model could have a relatively long run time (up to 1 hour), but it should be transparent enough to support manual calibration efforts. Rank et al. (2016) reflected on the importance of having the correct level of model complexity in semiconductor fab simulation and came to the conclusion that the different tasks need different levels of simplification. The authors of this article highlighted automated material handling systems especially. Our focus is on a fab simulation as a whole. The main goal of the research is to investigate how many tool sets could be substituted by constant delays. Therefore, we consider a problem of gradual simplification. As a basis for the research, the well-known MIMAC Datasets (1997) were used. MIMAC dataset 5 was chosen, because it has the biggest quantity of process flows in comparison to other sets. For the fab simulation model, AutoSched AP from Applied Material, version 10.1 is used.

This paper is organized as follows. Several ideas from related works are presented in Section 2. The concept of the research, including artificial process flow building and tool sets substitution, are described in Section 3. The logic of the experiments and lot distribution diagrams are shown in Section 4. Section 5 includes summarized experimental results. An issue in delay calculation is discussed in Section 6.

978-1-7281-3283-9/19/$31.00 ©2019 IEEE

2 RELATED WORK

An overview of the simulation model simplification was already described in (Van der Zee 2019). As pointed out by the author of this article, simplification "is still very much green field".

Rose (1999) presented results for FIFO (First In First Out) and CR (Critical Ratio) dispatching rules. He noticed that an overtaking effect takes place for the CR rule. In his paper, CR is equal to (due date – current time)/remaining processing time. We used the same formula because it is used in the AutoSched AP. Rose (2000) later described an overtaking behavior for random delays. To avoid this, we decided to use the constant delays. An interarrival approach (Sprenger and Rose 2010) could help to overcome the overtaking behavior for the random delay case, but it does not work for non-FIFO dispatching rules. Furthermore, Hung and Leachman (1999) noticed that, because of negative correlation, simplified models with constant delays are more accurate than models with random delays. Additionally, model maintainability needs fewer efforts for the constant delays.

In our paper we consider two types of simplification: process flow aggregation and tool sets substitution. Piplani and Puah (2004) presented an aggregation of routings by using representative flows. We developed another approach: an artificial process flow which comprises all unique operations from aggregated process flows. A typical criterion for tool sets substitution is workstation utilization (e.g. Rose (1999), Jain et al. (2000), Johnson et al. (2005)). However, Hung and Leachman (1999) used the standard deviation of lot waiting time. Piplani and Puah (2004) used the workstation throughput rate. The existence of different approaches suggested a need for deeper analysis of the substituted tool sets.

Johnson et al. (2005) demonstrated a gradual simplification of tandem-10 M/M/1 and tandem-10 M/M10 models. Another approach is presented by Völker and Gmilkowsky (2003). They focused on the operations themselves and tried to compensate the decreasing utilization through capacity reservations. They found "pathological errors" when using capacity reservations. Only the elimination of complete capacity was found to avoid these errors. In our paper, we used a substitution which includes the elimination of complete capacity.

To analyze lot cycle distributions, we introduced summarized absolute divergence (SAD), which is similar to the "similarity between two histograms" $\overline{HD}_{\text{abs}}$ (Ewen et al. 2017), but SAD = $2 * \overline{HD}_{\text{abs}}$, and the cumulative SAD suits better to comparison with another criterion, the Kolmogorov-Smirnov test.

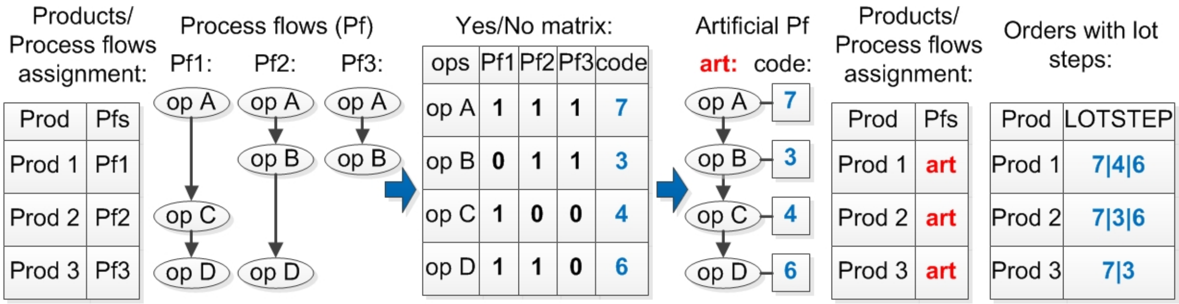
Despite the fact that Hood (1990) pointed out the importance of operator modeling, we decided to ignore the influence of operators. Our aim was to investigate the influence of tool set substitution and it would have been difficult to distinguish effects from tool set substitution and operator reallocation. Moreover, since the 1990s, semiconductor fabs have become much more automated and, thus, the operator factor no longer plays an important role.

3 DESIGN OF EXPERIMENTS

3.1 Artificial Process Flow Concept

The main idea of our simplification approach is a substitution of the different operations within the process flows with delays. We also tried to create an artificial process flow to reduce the quantity of the process flows. To achieve the artificial process flow, identical operations were determined. Two operations are considered as the same if they have identical tools set, load/unload time, processing time, batch ID, product setup, and setup group. An operation could be identical with one or zero operations in the other process flow. The operation order in each process flow was taken into account. Using this method, 590 unique operations were identified among 4176 operations in 24 process flows.

Each operation in artificial process flow was codified. Based on the process flow operations (see Figure 1), a "yes/no" matrix was determined. In the matrix the values are set to the following rules: 1, if we have the operation in the particular process flow, and 0, if it is not present. In this way we generated a binary code for each row in the matrix, which was converted to decimal code for length considerations (e.g. for operation B we have a triplet "011" and code 3). In the end we had created an artificial process flow with all operations and codes, and we used the same codes in the order table for each product. The codes are represented by the standard entity LOTSTEP in the simulation tool.



Prod	Pfs
Prod 1	Pf1
Prod 2	Pf2
Prod 3	Pf3

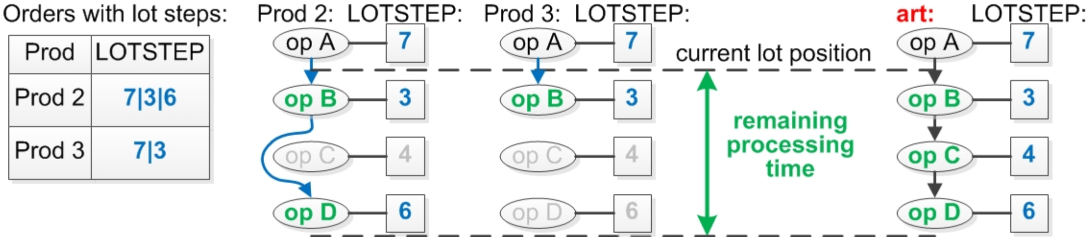
ops	Pf1	Pf2	Pf3	code
op A	1	1	1	7
op B	0	1	1	3
op C	1	0	0	4
op D	1	1	0	6

Prod	Pfs
Prod 1	art
Prod 2	art
Prod 3	art

Prod	LOTSTEP
Prod 1	7|4|6
Prod 2	7|3|6
Prod 3	7|3

Figure 1: Making and using an artificial process flow.

Figure 2 illustrates how a particular product paths through the artificial process flow. We used this artificial process flow in all of the experiments with the FIFO dispatching rule. Unfortunately, the usage of the artificial process flow for the built-in CR dispatching rule could not be used because the simulation tool (AutoSched AP 10.1) does not consider the codes (LOTSTEP) in the calculation of the remaining processing time. The following illustrates this point. Let us consider two lots after operation A: prod 2 and prod 3 (see Figure 2). The correct calculation of the remaining processing time for prod 2 is a sum of processing time of operations B and D, while for prod 3 it is the processing time of operation B only. But as long as we have only one list of operations for all process flows ("art" in Figure 2), a simulation engine considers the whole list without taking into account codes (LOTSTEP) and calculates the remaining processing time for both lots as a sum of the processing time of operations B, C, and D. The simulation tool allows custom-written dispatching rules, but this requires additional programming effort. Our aim in this research was to use standard simulation mechanisms. Therefore, for the experiments with the CR dispatching rule, we used all process flows and not artificial ones.



Prod	LOTSTEP
Prod 2	7|3|6
Prod 3	7|3

Figure 2: Product paths in the artificial process flow and remaining processing time.

3.2 Tool Sets Substitution and System Configurations for the Experiments

The main idea of the experiments presented here is to investigate how many tool sets could be substituted by delays. At first the delays were summarized for each process flow and placed at the end of the flow (see Figure 3). The delay for each of the tool sets is based on the sum of load time, unload time, processing time, and average waiting time. The average waiting time was extracted from the standard model reports for tool sets. In reality, these data would be historical production data, but we used a detailed model as a source. Thus, on the one side we have process flow specific values for each delay: all times are calculated according to the substituted operations in the process flow. On the other side, the value of the average waiting time is an average for all process flows for the particular tool set. For example, to calculate delay 1 and delay 2 in Figure 3, we consider the same average waiting time as for tool set 1. It is not possible to determine the average value for the tool set with the process specific granularity based on the standard model reports. This leads to certain issues which will be discussed in Section 6.

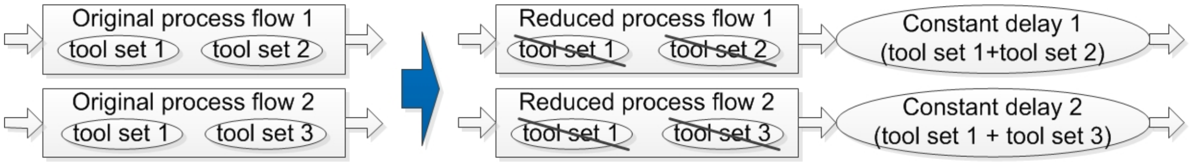


Figure 3: Tool sets substitution.

Two criteria were used to distinguish all tool sets: idle time of a tool set (IDLE) and a "queue ratio" (average queue length before a tool set / max batch size) (Figure 4). The first criterion is common for simulation experiments and means utilization of the tool set. We used the idle time and not utilization time because, in the simulation tool reports, utilization time does not include tool set downs. As for the second criterion, we used it mostly for batch tool sets with an idle time of less than 10% to distinguish between batch tool sets whose batches are not usually full ("half full"), and those which have almost full batches ("true bottlenecks"). In this case, non-batch tool sets with idle time less than 10% will be considered as "true bottlenecks", because they have a long queue in front of them. We also used this criterion for tool sets with idle time bigger than 10%, which have relatively big queues in front of them ("big queue"), to distinguish them from tool sets with an idle time bigger than 10%, which have relatively small queues ("normal"). The value of 0.9 for the queue ratio was chosen by intuition. This way we have three substituted tool set groups: (1) "normal", (2) "half full", (3) "big queue". Additionally, in the group "big queue" among all tool sets, two sub groups were determined: without setup and with setup (FIFO: 9 and 3; CR: 7 and 3 tool sets, respectively).

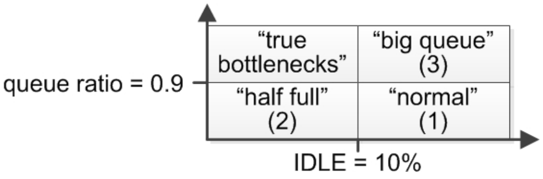


Figure 4: Substituted tool set groups.

The following system configurations were determined for the experiments:

- all – all process flows with all operations (base case, i.e. an original detailed model)
- art – artificial process flow (only for FIFO)
- 50, 40, 30, 20, 10 – "normal" tool sets were substituted (group 1), which have an idle time bigger than 50%, 40%, 30%, 20%, 10%, respectively
- A – additionally to "10", "half full" tool sets were substituted (group 2)
- B – additionally to "10", all "big queue" tool sets were substituted (group 3)
- C – additionally to "A", tool sets without setup from the "big queue" group were substituted (3)
- D – additionally to "A", all "big queue" tool sets were substituted (group 3)
- E – additionally to "C", three "true bottlenecks" without setup were substituted

4 EXPERIMENTS

We consider a steady state single load scenario. For each of the experiments, a simulation run is 11 years. We needed such a large horizon to gather enough statistics for rare events: combinations of monthly and quarterly down events. The first year results were not considered to avoid initialization bias effects. FIFO and CR dispatching rules were used. The down event streams were fixed for all experiments. Thus, we could implement only one simulation run for each of the experiments, but were able to extract comparable data despite this fact. The original MIMAC dataset 5 was changed to ignore the influence of operators: operators quantity was fixed higher than in the original data. Table 1 summarizes some information about the system configurations and experiments.

Table 1: Configurations and experiments.

Exp	FIFO (First In First Out)							CR (Critical Ratio)					
	Substituted tool set group			Run time	Number of operations			Substituted tool set group			Run time	Number of operations	
	1	2	3	(mm:ss)	#	% (all)	% (art)	1	2	3	(mm:ss)	#	% (all)
all				25:04	4176	100.00					25:45	4176	100,00
art				25:33	524	12.55	100.00	n/a					
50	27			20:33	460	11.02	87.79	28			20:38	3278	78.50
40	35			18:47	410	9.82	78.24	34			18:44	3111	74.50
30	43			15:51	346	8.29	66.03	45			15:55	2534	60.68
20	47			15:53	319	7.64	60.88	49			14:41	2441	58.45
10	54			11:58	256	6.13	48.85	58			11:11	1591	38.10
A	54	13		05:46	150	3.59	28.63	58	11		06:08	607	14.54
B	54		9+3	06:47	174	4.17	33.21	58		7+3	05:40	1196	28.64
C	54	13	9	02:31	105	2.51	20.04	58	11	7	02:38	288	6.90
D	54	13	9+3	01:36	68	1.63	12.98	58	11	7+3	01:42	212	5.08
E	54+3	13	9	01:41	96	2.30	18.32	58+3	11	7	01:49	203	4.86

4.1 Experiments with FIFO Dispatching Rule

In Figure 5 (left diagram) the lot cycle time distribution diagrams for “all” and “art” experiments are represented. The model results with the artificial process flow (“art”) differentiate slightly from the detailed model with all process flows (“all” – base case). The “art” curve is slightly shifted to the right relative to the “all”. We suppose that the difference is caused by internal lot ordering algorithms of the simulation tool.

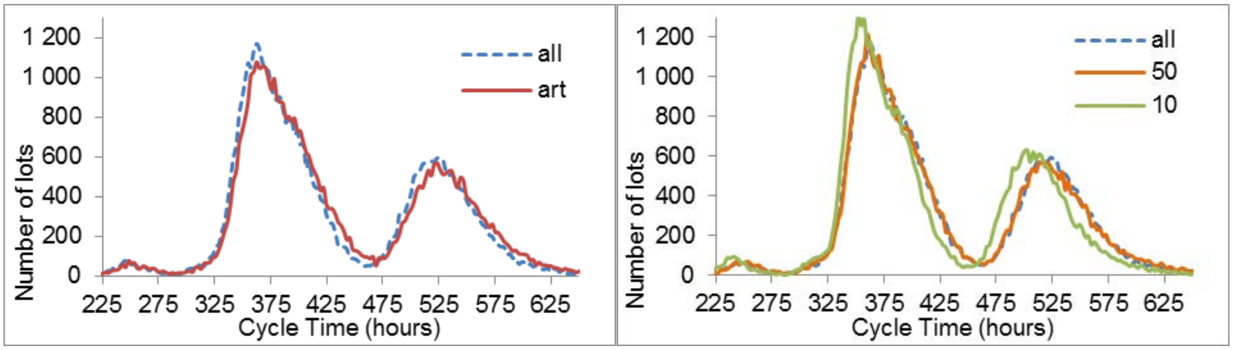


Figure 5: Lot cycle time distributions (FIFO): all – all process flows; art – artificial process flow; 50 – “art” with substituted IDLE>50% tool sets; 10 – “art” with substituted IDLE>10% tool sets.

On the right side of Figure 5 lot cycle time distribution diagrams for “all”, “50”, and “10” experiments are represented. The results for “40”, “30”, and “20” experiments are not shown in the Figure for the purpose of simplification. Some numerical values of these experiments can be found in Table 2 (see Section 5). We can see that “10” is slightly shifted to the left relative to “50”. There is the same directional shift between “50” and “art”, which is why “50” is closer to “all” and more accurate than “art”. We suppose that this shift to the left happened during the simplification because the interdependencies between operations were ignored when we substituted tool sets with delays. The investigation of the interdependencies is not considered in this paper. Here we present the results of different system configuration investigations and attempt to determine directions for future research.

Based on our practical experience, it was clear that the substitution of the most idle tool sets from group 1 (“normal”) do not have a big impact on the lot cycle time distribution. It is much more interesting to see what will happen with the other two groups, namely “half full” and “big queue” (see Figure 6).

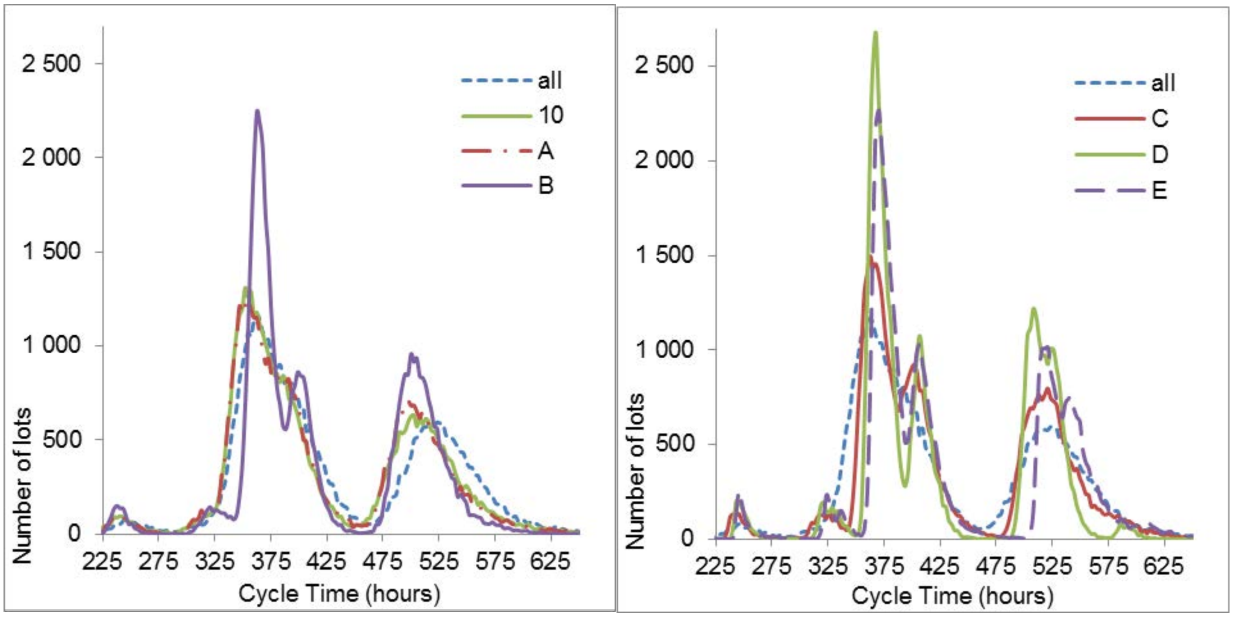


Figure 6: Lot cycle time distributions (FIFO): all – all process flows; 10 – “art” with substituted IDLE>10% tool sets; A – “10” with substituted “half full” tool sets; B – “10” with substituted “big queue” tool sets; C – “A” with substituted tool sets without setup from the group “big queue”; D – “A” with substituted “big queue” tool sets; E – “C” with substituted three “true bottlenecks” without setup.

On the left side of Figure 6 the comparison of the “A” (substituted “half full”) and “B” (substituted “big queue”) configurations is drawn (with “all” and “10” as references). The divergence between “10” and “A” is very small. Numbers are provided in the analytics in Table 2 (see Section 5). Thus, the contribution of the “half full” tool sets to the resulting lot cycle time diagram is small in our case. We want to remind the reader that “half full” tool sets are those batch tool sets with an idle time less than 10% but are not fully loaded. In other words, the “half full” are not bottlenecks despite the relatively high utilization. Another interesting observation is the difference between “10” and “B”. It is shown that “big queue” tool sets have a large impact on the lot cycle time distribution, and without them (“B”) the lot cycle distribution becomes narrower. This means that those tool sets contribute to the variability of the system significantly, in spite of the fact that these tool sets are not highly utilized (idle time is more than 10%).

Afterwards, we looked closer at the “big queue” tool set group and found that there are three tool sets with setup. It was assumed that those three tool sets have the biggest contribution in the whole group. A confirmation of this can be found on the right side in Figure 6, when comparing experiments “C” and “D” (“C” consists of four “true bottlenecks”, three tool sets with setup from “big queue”, and delays; “D” consists of only four “true bottlenecks” and delays). It is also interesting to see that the remaining “big queue” tool sets have a valuable impact on variability (compare “C” and “all”).

We wanted to investigate whether the impact of tool sets with setup is as big (or even bigger) as the impact of “true bottlenecks”. For this reason, we performed experiment “E”, where three of the “true bottlenecks” without setup are substituted by delays. In the end, experiment “E” has a configuration with only four tool sets with setup and delays. It is interesting to see that the curve of “E” is closer to “all” than “D”. This means that “big queue” tool sets with setup could have an even bigger impact than “true bottlenecks” without setup.

4.2 Experiments with CR Dispatching Rule

At first we conducted all experiments with the CR dispatching rule according to the tool set substitution concept, which is shown in Figure 3, and we received unexpected results (see left side in Figure 7). The left peak of the “all” curve was relatively well covered by curves “50” and “10”, but the right one was not well matched. Similar statements for the right peak effect were described in (Rose 1999). Rose pointed out an overtaking effect as “lots being processed earlier at the bottleneck than lots that are closer to their due date”. This occurs because of the reordering of lots at a smaller quantity of work centers in the simplified model (in Rose’s case at the bottleneck only). However the “overtaking behavior” described by (Rose 2000) for a random delay does not happen in our case because we use constant delays.

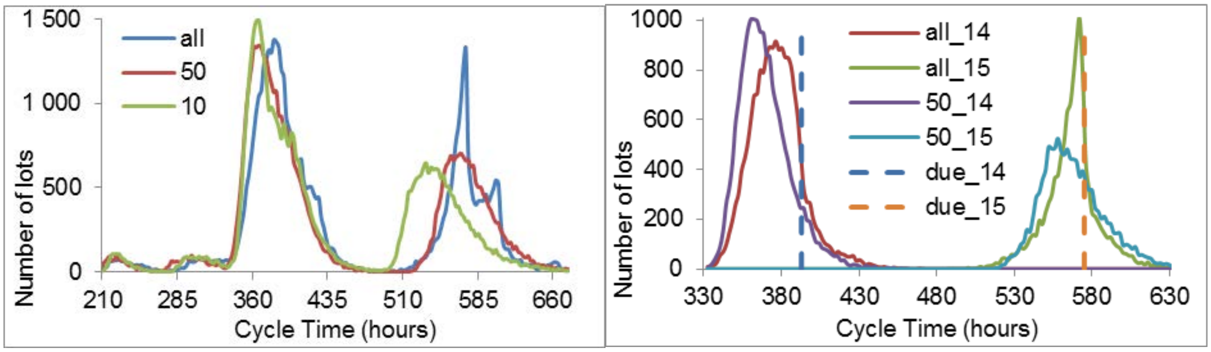


Figure 7: Lot cycle time distribution (CR). Left: Substitution with delays at the end of the process flows. Right: Inaccuracy illustration for process flows 14 and 15 (due – due date).

To understand the differences in the peak covering, the lot cycle distributions for each process flow were analyzed. In the right diagram of Figure 7 we see the lot cycle time distributions for process flows 14 and 15 for “all” and “50” experiments with their due dates (due_14 and due_15). The distributions all_14 and all_15 are asymmetric and, moreover, as long as distribution 15 is closer to its due date, the asymmetry is bigger for 15 than for 14. This shows the consequences of the CR rule influence: an “acceleration/deceleration” effect (it accelerates a lot when it is close to its due date at the expense of a lot which is not so close to its due date). We do not see this effect for “50”. The reason for this is that at the end of each process flow we put delays which are big enough and that do not contain the “acceleration/deceleration” property, because they are only large constant delays. The rest of the process flow before the delays does have it, but it does not perform optimally, because the “acceleration/deceleration” effect is higher when a lot comes closer to the end of the process flow and its due date. The solution to this problem is to put delays not at the end (as it is shown in Figure 3) but at the beginning of the each process flow. The results are shown in Figure 8: both peaks of “all” are covered better by “50” and “10”. Thus we could reduce the overtaking effect in our simplified models.

The results for configurations “C”, “D”, and “E” look similar to Figure 6. However, there is a difference: the left peak is much higher than in Figure 6 due to the “acceleration/deceleration” property of the CR rule and the reduced variability in the more simplified models “D” and “E” (more narrow distributions).

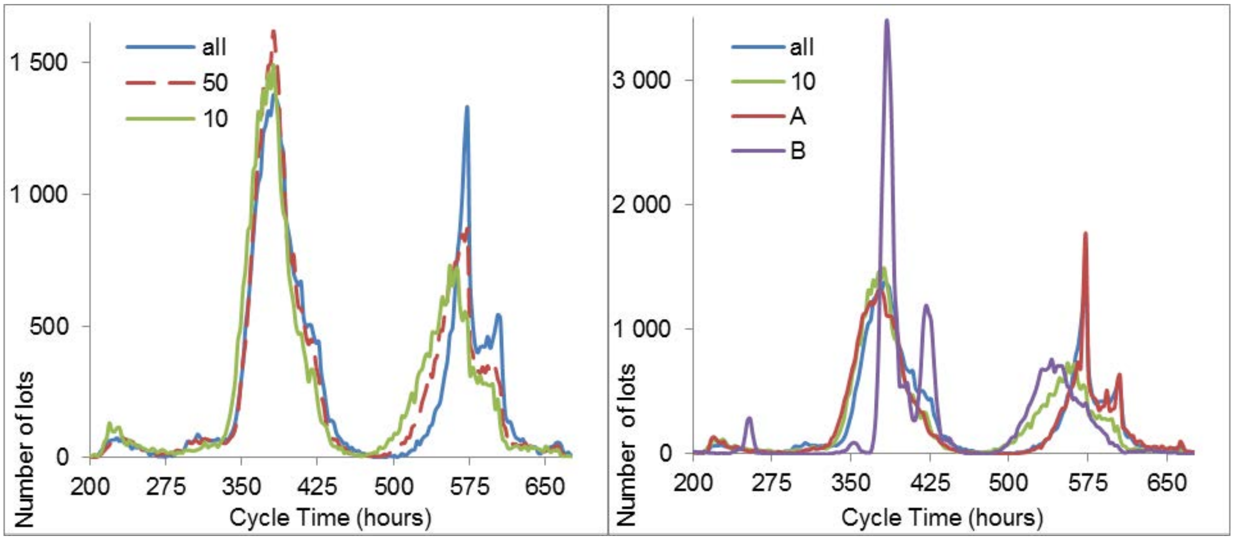


Figure 8: Lot cycle times distributions (CR). Left and right: Substitution with delays at the beginning of the process flows.

5 EXPERIMENTAL RESULTS

To analyze the experimental results we could not use the "classical" statistical tests t-test, z-test, F-test, or χ^2 test because the distributions in our case are far from normal, F-distribution, or χ^2-distribution. Therefore we used the Kolmogorov-Smirnov test (K-S) as a nonparametric test. An advantage of this test is that it can be used for any distribution, while a disadvantage is that it only measures the largest difference at one "point". Additionally we wanted to use an aggregated criterion in order to compare all distributions with the base case. Unfortunately, it is not possible to use a "classical" approach for machine learning the Kullback-Leibler divergence in our case, because some distributions have zero values in the areas where the base case distribution has non-zero values (The Kullback–Leibler divergence is defined only if a zero value of a distribution implies a zero value of another distribution). For this reason, we used summarized absolute divergence (SAD). For a short discussion about the criteria, see Appendix A. In Table 2 and Figure 9 the experimental results are represented.

Table 2: Experimental results summary.

	FIFO (First In First Out)					CR (Critical Ratio)						
Exp	Substituted tool sets group			FIFO (artificial process flow)		Substituted tool sets group			CR_e (delays at the end)		CR_b (delays at the beginning)	
	1	2	3	K-S	SAD	1	2	3	K-S	SAD	K-S	SAD
art				0.0383	0.1385				-	-	-	-
50	27			0.0115	0.0688	28			0.0898	0.3216	0.0724	0.2220
40	35			0.0106	0.0798	34			0.0992	0.3630	0.0757	0.2501
30	43			0.0378	0.1533	45			0.0914	0.3741	0.1034	0.3249
20	47			0.0306	0.1411	49			0.1846	0.4989	0.0994	0.3373
10	54			0.0687	0.2689	58			0.1686	0.5453	0.1228	0.4139
A	54	13		0.0850	0.3054	58	11		0.2257	0.7170	0.0806	0.2517
B	54		9+3	0.1108	0.5219	58		7+3	0.1901	0.7900	0.1867	0.8912
C	54	13	9	0.0732	0.2965	58	11	7	0.1103	0.4223	0.1031	0.4154
D	54	13	9+3	0.1237	0.6760	58	11	7+3	0.2108	0.8033	0.2058	0.9050
E	54+3	13	9	0.1666	0.5886	58+3	11	7	0.2252	0.6812	0.2507	0.8010

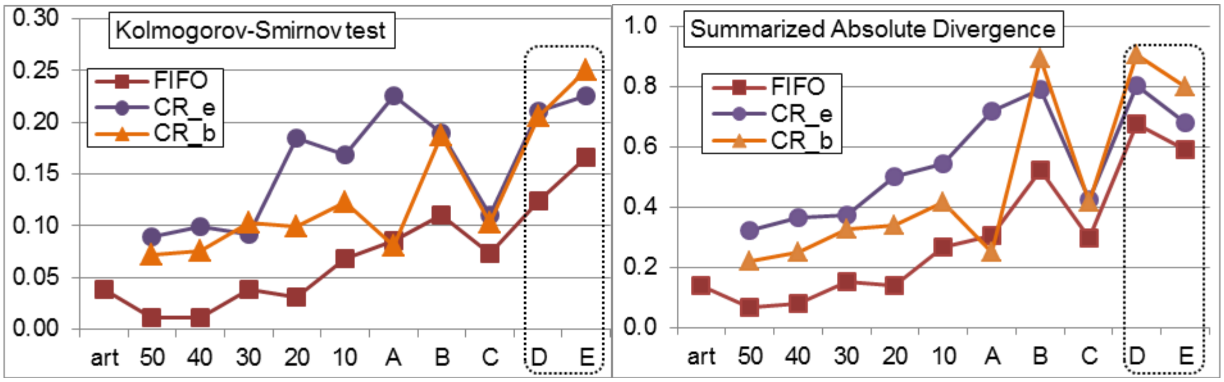


Figure 9: Experimental results: Kolmogorov-Smirnov test and summarized absolute divergence.

Most of the experimental observations under the FIFO rule were described above to provide a logic of experiments (see Section 4.1). The main finding here is the small influence of the "half full" tool sets (A) and valuable impact of the "big queue" tool sets with setup (compare B and C, D and E). It means that the work center utilization could not be a single criterion for the substitution (as it is described in most of the above mentioned papers). The waiting time distribution should also be considered and, despite the correlation with utilization (Hung and Leachman 1999), it should be considered separately. It is especially essential for tool sets with setup (compare B and C). Moreover, it is not possible to consider only the number of operations as a reduction criterion (Völker and Gmilkowsky 2003). It is important to take into account which kind of operations were reduced. A more detailed look at the form of distribution should be considered, which our future research will focus on.

The accuracy of the models for the FIFO rule is much better than for the CR rule. This is probably a consequence of an imprecise delay calculation (see Section 6). Nevertheless, we had a good approximation in the case of the CR rule. Based on diagrams from Sprenger and Rose (2010), we could conclude that our approach overcomes their delay approach for the CR rule because the overtaking effect was reduced. It was achieved thanks to a more detailed model and constant delays. For the most configurations the accuracy for the CR variant with delays at the beginning (CR_b) is better than the CR variant with delays at the end (CR_e). The reason for this was illustrated in Figures 7 and 8: the "acceleration/deceleration" effect of the CR rule.

Generally, Table 2 and Figure 9 show a relatively small degradation from "50" to "A" configurations. We concluded from this that the "normal" and "half full" substitution tool set groups are not as important for model accuracy. The experiments "B", "C", "D", and "E" with group "big queue" show that within the "big queue" group the most influential tool sets are those which have setup.

6 DISCUSSION

In Figure 8 we could see quite obvious differences for the right peak. It is also very interesting to see that "A" is closer to "all" than "10" is, especially at the right peak. Both "A" and "B" exhibit the same behavior at the right peak. "A" is closer to "all" than "B". From this we can conclude that different simplified models have different accuracies, which is not strictly correlated with model complexity. We did not find such an effect in most of above mentioned literature. Only Völker and Gmilkowsky (2003) noted it, calling the behavior "pathological". But they observed it for configurations with capacity reservation, and we did not use any reservation. We assumed that this effect is a consequence of calculations not being precise enough for the delays as averages for the tool sets and the lack of product-specific averages for the tool sets (see Section 3.2). To prove this idea, we performed several additional experiments with manual calibration of the delays, but only for one process flow – 15.

In Figure 10 the results of manual adjustments for process flow 15 are represented. On the left side, curve 10_15 represents the lot cycle time distribution only for process flow 15 for the configuration "10"

(CR_b case). The delay length for process flow 15 was alternated. After several calibration experiments it was found that if we added 60 more hours to the delay, which was calculated for configuration “10” (219.3 hours), we would get curve 10_15_adj, which is much closer to curve all_15 (base case). The same type of adjustment was done for configuration “B”. In this case, 50 hours (B_14_adj) were added to 291.8 hours (B_15). Additionally, several calibration experiments for “A” were carried out and it was found that there was no need to add or subtract anything. For this reason, it is not shown in Figure 10. Thus we could see that the delay calculation based on standard tool set reports was not precise enough. Making it more precise would be a question for future research activities.

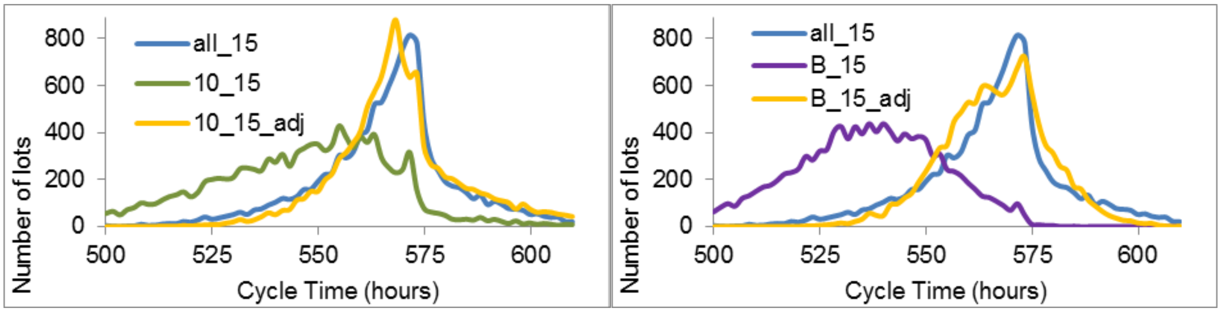


Figure 10: Lot cycle times distributions (CR_b case). Manual adjustments for the process flow 15.

7 CONCLUSIONS

In this paper we considered a simplification of process flow operations modeling through substitution for constant delays. For the substitution needs, we carried out a tool sets analysis based on the data from a detailed model. All tool sets were divided into four groups: “normal”, “half full”, “big queue”, and “true bottlenecks”. The tool sets from the groups “normal” and “half full” were not found to have a significant impact on model accuracy. Among tool sets from the group “big queue”, tool sets with setup contribute more to the variability (and as a consequence, to model accuracy) than tool sets without setup from this group. Based on experimental results, we assumed that work center utilization could not be the only criterion for the substitution. The waiting time distribution should also be considered. Experiments with the Critical Ratio dispatching rule revealed a weakness in our initial concept. They showed the importance of the delay position in each process flow (at the beginning, and not at the end). They also demonstrated possible inaccuracies when using waiting time calculations based on standard simulation reports as averages for the tool sets and not product-specific averages for the tool sets.

In further research we plan to enhance our approach for variable workload scenarios. We will continue the development of the artificial process flow concept and plan to implement some aggregated operations in the artificial process flow. Another area of interest includes a deeper analysis of the substituted tool set groups. We will continue the experiments with different values of the “queue ratio”. Additionally, we will disclose the more precise delay calculation. Eventually, we plan to implement these ideas in a real Infineon simulation model.

ACKNOWLEDGMENTS

A part of the work has been performed in the project iDev40. The iDev40 project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 783163. The JU receives support from the European Union’s Horizon 2020 research and innovation programme. It is co-funded by the consortium members, grants from Austria, Germany, Belgium, Italy, Spain and Romania.

APPENDIX A

Additionally, we would like to discuss the criteria which we used for model evaluation. Kolmogorov-Smirnov test: $KS = \sup|CF_n(x) - CF(x)|$, where $CF_n(x)$ cumulative distribution functions "50", "40", … "E", and $CF(x)$ – cumulative distribution function for the base case ("all"). The value of KS is changed from 0 when two distributions are identical, to 1, when two distributions have nothing in common. Summarized absolute divergence: $S_{ad} = \sum_x |F_n(x) - F(x)|$. The value of S_{ad} is the area between two probability density functions $F_n(x)$ and $F(x)$. It is changed from 0, when two probability density functions are identical, to 2, when the functions have nothing in common (compare with a "similarity between two histograms" – $\overline{HD}_{\text{abs}}$ from Ewen et al. (2017), SAD = $2 * \overline{HD}_{\text{abs}}$).

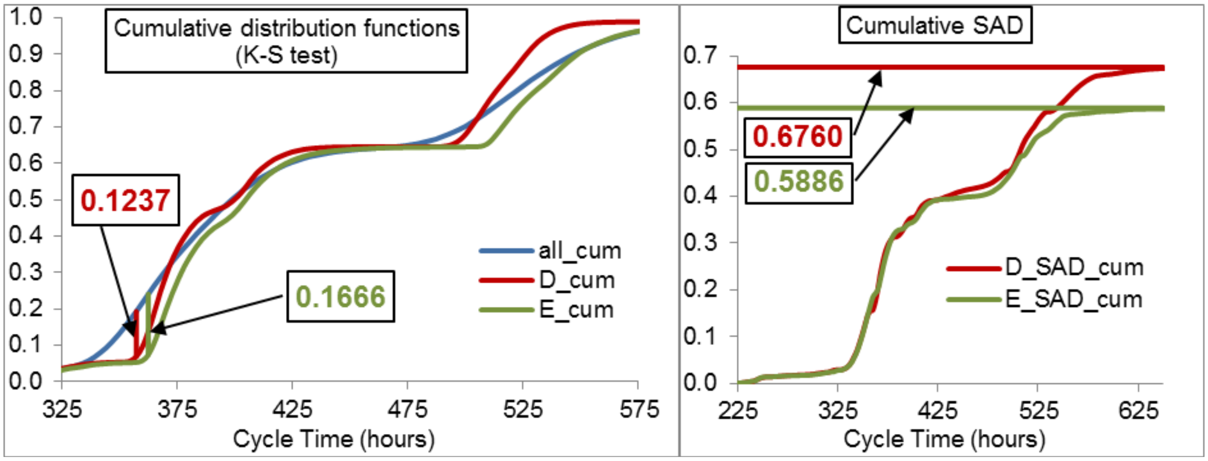


Figure 11: Comparison of two evaluation criteria for the experimental results (FIFO). Left: Kolmogorov-Smirnov test. Right: cumulative summarized absolute divergence.

KS and SAD usually correlate with each other, but sometimes there is disagreement, for example experiments "D" and "E" (see dotted area in Figure 9, compare with Figure 6). Figure 11 illustrates how the criteria are calculated. The Kolmogorov-Smirnov test evaluates the divergence in a limited fashion – it only considers the maximum divergence between *cumulative* distribution functions. Summarized absolute divergence generalizes the divergences for the probability density functions *among the whole range of values*. We assume that the summarized absolute divergence represents the difference between modeling results better than the Kolmogorov-Smirnov test.

REFERENCES

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.

Hood, S. J. 1990. "Detail vs. Simplifying Assumptions for Simulating Semiconductor Manufacturing Lines". In *Proceedings of the Ninth IEEE/CHMT International Symposium on Electronic Manufacturing Technology, Competitive Manufacturing for the Next Decade*, October 1st – 3d, Washington, DC, USA, 103-108.

Hung, Y. F. and R. C. Leachman. 1999. "Reduced Simulation Models of Wafer Fabrication Facilities". *International Journal of Production Research* 37(12): 2685-2701.

Jain, S., B. P. Gan, C. C. Lim, and Y. H. Low. 2000. "Bottleneck Based Modelling of Semiconductor Supply Chains". In *Proceedings of the International Conference on Modeling and Analysis of Semiconductor Manufacturing*.

Johnson, R. T., J. W. Fowler and G. T. Mackulak. 2005. "A Discrete Event Simulation Model Simplification Technique". In *Proceedings of the 2005 Winter Simulation Conference*, edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 2172–2176. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

MIMAC Datasets. 1997. http://p2schedgen.fernuni-hagen.de/index.php?id=296, accessed 14th June 2019.

Piplani, R. and S. A. Puah. 2004. "Simplification Strategies for Simulation Models of Semiconductor Facilities". *Journal of Manufacturing Technology Management* 15(7): 618-625.

Rank, S., C. Hammel, T. Schmidt, J. Müller, A. Wenzel, R. Lasch, and G. Schneider. 2016. "The Correct Level of Model Complexity in Semiconductor Fab Simulation – Lessons Learned from Practice". In *Proceedings of the 27th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC) Saratoga Springs 16-19 May 2016*, 133-139. Saratoga Springs, New York: Institute of Electrical and Electronics Engineers, Inc.

Rose, O. 1999. "Estimation of the Cycle Time Distribution of a Wafer Fab by a Simple Simulation Model". In *Proceedings of the Semiconductor Manufacturing Operational Modeling and Simulation conference*, San Francisco, CA, USA,133-138.

Rose, O. 2000. "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.

Sprenger, R., and O. Rose. 2010. "On the Simplification of Semiconductor Wafer Factory Simulation Models". In *Conceptual Modeling for Discrete-Event Simulation*, edited by S. Robinson, R. Brooks, K. Kotiadis, and D.J. van der Zee, 451-470. Boca Raton: CRC Press.

Van der Zee, D. J. 2019. "Model Simplification in Manufacturing Simulation – Review and Framework". *Computers & Industrial Engineering*, 127: 1056-1067.

Völker S. and P. Gmilkowsky. 2003. "Reduced Discrete-Event Simulation Models for Medium-Term Production Scheduling". *Systems Analysis Modeling Simulation* 43(7): 867-883.

AUTHOR BIOGRAPHIES

IGOR STOGNIY is a PhD-student at the Technische Universität Dresden (Germany). He received his M.Sc in Automation and Control in 2007 from Bauman Moscow State Technical University (Russia). Since 2012 he has been working at the TU Dresden as a scientific researcher in the area of simulation. His email address is igor.stogniy@tu-dresden.de

WOLFGANG SCHOLL works as a Senior Staff Expert for modeling and simulation for Infineon Technologies in Dresden (Germany). He studied physics at the Technical University of Chemnitz (Germany) and graduated in solid-state physics in 1984. From 1984 to 1995 he worked as a process engineer for ZMD in Dresden. In 1996 he joined Infineon Technologies (former SIMEC) and worked in the field of capacity planning. Since 2003 he has been responsible for fab simulation. He supervises development and roll-out projects and is also a member of the Supply Chain Simulation community. His email address is wolfgang.scholl@infineon.com.