Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

SENSITIVITY ANALYSIS OF THE SMT2020 TESTBED FOR RISK-BASED MAINTENANCE IN SEMICONDUCTOR MANUFACTURING

Philipp Andelfinger[1], Shuaibing Lu[2], Chew Wye Chan[3], Fei Fei Zhang[3], Boon Ping Gan[3],
Jie Zhang[1], and Wentong Cai[1]

[1]College of Computing and Data Science, Nanyang Technological University, SINGAPORE
[2]College of Computer Science, Beijing University of Technology, Beijing, CHINA
[3]D-SIMLAB Technologies Pte Ltd, SINGAPORE

ABSTRACT

A key concern for factory operators is the prioritization of machine groups for preventive maintenance to reduce downtimes of the most critical machine groups. Typically, such maintenance decisions are made based on heuristic rules supported by discrete-event simulations. Here, we present a large-scale sensitivity analysis that elucidates the relationship between machine group features and factory-level performance indicators in the presence of unplanned tool downtimes in SMT2020. From extensive simulation data, we extract predictors that permit a machine group ranking according to their criticality. An exhaustive combinatorial evaluation of 11 machine group features allows us to identify the most predictive features at different subset sizes. By evaluating and visualizing predictions made using linear regression models and neural networks against the ground truth data, we observe that to accurately rank machine groups, capturing nonlinearities is vastly more important than the size of the feature set.

1 INTRODUCTION

The prioritization of maintenance efforts is a key consideration for operators of manufacturing plants (Gopalakrishnan and Skoogh 2018). The notion of risk-based or criticality-based maintenance comprises systematic approaches for targeted maintenance in order to minimize risk in the form of the probability and impact of equipment failures. Tool criticality measures that combine static or dynamic machine attributes (Mönch and Zimmermann 2007) can drive heuristics based on which maintenance decisions can be made.

The SMT2020 testbed (Kopp, Hassoun, Kalir, and Mönch 2020b) is a commonly used simulation model representing typical characteristics of semiconductor fabrication plants. It is intended as a sufficiently realistic representation to close the gap between scientific research and real-world manufacturing practice, covering various details such as batch processing, cascading, lot-to-lens dedications, and reentrant process flows. The model is widely used for evaluating optimization methods for key aspects of semiconductor manufacturing such as scheduling and dispatching using heuristics (Mönch and Zimmermann 2007), meta-heuristics (Kovács, Tassel, Ali, El-Kholany, Gebser, and Seidel 2022), and learning-based methods (Tassel, Kovács, Gebser, Schekotihin, Stöckermann, and Seidel 2023). The dataset studied in the present paper represents a low-volume/high-mix wafer fabrication plant covering ten routes with up to 583 steps processed by 105 machine groups comprised of 913 machines in total. Times between tool breakdowns are drawn from exponential distributions. Given its role in evaluating novel methods, understanding SMT2020's sensitivity to changes to its parameters is important.

In the present work, we aim to identify small, yet sufficient, sets of machine group attributes to predict their criticality. By ranking machine groups according to their criticality, operators can prioritize maintenance efforts to minimize the risk to the factory's productivity. At the same time, studying the relationship between machine group features and factory-level indicators improves the understanding of SMT2020's behavior and the achievable generality of machine learning models trained using this testbed.

More specifically, we analyze the sensitivity of the factory throughput to decreases in individual machine groups' mean times to failure. Extensive simulation results serve as ground truth data describing the true effect of increased breakdowns. To understand the predictive power of the machine group features, we train linear regression models and neural networks on all combinations of up to 11 features. By evaluating the prediction accuracy and the best-performing features at each feature subset size, we show that when nonlinear relationships are captured, high-quality predictions can be made even based on only a few features. In particular, neural network-driven rankings correctly identify 74% or 81% of the ten most critical machine groups using only two or five features, respectively.

2 RELATED WORK

Identifying the most critical machine groups or machines is key to maintaining factory performance, prioritizing maintenance, and for optimizing resource allocation in semiconductor manufacturing. In the following, we summarize existing research aiming to determine highly predictive features for this purpose.

2.1 Features for Criticality Analysis

Yu et al. (Yu, Li, Zhao, Liu, and Lin 2021) conducted a correlation analysis between fab-level and equipment-level performance indicators based on data from a semiconductor manufacturer in Shanghai, and identified both linear and complex relationships between movement, work-in-process (WIP), and queue length in a certain time period. Anthouard et al. (Anthouard, Borodin, Dauzère-Pérès, Christ, and Roussel 2022) studied machine criticality from the perspective of time constraint tunnel management by comparing a set of static and dynamic criticality measures based on multiple machine features such as processing time and cycle time. They evaluated the effectiveness of each measure using the average Spearman's rank correlation with constraint violations observed in simulations. Dybowski et al. (Dybowski, Sander, and Sprenger 2023) analyzed the impact of WIP, moves, and machine availability on time constraints between processing steps, identifying machine uptime and mean offline time as dominant factors influencing system throughput and stability, and proposed a WIP-based control approach using these features. Mönch et al. (Mönch and Zimmermann 2007) proposed a hierarchical multi-layer approach that integrates multiple machine-level features, including processing time, flow factor, and remaining steps, into weighted machine criticality measures to provide heuristic guidance for shifting bottleneck scheduling and to reduce total weighted tardiness.

Unlike these works, which put a stronger emphasis on time constraint violations, we aim to identify the dominant features on the equipment group level that serve to predict the fab-level throughput and average cycle time. Further, we specifically consider the SMT2020 testbed.

2.2 Non-Linear Feature Analysis

Some existing studies explore the impact of features on system performance using machine learning methods. Senoner et al. (Senoner, Netland, and Feuerriegel 2022) proposed the use of nonlinear modeling (gradient boosting and random forest meta models) with Shapley additive explanations to infer the criticality between the production parameters and the process quality of a semiconductor manufacturing system. In contrast to our work, their focus is on the prediction quality using a wide set of features rather than insights into small, yet highly predictive feature subsets.

Wang et al. (Wang, Zhang, and Wang 2018) proposed a regression-based feature selection method that combines discretization and adaptive logistic regression to identify the key predictors for the fab-level cycle time, and demonstrated that features related to queue lengths, processing time, and machine utilization are the dominant factors. Meidan et al. (Meidan, Lerner, Rabinowitz, and Hassoun 2011) proposed a data-driven method combining conditional mutual information maximization and a selective naive Bayesian classifier based on the SEMATECH dataset, and identified machine idleness, machine availability, number of previous tool loops, and standard deviation of the queue length as the most critical predictors of overall queue time.

Although our predictions rely on much smaller feature sets, these previous results are compatible with our findings, in which we identify the machine groups' queue times, utilizations, and processing times to be among the most predictive features for the effect of increased breakdowns.

2.3 Features for Dispatching and Scheduling

Some studies consider the importance of features indirectly by considering the impact when developing strategies for scheduling and planning to improve system throughput. Kovács et al. (Kovács, Tassel, and Gebser 2023) proposed an optimized dispatching strategy using genetic programming based on hierarchical priorities of lot-level features to improve the throughput. Kopp et al. (Kopp, Hassoun, Kalir, and Mönch 2020a) proposed critical queue time-aware dispatching strategies that combine both lot-level (e.g., remaining process steps, end date) and machine-level (e.g., processing time, flow factor) features to reduce critical queue time violations without large degradations in cycle time or throughput. Ali et al. (Ali, Qaiser, El-Kholany, Eftekhari, Gebser, Leitner, and Friedrich 2024) developed a greedy search-based ant colony optimization algorithm for assigning operations to machines based on the tool group, processing time, and operation sequencing, to solve large-scale flexible job shop scheduling problems in semiconductor manufacturing for minimizing the makespan. Tassel et al. (Tassel, Kovács, Gebser, Schekotihin, Stöckermann, and Seidel 2023) proposed an adaptive scheduling framework for semiconductor manufacturing based on a set of thirteen features (including critical ratio, remaining time, waiting time, and maximum batch size), and employed deep reinforcement learning combined with self-supervised learning to improve fab throughput and reduce completion times. While some of these works employ similar features as ours, their focus is not on identifying predictive feature subsets but on harnessing these features for optimization purposes.

3 METHODS

Our goal is to determine concise predictors for the criticality of machine groups. Here, we define criticality as the negative impact on factory-level statistics if a machine group's unscheduled downtime increases. By prioritizing maintenance based on the machine groups' relative criticality, risks to the factory's productivity can be reduced. On a high level, our approach is to simulate large numbers of what-if scenarios, each representing a sharp increase in a single machine group's probability for tool breakdowns. Subsequently, we study the ability to predict the effects on factory-level statistics purely from (subsets of) the affected machine group's features.

More specifically, in the *data generation* step, we first execute a baseline simulation run. After a warmup period spanning thirty days of simulation time, machine group features are gathered. At termination after one year, we collect the factory-level statistics. Now, we run a large number of what-if simulations using the same random number seed, one per machine group. In each simulation, after the warmup period, the selected machine group's mean time to failure, which applies to all the tools in the group, is divided by four. The differences between the what-if simulations' final factory-level statistics and those from the baseline simulation quantify the impact of the machine groups' higher failure rates. Each comparison to the baseline provides one datapoint comprised of the features of the selected machine group and the resulting change in the factory-level statistics.

This process is repeated for different random number seeds. To increase the diversity in the generated data, we vary the configured machine group processing speeds by multiplying them by random factors chosen uniformly in $[0.9, 1.1]$ for each set of runs with the same seed.

In the *prediction* step, we study the ability to predict factory-level statistics from small subsets of a target machine group's features and to rank machine groups accordingly.

Commonly, L1 and/or L2 regularization are used to encourage sparsity in regression models, common forms being Lasso and Ridge regression (Tibshirani 1996). Here, instead of employing regularization, we take a combinatorial approach and generate regression models for all possible feature sets of a given size. Given N features, this approach requires $\binom{N}{k}$ model generations and evaluations for each feature set

Table 1: List of all considered machine group features and their Pearson correlation with changes in factory-level performance indicators when a machine group's breakdown probability is increased.

Machine Group Feature		**Corr. with Factory-Level Indicator**	
Feature	**Description**	**ΔThroughput**	**ΔCycle Time**
Tool count	Number of equipments	-0.15	0.18
Utilization	Ratio of active time to total time	-0.28	0.32
Average work-in-process	Avg. number of lots in progress	-0.17	0.21
Average queue time	**Avg. time lots wait before processing**	**-0.47**	**0.46**
Average processing time	Avg. time to process a lot	0.08	-0.09
Average cycle time	Avg. total time per lot	-0.27	0.25
Total queue time	**Total time all lots spent waiting**	**-0.56**	**0.58**
Total processing Time	Total time spent on processing	-0.15	0.18
Total cycle time	**Total end-to-end time for all lots**	**-0.50**	**0.53**
Average route position	Avg. position in processed lots' routes	0.17	-0.26
Daily completed lots	Lots finished per day	-0.11	0.18

size k. From this, we can identify the most predictive features at each set size without relying on tunable regularization parameters. To study the dependence of the prediction accuracy on capturing nonlinear relationships, we train both linear regression models and neural networks. Finally, for each k, we select the feature subsets that led to the most accurate predictions of the factory-level statistics.

4 RESULTS

The simulations were carried out using PySCFabSim (Kovács, Tassel, Ali, El-Kholany, Gebser, and Seidel 2022), a Python-based simulator that supports fast just-in-time compiled execution via PyPy. The machine group features were collected after 30 days of warmup time, after which the simulation continued until the end of the year. The factory throughput and cycle time were computed as averages across the entire simulation duration of one year. In total, 14 873 simulation runs were carried out based on 140 unique seeds, resulting in 14 734 pairs comprised of machine group features and changes in factory-level statistics after decreasing the machine groups' mean time to failure by a factor of four from 168 to 42 hours. Our simulations used SMT2020's low-volume/high-mix (LV/HM) dataset which comprises routes for 10 different products. Dispatching was performed using the critical ratio rule.

The machine group features were normalized based on the mean and standard deviation across all machine groups in the same simulation run. All plots show the normalized features. The predictions employ linear regression via Python's SciPy library, and a neural network with a single hidden layer comprised of 64 neurons via PyTorch. The reported prediction accuracies refer to a test set excluded during training, which was comprised of the simulation outputs for one fifth of all covered random number seeds.

We carried out the sensitivity analysis and predictions both for factory-level daily completed lots and dynamic cycle time. However, since the trends and conclusions were largely consistent, most results are reported for the factory throughput only.

We executed the simulations and model training on a machine equipped with an AMD EPYC 7742 processor and 256GiB of RAM running Ubuntu 20.04.6 LTS. The execution time was around 4 minutes per simulation run, resulting in a compute time of about 1000 CPU core hours for the overall ensemble.

4.1 SENSITIVITY ANALYSIS

The complete set of machine group features collected from our simulations is listed in Table 1. For the queue, processing, and cycle times, we gathered both the average per lot that arrived at the respective machine group, and the sum over all lots. As the sum times are the products of the average times and the number of lots, they can be seen as combined features capturing both of these aspects. The average route position was calculated by averaging over the relative positions of all arriving lots on their respective routes, making it a dynamic rather than static machine group feature. The table also shows the correlation

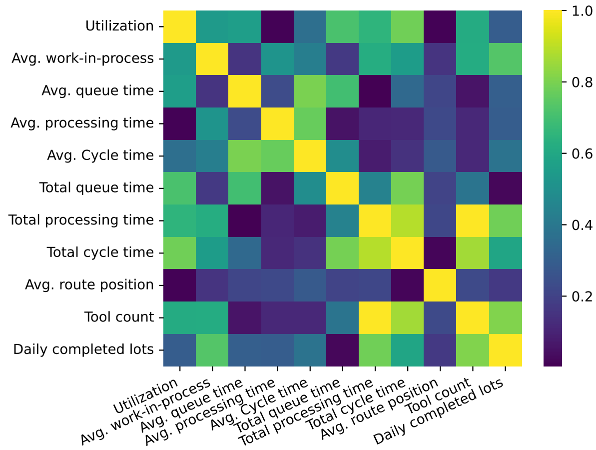
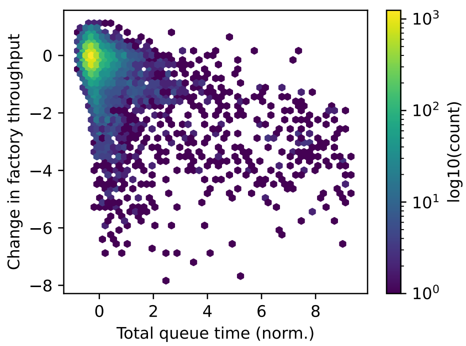
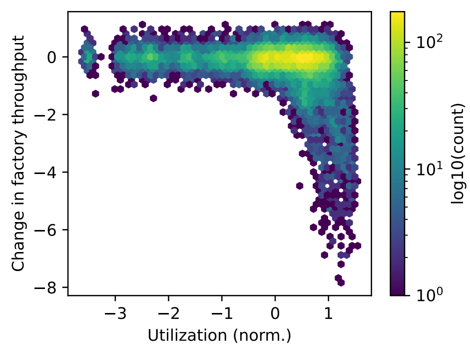


Figure 1: Absolute Pearson correlation among the machine group features.



(a) Total queue time, $\rho = -0.56$.



(b) Utilization, $\rho = -0.28$

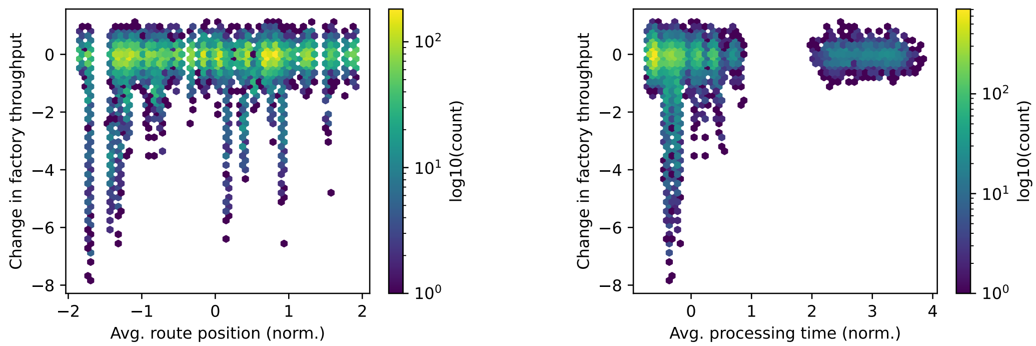
Figure 2: Relationships between machine group features and change in factory throughput.

of the machine group features with the observed change in factory throughput and average cycle time when a machine group's breakdown probability is increased.

To better understand the overall set of machine group features, we study the correlation among all pairs of features, which is shown in Figure 1. As we are interested in the magnitude of the correlation, we show the absolute values of the Pearson correlation coefficients. First, we find that many of the machine group features are largely uncorrelated or only weakly correlated. In line with expectations, the most strongly correlated features pairs are tool count and total processing time, total queue and total cycle time, and utilization and average work-in-process.

We next consider the relationships between individual machine group features more closely together with the effects of factory throughput when a machine group's breakdown probability is increased. Figure 2 shows binning plots for the machine group features of total queue time or utilization, illustrating both the density of the data points and their correlation, with the Pearson correlation coefficient ρ shown in the caption. We can see that there is a substantial linear component to the relationship between total queue time and the change in throughput, although a significant spread remains for any given value on the x-axis. In contrast, the change in throughput depends on the machine group's utilization in a highly nonlinear way. Pronounced changes in throughput is only seen at the highest utilizations, and even at very high utilizations there are cases where no substantial change is observed. In other words, high utilization is necessary, but not sufficient, for a strong effect on throughput.

Distinct patterns are observed in the plots for average processing time and route position shown in Figure 3. The average processing time exhibits two clusters, the cluster with the higher processing times



(a) Route position, $\rho = 0.17$ (b) Avg. processing time, $\rho = 0.08$

Figure 3: Relationships between machine group features and change in factory throughput.

Table 2: Best-performing machine feature subsets to predict factory throughput.

Linear regression		
#Features	MSE	Features
1	0.454	Total queue time
2	**0.434**	**Total queue time, route position**
3	0.416	Avg. queue time, total queue time, route position
4	0.409	Avg. processing time, avg. cycle time, total queue time, route position
5	**0.407**	**Avg. processing time, avg. cycle time, total queue time, route position, tool count**

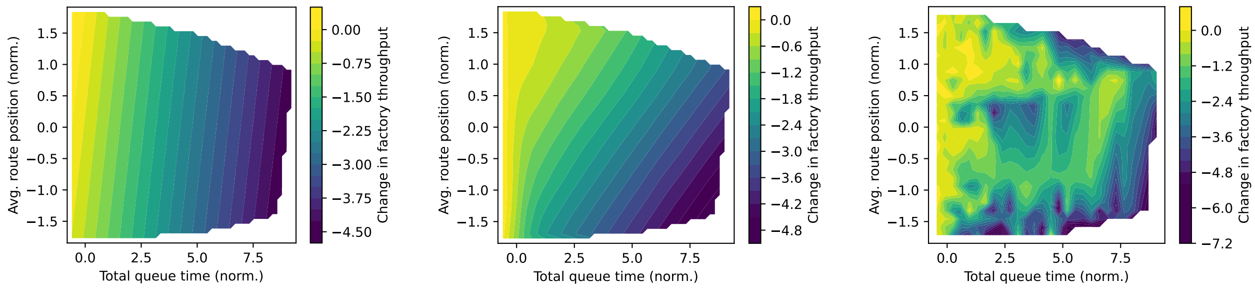
Neural network		
#Features	MSE	Features
1	0.441	Total queue time
2	**0.277**	**Utilization, avg. processing time**
3	0.210	Utilization, avg. processing time, route position
4	0.176	Utilization, avg. work-in-process, avg. processing time, route position
5	**0.170**	**Utilization, avg. work-in-process, avg. processing time, route position, tool count**

Table 3: Best-performing machine feature subsets to predict factory cycle time

Linear regression		
#Features	MSE	Features
1	3.351	Total queue time
2	3.020	Total queue time, route position
3	2.914	Avg. queue time, total queue time, route position
4	2.869	Avg. processing time, avg. cycle time, total queue time, route position
5	2.837	Avg. processing time, avg. cycle time, total queue time, route position, tool count

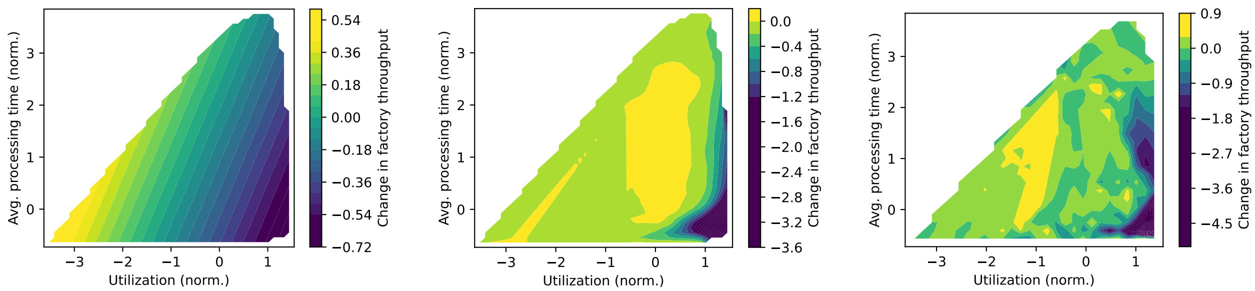
Neural network		
#Features	MSE	Features
1	3.139	Total queue time
2	1.691	Avg. processing time, total cycle time
3	0.900	Utilization, avg. processing time, route position
4	0.601	Utilization, total processing time, route position, daily lots
5	0.596	Utilization, avg. queue time, total processing time, route position, daily lots

being defined by the diffusion groups. The largest change in factory throughput is caused by machine groups with relatively low processing times. Considering the relationship between the route position and the change in throughput, a gradually diminishing effect would be expected with increasing route positions, as increased queuing in the beginning of a route is likely to propagate downstream. While the largest changes in throughput are seen at lower route positions, the scattered vertical lines suggest that other factors dominate.



(a) Linear regression (b) Neural network (c) Ground truth

Figure 4: Predictions based on total queue time and route position compared to ground truth.



(a) Linear regression (b) Neural network (c) Ground truth

Figure 5: Predictions based on utilization and average processing time compared to ground truth.

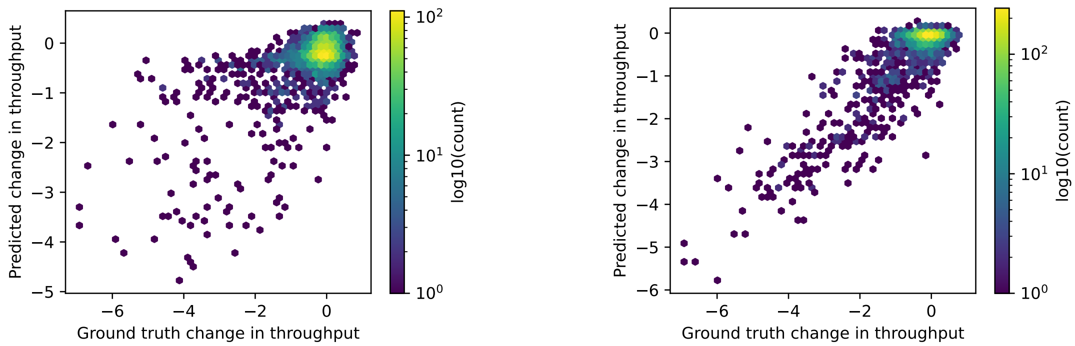
4.2 CRITICALITY PREDICTORS

We now consider the predictive power of restricted subsets of the machine group features. Table 2 shows the prediction accuracy for the factory throughput achieved using up to five features, which was determined by training models using all possible feature subsets and selecting the subset with the lowest mean squared error (MSE).

We are interested both in the achievable accuracy and in most predictive feature subsets. First, we observe that the prediction accuracy using a single feature is similar between the linear regression and neural network model, and that in both cases total queue time allowed for the best prediction. Both of these observations are in line with the sensitivity analysis as the relationship between total queue time and the change in throughput has a strong linear component, allowing for reasonable predictions even with the linear model.

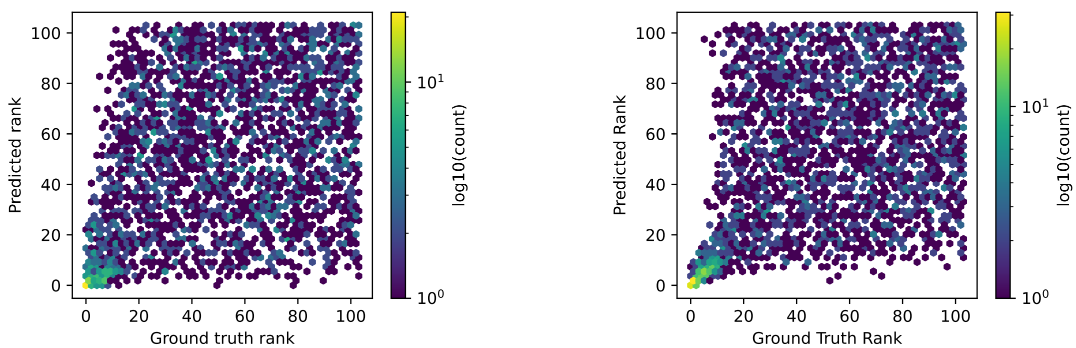
As the number of features increases, we see that the prediction quality of the linear regression model increases only slightly, while the neural network's accuracy improves sharply. The most predictive feature subsets also differ significantly. In line with the sensitivity analysis, the linear regression model remains reliant on the total queue time even with multiple features, whereas the neural network employs the largely nonlinear machine group utilization together with the average processing time and average route position. The results for the factory's average cycle time shown in Table 3 follow the same trends.

Figure 4 visualizes the predictions using the machine groups' total queue time and route position. It is evident that the predictions of the linear regression model (MSE: 0.43) depend largely on the total queue time, with only a minor influence by the route position. The neural network (MSE: 0.38), on the other hand, more closely approximates the ground truth results. In Figure 5, we see the same comparison when predicting the change in throughput using the machine groups' utilizations and average processing times. Here, linear regression (MSE: 0.48) is clearly insufficient to capture the localized peak when high



(a) Linear regression (b) Neural network

Figure 6: Comparison of factory-level statistic predicted via five best-performing machine group features and ground truth.



(a) Linear regression (b) Neural network

Figure 7: Comparison of machine ranks predicted via five best-performing machine group features and ground truth. The most critical machines are on the left.

utilizations coincide with low processing times. For this feature combination, the neural network achieved its best two-feature accuracy, with an MSE of 0.28.

We now compare the prediction accuracies more directly by plotting binned predictions of the change in factory throughput against the ground truth. Figure 6 shows the results for five-feature predictions based on the best-performing feature sets using linear regression and neural networks. We see that in a large number of cases, there is only a negligible change in throughput, which both models predict reasonably well. As the impact on the throughput increases, the difference in prediction quality is visible, with the linear regression results becoming more dispersed and only in rough accordance with the data. Compared to the neural network, the linear regression model's inability to capture nonlinearities translates to a nonlinear relationship between prediction and ground truth.

Based on the predictions, we can now rank the machine groups by their criticality to study the ability of the models to guide maintenance prioritizations. In Figure 7, we see that in accordance with the previous results, the neural network is substantially more successful in correctly ranking the machines. The light colors at the bottom right show that for the most critical machines, the predictions and ground truth ranks are in close alignment. As the machine groups' criticalities decrease, the change in throughput approaches zero and the relative ranking becomes increasingly volatile and hard to predict, but also less relevant to operators. To more directly quantify the quality of the rankings, we calculated the intersection of ground truth and predicted top k ranks for each of the 140 unique seeds, with $k \in \{5, 10, 20\}$. Using the five best-performing features, the linear regression's ranking agreed with the ground truth for averages of 2.4, 5.6, and 10.1 machine groups, respectively, corresponding to 48%, 56%, and 50% of the machine groups. Using the neural network, averages of 3.9, 8.1, and 13.0 of the top k machine groups were in agreement, corresponding to 78%, 81% and 65% of correctly identified critical machine groups. Even with only two features – utilization and average processing time – the neural network is in accordance with the ground truth for 66%, 74%, and 63% of the machine groups.

When increasing the number of features beyond five features, the prediction accuracy using the linear regression model stagnates around an MSE of 0.39 from eight features onwards. The neural network improves up to eight features, reaching an MSE of 0.16. Remarkably, the neural network predictions based on only two features outperform the linear regression model using all 11 features. Since neither of the models substantially improved beyond five features, we conclude that with the chosen overall feature set, a small feature subset suffices as long as nonlinearities are captured.

5 CONCLUSIONS

We presented a sensitivity analysis of the SMT2020 testbed under increased tool breakdown rates. From the results of a combinatorial evaluation of various machine group features' predictive power, we observed that two to five machine group features suffice to guide reasonably accurate rankings of the machine groups by their criticality to factory-level performance indicators. Out of the 11 considered machine group features, the total queue time was the most predictive individual feature. Since the relationships between features and performance indicators are highly nonlinear, a neural network trained on only two machine group features – utilization and average processing time – outperformed a linear regression model using our full machine group feature set.

A potential limitation of our study is that certain features may identify critical machine groups directly, which would make the model specific to the SMT2020's topology and configuration. Potential indications for this were observed in the form of distinct patterns when visually studying the relationship between features such as the relative positions of machine groups along a route and changes in the factory throughput. These patterns could indicate hidden variables not represented directly in our feature set. Considering additional features could allow for more accurate predictions, e.g., by accounting for the effect of required batch sizes and the associated dynamic waiting times explicitly.

Further, although we increased the diversity of the model training data by varying SMT2020's machine speeds, the fixed route structures and job arrival patterns may persistently lead to certain machine groups being the most critical ones. Nevertheless, even with these limitations, we believe that our findings further the understanding of SMT2020's sensitivity, which is important given the model's central role in evaluating manufacturing optimization methods.

A natural future direction is to study extensions of SMT2020 such as SMAT2020 (Lee, Jeon, and Park 2023), which integrates automated material handling systems in the simulation model. Our analysis focused on the factory-level throughput and average cycle time, with the high correlation between the two leading to largely the same conclusions. Real-world operations must consider aspects such as product due dates, temporal constraints between process steps (Kopp, Hassoun, Kalir, and Mönch 2020a), and profitability concerns. An important avenue for future work is to study whether our results extend to predictions of more complex performance indicators.

ACKNOWLEDGMENTS

This research is supported in part by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG3-RP-2022-031) and in part by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG2-100E-2024-130).

REFERENCES

Ali, R., S. Qaiser, M. El-Kholany, P. Eftekhari, M. Gebser, S. Leitner *et al*. 2024. "A Greedy Search Based Ant Colony Optimization Algorithm for Large-Scale Semiconductor Production". In *Proceedings of the 14th International Conference on Simulation and Modeling Methodologies, Technologies and Applications - Volume 1: SIMULTECH*, 138–149. INSTICC: SciTePress https://doi.org/10.5220/0012813100003758.

Anthouard, B., V. Borodin, S. Dauzère-Pérès, Q. Christ, and R. Roussel. 2022. "Criticality Measures for Time Constraint Tunnels in Semiconductor Manufacturing". In *2022 Winter Simulation Conference (WSC)*, 3326–3337. IEEE.

Dybowski, N., M. Sander, and R. Sprenger. 2023. "Optimization of Timelinks in Semiconductor Manufacturing". In *2023 Winter Simulation Conference (WSC)*, 2160–2171. IEEE.

Gopalakrishnan, M., and A. Skoogh. 2018. "Machine criticality based maintenance prioritization: Identifying productivity improvement potential". *International Journal of Productivity and Performance Management* 67(4):654–672.

Kopp, D., M. Hassoun, A. Kalir, and L. Mönch. 2020a. "Integrating critical queue time constraints into smt2020 simulation models". In *2020 Winter Simulation Conference (WSC)*, 1813–1824. IEEE.

Kopp, D., M. Hassoun, A. Kalir, and L. Mönch. 2020b. "SMT2020—A semiconductor manufacturing testbed". *IEEE Transactions on Semiconductor Manufacturing* 33(4):522–531.

Kovács, B., P. Tassel, R. Ali, M. El-Kholany, M. Gebser, and G. Seidel. 2022. "A customizable simulator for artificial intelligence research to schedule semiconductor fabs". In *33rd Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, 1–6. IEEE.

Kovács, B., P. Tassel, and M. Gebser. 2023. "Optimizing Dispatching Strategies for Semiconductor Manufacturing Facilities with Genetic Programming". In *Proceedings of the Genetic and Evolutionary Computation Conference*, 1374–1382.

Lee, K., W. Jeon, and S. Park. 2023. "Multi-Resolution Modeling Method for Automated Material Handling Systems in Semiconductor FABs". In *2023 Winter Simulation Conference (WSC)*, 2345–2356. IEEE.

Meidan, Y., B. Lerner, G. Rabinowitz, and M. Hassoun. 2011. "Cycle-time key factor identification and prediction in semiconductor manufacturing using machine learning and data mining". *IEEE transactions on semiconductor manufacturing* 24(2):237–248.

Mönch, L., and J. Zimmermann. 2007. "Simulation-based assessment of machine criticality measures for a shifting bottleneck scheduling approach in complex manufacturing systems". *Computers in Industry* 58(7):644–655.

Senoner, J., T. Netland, and S. Feuerriegel. 2022. "Using explainable artificial intelligence to improve process quality: evidence from semiconductor manufacturing". *Management Science* 68(8):5704–5723.

Tassel, P., B. Kovács, M. Gebser, K. Schekotihin, P. Stöckermann, and G. Seidel. 2023. "Semiconductor Fab Scheduling with Self-Supervised and Reinforcement Learning". In *2023 Winter Simulation Conference (WSC)*, 1924–1935. IEEE.

Tibshirani, R. 1996. "Regression shrinkage and selection via the lasso". *Journal of the Royal Statistical Society Series B: Statistical Methodology* 58(1):267–288.

Wang, J., J. Zhang, and X. Wang. 2018. "A data driven cycle time prediction with feature selection in a semiconductor wafer fabrication system". *IEEE Transactions on Semiconductor Manufacturing* 31(1):173–182.

Yu, Q., L. Li, H. Zhao, Y. Liu, and K.-Y. Lin. 2021. "Evaluation system and correlation analysis for determining the performance of a semiconductor manufacturing system". *Complex System Modeling and Simulation* 1(3):218–231.

AUTHOR BIOGRAPHIES

PHILIPP ANDELFINGER is a Senior Research Fellow in the College of Computing and Data Science at Nanyang Technological University (NTU), Singapore. He received his Ph.D. degree in computer science from the Karlsruhe Institute of Technology, Germany, in 2016. His research focuses on simulation-based optimization, differentiable simulation, and parallel simulation on heterogeneous hardware platforms. His email address is philipp.andelfinger@ntu.edu.sg.

SHUAIBING LU is a lecturer in the College of Computer Science, Beijing University of Technology, Beijing, China. She received her Ph.D. degree in computer science and technology from Jilin University, Changchun, in 2019. She is supported by the China Scholarship Council as a visiting scholar supervised by Prof. Wentong Cai in the College of Computing and Data Science at NTU, Singapore. Her research focuses on distributed computing, cloud computing, and edge computing. Her email address is lushuaibing@bjut.edu.cn.

CHEW WYE CHAN is a Principal Software Engineer at D-SIMLAB Technologies (Singapore). He holds a Master of Computing degree from the National University of Singapore and a Ph.D. degree in computer science from Nanyang Technological University, Singapore. His research interests include machine learning, data science, and simulation-based optimization. His email address is chew.wye@d-simlab.com.

FEI FEI ZHANG is a Product Manager at D-SIMLAB Technologies (Singapore). He is responsible for lot scheduling and dispatching products for customers in the semiconductor industry. He earned a Bachelor of Computer Science from the National

University of Singapore. His email address is zhang.feifei@d-simlab.com.

BOON PING GAN is the CEO of D-SIMLAB Technologies (Singapore). He has been involved in simulation technology application and development since 1995, with a primary focus on developing parallel and distributed simulation technology for complex systems such as semiconductor manufacturing and aviation spare inventory management. He was also responsible for several operations improvement projects with wafer fabrication clients which concluded with multi-million dollar savings. He holds a Master of Applied Science degree, specializing in Computer Engineering. His email address is boonping@d-simlab.com.

JIE ZHANG received the Ph.D. degree from the Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, Canada. He is currently a Professor with the College of Computing and Data Science, Nanyang Technological University, Singapore. He has also been appointed as an Adjunct Fellow of Singapore Institute of Manufacturing Technology (SIMTech), A*STAR, since 2020. His papers have been published by top journals and conferences and won several best paper awards. He was a recipient of the Alumni Gold Medal in 2009. He is also active in serving research communities. His email address is zhangj@ntu.edu.sg.

WENTONG CAI is a Professor in the College of Computing and Data Science at Nanyang Technological University, Singapore. He received his Ph.D. in Computer Science from University of Exeter (UK) in 1991. His expertise is mainly in the areas of Modeling and Simulation and Parallel and Distributed Computing. He has published extensively in these areas and has received a number of best paper awards at the international conferences for his research in distributed simulation. He is an Associate Editor of the ACM Transactions on Modeling and Computer Simulation (TOMACS), an Editor of the Future Generation Computer Systems (FGCS), and in the Editorial Board of the Journal of Simulation. His email address is aswtcai@ntu.edu.sg.