Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

SIMULATION BASED EVALUATION OF DIFFERENT EMPTY VEHICLE MANAGEMENT STRATEGIES WITH CONSIDERING FUTURE TRANSPORT JOBS

Robert Schmaler
Thorsten Schmidt

Chair of Logistics Engineering
Technische Universität Dresden
Münchner Platz 3
Dresden, 01187, GERMANY

Matthias Schoeps
Joerg Luebke
Ralf Hupfer
Nikolas Schlaus

Factory Automation Support
GLOBALFOUNDRIES
Wilschdorfer Landstraße 101
Dresden, 01109, GERMANY

ABSTRACT

Even in today's semiconductor Fab the Automated Material Handling System (AMHS) is still the part one pays too little attention to. With having in mind that all the added value is done on the process tools this opinion might be comprehensible. But without the Overhead Hoist Transport (OHT) vehicles to be at the right time at the right place semiconductor manufacturing could not work efficiently. The following paper describes new empty vehicle management strategies as an important part of this on time AMHS delivery. Information about future upcoming transport jobs will be included to allocate this limited resource proactively and to achieve goals such as minimizing the tool waiting time for empty vehicles or the total number of dispatch moves. Different scenarios with changing input parameters will be tested and compared by using a simulation model which was developed with the focus of representing empty vehicle balancing functionalities.

1 INTRODUCTION

Semiconductor manufacturing is a highly expensive industry. The majority of the investment is for different process and measurement tools. A single tool may cost more than $10 Million. The number of tools can reach more than 300 within a single manufacturing building. Compared to the production system, the installation of a transport system seems to be quite inexpensive ($100-200 Million). As a consequence, the transportation system is often not the focus of attention.

As stated in Chen, Wang, and Chan (2017), the production capacity is limited to what the transport system can handle. This assumption can be supported by the following numbers. Since a single wafer has to undergo more than 400 process steps together with a varying number of measurement steps and transportation into different storages in between, it has to be handled by the transport system very often. Current Automated Material Handling Systems (AMHS) can handle more than 3000 transport steps per hour, as can be seen in Schmaler et al. (2016). Thus, with increasing demands of fab throughput also AMHS performance has to increase to fulfill production demands.

One of the main requirements for a complex transport system is the capability to react to constantly changing system states. A significant amount of empty vehicle dispatching moves has to be initiated to achieve this requirement. The semiconductor production system typically has tool and wafer recipe specific processing times. Although some tools have uncritical carrier exchange times (CET), low utilized tools can be dangerous for the competitiveness of a whole fab. CET describes the time from FOUP pickup request at a load port to the delivery of a new FOUP with a processing lot to the same load port. To

978-1-5386-3428-8/17/$31.00 ©2017 IEEE

achieve the goal of highly utilized tools and an optimized yield, on time delivery has to be guaranteed. One key factor to achieve this goal is the fast retrieval of FOUPs from the tools after a process completes. Therefore empty vehicle dispatching plays an important role in this topic.

Empty vehicle balancing in a semiconductor fab mainly focusses on two objectives. On the one hand is the instant availability of vehicles to pick up waiting FOUPs to reduce or even delete transport related tardiness for the production system. On the other hand, the total amount of dispatching moves has to be as low as possible to minimize negative influence on the transport of time critical production wafers. In other words, as much AMHS capacity as possible should be available for the production system to transport FOUPs and not to move empty vehicles between bays. The rail guided transport system reaches capacity limits with a certain amount of vehicle moves per time unit. Transportation times increase especially at critical points as at crossings of main tracks if vehicle traffic increases and vehicles hold each other back. Even a worst case scenario is possible where vehicles block each other into deadlock situations and the whole transportation process stops. Minimized number of empty vehicle moves also leads to a lower number of vehicles needed for a specific throughput. That is beneficial for the vehicle budget as well as for the robustness of the system. The less vehicle resources get lost due to empty vehicle balancing, the higher vehicle utilization is possible without seeing any impact on the retrieval times and thus increasing fab throughput.

Even though the requests for transport vehicles are highly dynamic, current approaches as used in running fabs mostly distribute empty vehicles within different areas regarding preset limitation values. These values usually include minimum amounts of empty vehicles for each area (Low Water Mark – LWM) as well as maximum vehicle amount (High Water Mark – HWM). These values are usually based on empirical knowledge, are static and therefore not able to react to changing system states automatically. That means the current fab situation is not reflected within these values. Looking at already available, but not yet used detailed system information this approach does not seem to meet today's possibilities anymore. Semiconductor manufacturing with its complex processes and control systems seems to be a prime example of all the potentials and developments coming along with keywords as "Industry 4.0" as well as "Big Data". Different systems from different control layers communicate with each other and produce a huge amount of data which mostly is not used to optimize vehicle control strategies yet. Even though some examples already exist showing how an integration of information from the transport into production system as well as the other way around can be beneficial for both sides. Examples can be found in Hammel et al. (2016) where transport amounts using AMHS-bottlenecks are reduced to increase transport system throughput and in Driessen et al. (2016) where existing information are used to increase tool utilization together with better vehicle dispatching. Approaches like that cannot be found in the field of empty vehicle management strategies yet.

Complexity and multi-layer information system structure (process tools, RTD, MES, MCS) of a semiconductor manufacturing facility lead on one hand to the availability of all kinds of information, but is on the other hand the reason why the exchange of suitable information is that complex, too. Detailed information and current states of processing tools are available to the responsible system engineer (e.g. equipment engineer for process time information). But already at the next higher level responsible for wafer distribution among different process and measurement tools (Real-Time Dispatching – RTD) all of these detailed process information like process time and process end time are either not available or simply not used. Tool throughput is calculated based on the simulation with self-developed tool models and wafers are assigned to average values resulting from these models.

The same problem can be observed by looking at the connection of the transport system to the level above. Neither information about when a process tool needs an empty vehicle nor information about when an empty or loaded vehicle will arrive at the destination port is currently considered within dispatching decisions. Transport request rates and sequence are highly dependent on the tool type (single wafer, batch tools, multi-chamber, etc. – for more information see Wang, Wang, and Chueh (2016)). But still, the significant amount of transport requests starting at a tool (which is about the half of the total transports generated in a fab) can be forecasted within a suitable time range because they are already available. At least this kind of transport information could be used to optimize empty vehicle management strategies. The focus of this paper is to show a new empty vehicle dispatching approach using forecast data and how this additional information could be beneficial for the transport system. Furthermore different degrees of information availability and the influence on empty vehicle dispatching strategies will be investigated using simulation experiments.

The next sections of this paper are organized as follows: Section 2 shows an introductory example about empty vehicle dispatching strategies and how they influence AMHS throughput. The literature is reviewed regarding current approaches dealing with this topic in Section 3. In Section 4 the algorithm to dispatch empty vehicles based on forecast information is introduced as well as different vehicle types and dispatching objectives. The next part presents explanation of a simulation tool for evaluating different empty vehicle management strategies together with the results of different simulation experiments regarding available information and optimization goal (Section 5). The paper finishes with a conclusion and gives an overview of future research topics.

2 EXAMPLE

To present the reader a first rough idea about empty vehicle management properties and necessity a comparison of three basic scenarios will be described next. The first scenario is the simulation of vehicle movement based on real fab transport data without any dispatching functionality (S1). Vehicles will only move based on their transport requests. The second one introduces first simple dispatching ideas. As an extension to S1, the second one calls empty vehicles from other bays when needed. That means if a transport request can't find an empty vehicle in the current bay, the dispatcher searches for empty ones through the other bays (S2). The search order is based on a simple time matrix between all existing bays. The third scenario in turn is an extension of scenario S2, but empty vehicles won't be called if other loaded vehicles are already on their way to the searching bay (S3). The reference transport system is a unified one with the following properties:

- 25 bays (20 intra and 5 inter) along 12 km of track
- 500 tools with 10.000 storage locations
- 3.500 transports per hour achieved by 300 vehicles

The simulation model used for the studies in this paper is an event triggered Java model which especially was developed to compare different empty vehicle management strategies. The reason why this was favored over a full AMHS dynamic simulation model is that other influences like traffic jams or layout irregularities (tools placed at bay exits) do not dilute the core of the dispatching functionality in the Java model. These and interdependent functionalities of the real system weaken the evaluation of the empty vehicle dispatching strategy. The model hast been validated by running different pre-defined test scenarios and comparing the results of the model with the expectations. Performance parameters used to compare the different simulation scenarios are the fulfillment of production transports, the average wait time for empty vehicles, the amount of empty vehicle dispatch moves made, a counter how often no vehicle was either in the bay or on the way when a transport request came up (missing vehicle) and finally, how often no vehicle was physically in the bay but at least on the way (waiting transports). As easily seen in Figure 1 without any dispatching strategy the throughput demand can't be fulfilled. From S2 and S3 the two different objectives of minimization of dispatch moves and reduction of wait time are already achieved differently, too. On the one hand, S2 initiates more dispatch moves because no matter if vehicles are heading to the necessary bay, if they have not arrived already new vehicles will be dispatched. On the other hand, these dispatch moves result in a reduced waiting time compared to S3, where dispatching is not initiated if there are vehicles which are heading to the necessary bay. This results in less dispatching transports.

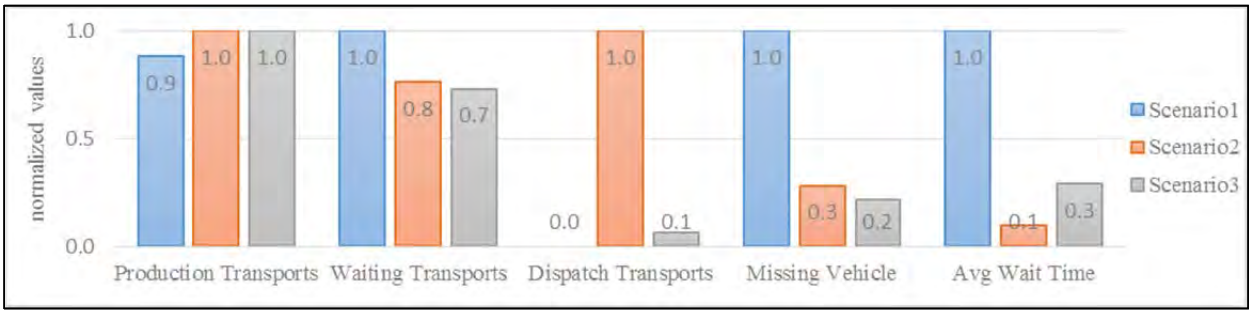

Figure 1: Comparison no dispatching vs basic dispatching strategies (normalized values where 1.0 is maximum value and other values as ratios according to maximum value).

Regarding the experiments above a new strategy to fulfill both objectives of minimizing wait time and dispatch moves has to be found. According to this question, the following section describes existing research in the field of vehicle dispatching and vehicle fleet sizing.

3 LITERATURE REVIEW

Lin, Wang, and Wu (2003) introduced a classification of four vehicle types regarding different transport demands (intrabay only, between interbay stocker or bay to bay). In the extension by Lin, Wang and, Young (2004) the possibility of virtual vehicles was added where the vehicle type can change and is not fixed anymore. That means vehicles can be related to a certain type depending on the demand. In this paper, control limits are mentioned, but it is not told how to set them. Additionally, vehicles are all assigned to a fixed bay and they have to go back to their home bay after delivering anywhere else. One main problem of this approach which makes it quite inefficient to use for a unified system is the high part of vehicle utilization needed for empty vehicles to go back home. Instead of using different vehicle types Kiba et al. (2010) compared the classical nearest neighbor dispatching strategy to a so-called minimum service policy. This policy defines Low Water Marks (LWM) which serve as a lower control limit to define minimum necessary available vehicles per bay. If the vehicle amount falls below this control limit vehicles can be called from other bays which have more vehicles than defined by the LWM. Carrier Exchange Time (CET) was used to compare both dispatching strategies. The study shows that minimum service policy outperforms the classical strategy. The authors also state that LWMs of different bays have interdependencies as well as they can influence CET of tools, and therefore fab performance. But it is neither shown how these interdependencies are considered nor how the LWMs are set. Additional High Water Marks (HWM) where mentioned in Jimenez et al. (2010) to lower the risk of creating traffic jams because of too many vehicles within a single bay. If HWM value is exceeded, bay pushes out empty vehicles to other bays. Two strategies called bay dedication and zone balancing are compared. Bay dedication forces vehicles to go back to source bay (see Lin, Wang and, Young 2004) whereas zone balancing parks empty vehicles right where they finished their last transport request. According to the control limits (LWM, HWM), dynamic vehicle exchange is possible to balance empty vehicles. The water mark based strategy outperforms the other one for each performance indicator. That is the reason why this approach is currently used in most fabs. However, the topic of water mark calculation and bay dynamics was first directly addressed by Chaabane et al. (2013). Again a minimum service policy was used with including a degree of freedom to reduce vehicle amount available for empty vehicle dispatching. LWMs are calculated based on the same proportion as bay transports to total transports, reduced by the degree of freedom. This approach was even extended by including system dynamics (transport leaving vs. transports staying in the same bay). But nothing is written about the way of vehicle exchange when limits are reached. This was mentioned in Lin, Wu, and Huang (2013) where they introduced a Markov decision process model to solve the vehicle assignment problem dynamically. To increase calculation efficiency

Dynamic Vehicle Allocation Control (DVAC) to assign the optimal vehicles amount to each bay is described. This mathematical model solves for each possible system state the problem if it is better to dispatch empty vehicles or not. The objective is to reduce average FOUP waiting times. This approach sounds quite promising, but the example scenario only implies two bays. Even for this small scenario decisions cannot be made in real time but have to be calculated before and stored in a decision table. The authors remark that this approach reaches its calculation limits quite soon when considering bigger systems.

Other approaches deal with the vehicle fleet sizing problem by using simulation optimization. For example, Huang, Chang, and Lin (2012) tried to find the optimal vehicle allocation for each bay. They adapted the Convergent Optimization via Most-Promising-Area Stochastic Search (COMPASS) Algorithm to reduce the number of necessary simulation runs to find an optimal solution. They show good solutions to find an initial set of vehicles per bay, but with no functionality to react to changing system states without repeating the whole scenario of simulation optimization. The same problem can be obtained by other approaches as in Chang, Huang, and Yang (2014). They used the Simulation Sequential Metamodeling (SSM) method that is based on functional relations between input and output of a simulation model as for example the vehicle amount and the corresponding transport time. The meta-models are built to find the minimum necessary fleet size while fulfilling production based process to process time constraints. The reason why this approach is not applicable for a unified system is that the assumptions for bay wise meta-models do not hold with higher system dynamics. Other approaches focus on how to reduce necessary amount of simulation runs to find optimal fleet size by using evolutional algorithms as in Lin and Huang (2013). They aimed to find the right number of vehicles for each bay with the objective to maximize AMHS throughput. A Genetic Algorithm (GA) supported by Optimal Computing Budget Allocation (OCBA) to allocate simulation budget to different alternatives as defined by the GA was used. For large scale problems where not every single alternative can be simulated this combination generated good solutions in a reasonable time. They were able to reduce simulation time up to 60% with including OCBA to GA compared to GA only. But nothing is said about how to deal with stochastics when optimizing with GA only. Lin and Huang (2014) conducted a similar approach by using OCBA together with Particle Swarm Optimization (PSO) to find the best combination of vehicle amount and dispatching rules to run a photobay.

The importance of the current topic can be seen by looking at approaches as in Huang and Lin (2016) where the Pre-Dispatching Vehicle (PDV) method is described. The method analyzed for batch tools in the diffusion area the benefit of simultaneously assigning more than one vehicle to pick-up FOUPs from one tool or stocker load port only. These so called form-batch operations require more than one transport initiated at the same time, depending on the Work in Process (WIP) of the tool. But this kind of optimization approach is only applicable when empty vehicles are already in the bay.

To show the improvement of currently used water mark strategies compared to already presented basic scenarios (S2, S3), Figure 2 gives an overview of performance improvements. It can be seen that water mark based dispatching scenario (S4) increases dispatch transports, but leads to only 60% of waiting transports and a 20% reduced average waiting time compared to S3. In S5 additional rule sets determining source bays for dispatch moves have been included. This scenario represents the state of the art real system scenario and outperforms the other ones significantly. Thus, S5 will be the benchmark scenario for evaluating the new proposed algorithm using forecast data as described in the next section.

4 DISPATCH ALGORITHM WITH FORECAST HORIZON

This section introduces a new empty vehicle dispatching strategy which includes the possibility of using tool events as forecasts for dispatching decisions. The main problem of empty vehicle allocation decisions in wafer fabrication facilities, is the unpredictable behavior of the production system. The presented approach tries to show the benefit of implying at least minimum available tool request information and also introduces a solution for having all information regarding a short look ahead in the future available.

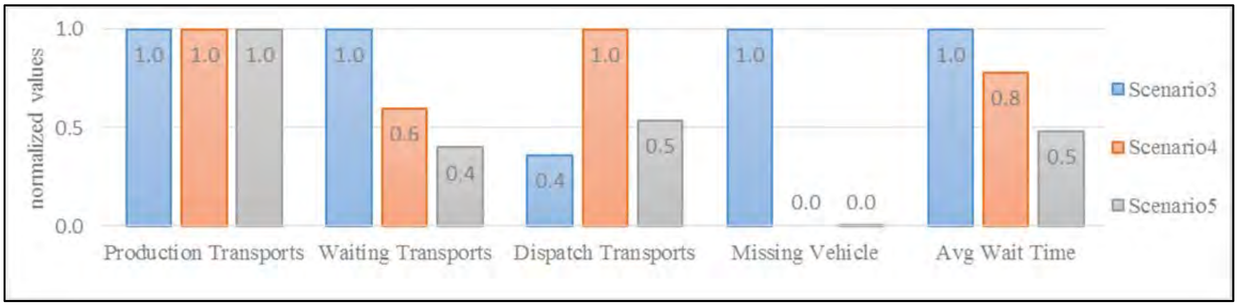

Figure 2: Comparison of basic dispatching strategy vs. water mark based strategies (normalized values where 1.0 is maximum value and other values as ratios according to maximum value).

This should be feasible with all the possibilities a semiconductor manufacturing execution system has.

As already mentioned earlier, empty vehicle dispatching mainly focuses on two objectives: minimizing waiting time and minimizing empty vehicle balancing moves. Regarding the degree of information availability, three scenarios will be displayed in this paper: having no forecast information, tool event information only and all information regarding a certain time span available. For information which is not available statistical values will be used. Table 1 shows the definition of different vehicle types regarding their status (idle, incoming, outgoing) and additional forecast information (e.g. time of arrival). The table also shows which vehicle type is available for which scenario as well as which type is necessary to achieve the different objectives of minimizing wait time or empty vehicle dispatch moves.

Idle vehicles are not loaded without an assigned transport job. These vehicles may be moving between parking points within a bay. Incoming vehicles are loaded with the destination of the transport job within the bay they are assigned to. A distinction is made between incoming vehicles type (1) and type (2) regarding their estimated time of arrival. Incoming(1) vehicles will arrive within a forecast horizon (FH) and Incoming(2) after that. The forecast horizon is the time span for which future transport job information is available. Experiments have shown that the average vehicle dispatch time is a suitable FH. According to equipment engineers, an average value of about 60s is also feasible to at least forecast transport request starting at tools. FutIncoming are attached to a transport job which is not yet generated at the decision point (t_0), but will be generated within the FH. Again this type of incoming vehicles is subdivided into FutIncoming(1) and FutIncoming(2) depending if the estimated time of arrival is before or after the end of the FH. It is clear that the FutIncoming types will probably be the hardest to get information about. But still, within a short forecast horizon even this information should be possible to get. FutOutgoing refers to vehicles where transport jobs will be created between t_0 and end of FH. FutIncoming vehicles require that not only information about source, but also about destination of future transports is available. For example a transport starting in Bay X with destination in Bay Y, starting after t_0 and estimated time of arrival after FH will be added to Bay X as FutOutgoing and to Bay Y as FutIncoming(2). If the destination information for future outgoing transports is available it necessarily results in a future incoming (1 or 2) transport for the destination bay. If this information is not available the vehicle type of future incoming is not available, too.

Figure 3 summarizes the different vehicle types in a time line with t_0 as decision point and t_0+FH as the end of the forecast horizon. The availability of the described vehicle types can be summarized as follows: Idle vehicles and Incoming(1+2) with their estimated time of arrival are already available (with the estimated time of arrival not being used yet). FutOutgoing transports can be determined by using tool events. The FutIncoming transports are kind of a future scenario because at first the data collection process where the FOUPs get their next destination would have to be revised. But in practice some tools using an early unload scenario already have this information (for more information refer to Rothe et al. 2015).

Table 1: Vehicle Status used in different Algorithms (X = used, O = not used).

Vehicle Status	No Forecast	Complete Forecast	Realistic Forecast	Minimizing Wait Time	Minimizing Dispatch Moves
Idle	X	X	X	X	X
Incoming(1)	X	X	X	X	X
Incoming(2)	X	X	X	O	X
FutIncoming(1)	O	X	O	X	X
FutIncoming(2)	O	X	O	O	X
FutOutgoing	O	X	X	X	X

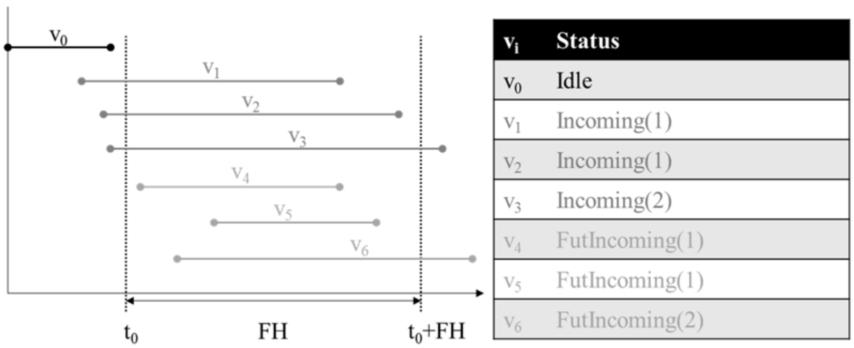

Figure 3: Time line visualization of different vehicle types.

Figure 4 shows the components and the sequence of the dispatch algorithm. The algorithm starts with the calculation of bay vehicle numbers using the different vehicle types as seen in Table 1. Depending on which dispatch scenario is used, vehicle type information for each bay will be included as a demand, surplus or not available for the certain scenario. The decision if a bay needs dispatching is made on three different levels. The first is to select these bays which have open transport requests (referred to as Outstanding), but not enough empty vehicles to serve these requests. The second level contains bays which have less available vehicles than the according reference value. The reference value is called Future Balance (FB) and works like a buffer to counteract uncertainty. For example, to describe future outgoing transports which can't be forecasted exactly and are only calculated by statistical averages (e.g. from storage transports) FB is dynamically updated by comparing bay waiting times as well as dispatch amount to system average. For complete forecast scenarios FB could be 0. Initial bay values for FB are calculated by using a genetic algorithm to reduce necessary simulation runs. This will be described in more detail in Section 4.

The third level dispatch decision considers the maximum vehicle amount per bay. If the total of incoming and idle vehicles reaches an upper control limit (UCL) the bay tries to push the necessary vehicle amount into another bay. The destination bay is found based on the push table which defines the push order. As an example the dispatch conditions (DC) for the three different level for the Minimizing Wait Time scenario (MWT) are calculated as follows:

DC1: Idle + Incoming(1) < Outstanding
DC2: Idle + Incoming(1) + FutIncoming(1) – Outstanding – FutOutgoing < FB
DC3: Idle + Incoming(1) – Outstanding > UCL

After introducing the idea of using forecast information for empty vehicle dispatching according to different optimization objectives as well as the main dispatcher steps the next section shows an evaluation of the described scenarios using simulation optimization.

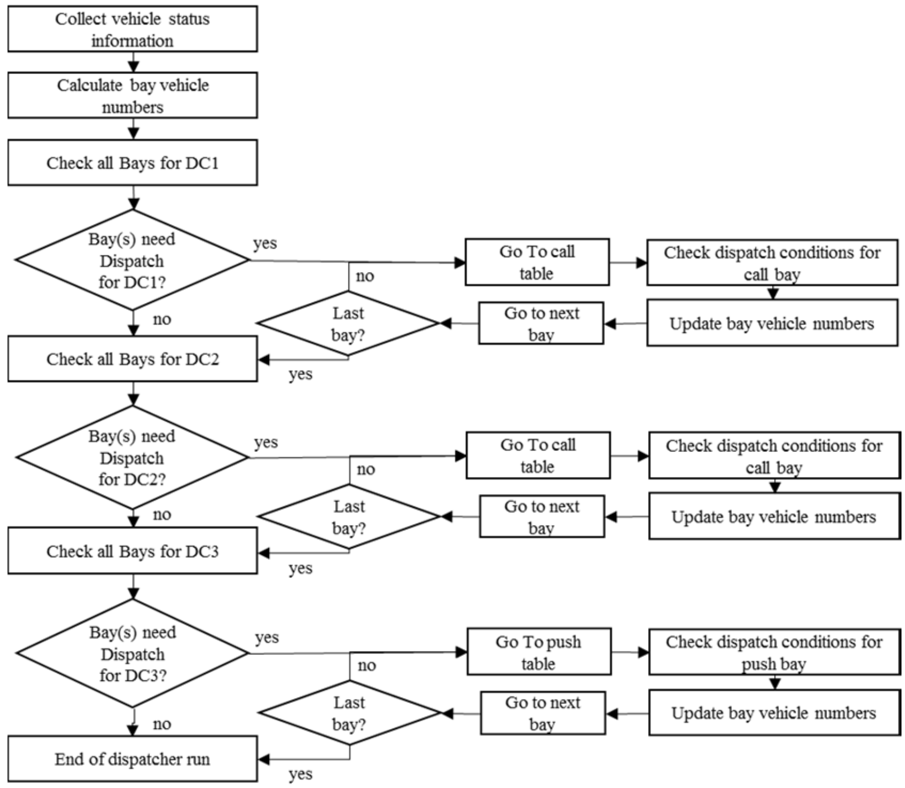

Figure 4: General Dispatcher steps.

5 SIMULATION EXPERIMENTS

In this section a simulation study is presented based on real fab transport data. The AMHS is a unified system consisting of 25 bays. The simulation model used is a java based event triggered one which explicitly represents empty vehicle balancing functionalities. Simulation of five different scenarios as described in Table 1 has been conducted to show differences between algorithms regarding information availability and optimization objective as well. The scenarios are as follows:

- No Forecast (NF) - similar to the watermark based and therefore the state of the art scenario
- Realistic Forecast (RF) - includes information about upcoming tool events
- Complete Forecast (CF) - includes all upcoming information for Forecast Horizon (FH)
- Minimizing Wait Time (MWT) - based on the CF scenario but does not include incoming type(2) transports because these would arrive after FH and thus reduce dispatch transports but increase wait time
- Minimizing Dispatch Transports (MDT) - the same as CF regarding the included information but with a different objective (minimizing dispatch transports more important than minimizing wait time)

Each simulation run used real fab transportation data of a 12 hour time range consisting of information about time and place of source and destination as well as the transport type (e.g. tool to tool or tool to storage). To find suitable initial values for the scenarios using forecast data, a genetic algorithm (GA) has been implemented (see e.g. Sivanandam and Deepa (2007) for more detailed information about GAs). It is important to find a suitable set of initial FB values for each scenario to be able to compare them. Otherwise it would be possible for a better algorithm to be outperformed by a worse one due to unfitting parameter sets. The genes in the GA represented the FB values for each bay as a decision variable x_i. The range of x_i depends on the scenario. The fitness which evaluates the performance of each simulation run depends on the scenario, too. While the performance of the MDT scenario was measured by the amount of dispatch transports the objective of the MWT scenario was to find the one with the lowest waiting times. Table 2 summarizes FB interval and fitness function for the different scenarios.

Table 2: Fitness function an FB interval for GA depending on scenario.

Scenario	RF	CF	MWT	MDT
Fitness function	Min WT + DT	Min WT + DT	Min WT	Min DT
FB Interval	0…8	0…2	0…15	0…8

The development of the average wait time, the amount of waiting and dispatch transports as well as the vehicle utilization is depicted in Figure 5. The GA stopped if there has not been any improvement in the "fittest individual" for three generations.

Figure 5: Development of performance parameters during optimization with GA to find suitable parameter sets (with normalized values).

Figure 6 shows the opposing trends of dispatch transports and wait time when optimizing one objective only (either minimize dispatch transports or average waiting time). The basic conclusion drawn from this is that reduced waiting times usually result in higher dispatch amounts and the other way around. The comparison of the performance parameters of the five previous described scenarios can be seen in Figure 7. The forecast scenarios optimizing a single objective achieve the best values for the corresponding parameter. MWT has minimal wait time and MDT has minimal amount of dispatch transports respectively. Nevertheless, the CF scenario trying to reduce both objectives reaches second best values for wait time and dispatch transports as well. Even the scenario using only information about upcoming tool events and thus only half of all possible future event information shows better results than the current state of the art water mark based method (NF).

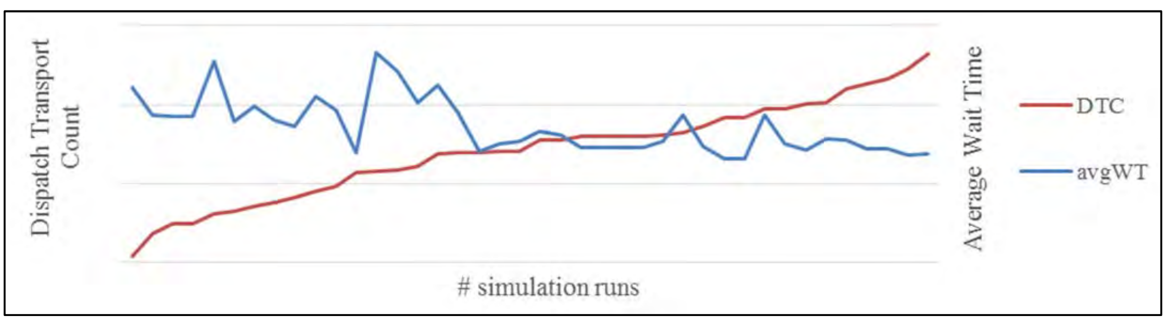

Figure 6: Development of Wait Time and Dispatch Transport Count while minimizing Wait Time.

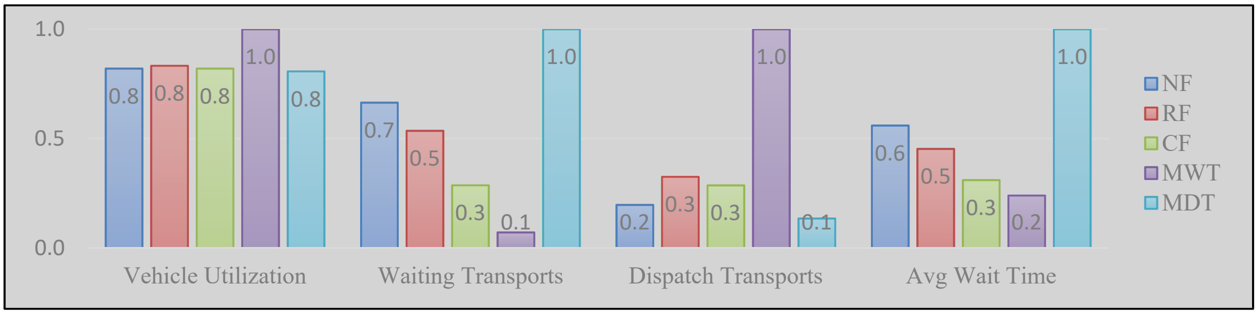

Figure 7: Dispatching scenarios with static values.

The last simulation experiment is aimed at introducing dynamically adjusting bay parameters (FB). The focus was on how dynamic values influence bay performance for the different scenarios. The control value to reduce or increase bay parameters for these experiments was the average bay waiting time and the relation to the average total waiting time over all bays. The results for the CF and RF scenarios can be seen in Figure 8. The experiments with the dynamic bay parameters outperform the ones with static values. The improvement from static to dynamic at the RF scenarios (RF_static vs RF_dyn) is bigger than at the CF ones. This can be explained with the different FB ranges. Since the FB for the CF scenario is between 0 and 2 there is not much room for dynamic adjustment. Dynamics have more room for improvement within the 0 to 8 range of the RF scenario.

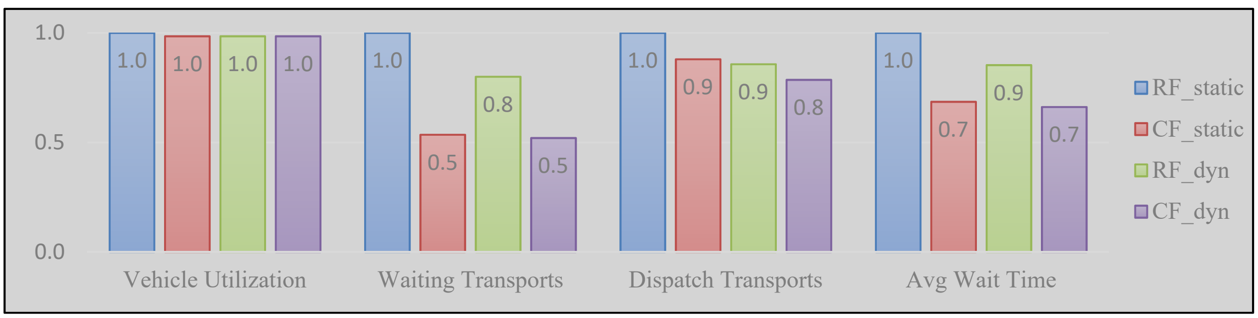

Figure 8: Dispatching scenarios with dynamic values.

6 CONCLUSION

The presented paper aimed at introducing the possibility of available forecast data into empty vehicle dispatching strategies for a unified AMHS of a semiconductor manufacturing facility. Different dispatching algorithms depending on the degree of information availability as well as the objective have been introduced. According to this distinction, four different algorithms (Realistic Forecast, Complete Forecast, Minimizing Waiting Time, Minimizing Dispatch Moves) have been compared to the current dispatch approach (static low/high water marks) in running fabs. Initial values for bay parameters serving as buffer values (Future Balance) to keep a minimum necessary amount of empty vehicles in each bay have been found by conducting simulation experiments using a genetic algorithm to find suitable values within an acceptable time. The results showed that reduced waiting times mostly result in higher dispatch transports and the other way around. It was also shown that the CF scenario with optimizing wait time and dispatch moves together is a good alternative to MDT and MWT strategies where only one objective is pursued. Finally, the RF could not achieve the same results as the CF strategy but still outperformed the state of the art scenario (NF).

In a second study the influence of dynamic bay parameters has been investigated and showed the improvements when introducing the availability of reacting to changing system states into empty vehicle dispatching scenarios.

Future studies should be dealing with the question which parameter to use for dynamic FB adjustment. Even a combination of average wait times and dispatch moves could be possible. Another research field is whether optimization techniques can be used to identify which source bay would be the best to serve a bay calling for empty vehicles. Of course the possibility of getting more detailed information about forecasting transport requests has to be investigated in more detail, too.

ACKNOWLEDGMENTS

The work was performed in the project "Verbesserte Layout- und Steuerungskonzepte von automatischen Transportsystemen in komplexen 28nm-Halbleiterfertigungen", co-funded by grants from the European Union (EFRE) and the Free State of Saxony (SAB).

REFERENCES

Chaabane, A. B., S. Dauzère-Pérès, C. Yugma, L. Rullière, and G. Lamiable. 2013. "Analyzing the Impact of Key Parameters of Vehicle Management Policies in a Unified AMHS". In *Proceedings of the 2013 Winter Simulation Conference*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, 3818–28. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chang, K.-H., Y.-H. Huang, and S.-P. Yang. 2014. "Vehicle Fleet Sizing for Automated Material Handling Systems to Minimize Cost Subject to Time Constraints". *IIE Transactions* 46 (3): 301–12.

Chen, W., Z. Wang, and F. TS Chan. 2017. "Robust Production Capacity Planning Under Uncertain Wafer Lots Transfer Probabilities for Semiconductor Automated Material Handling Systems". *European Journal of Operational Research*.

Driessen, J., R. Sjardijn, F. van Heukelom, C. van Roest, and M. Mom. 2016. "Wait-Time-Waste Improvement Opportunities and 'Smart Manufacturing' in Legacy 200mm Fabs". In *SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, 440–45. IEEE.

Hammel, C., R. Schmaler, T. Schmidt, J. Lübke, M. Schöps, U. Horn, and M. Mosinski. 2016. "Empowering Existing Automated Material Handling Systems to Rising Requirements: FA: Factory automation". In *SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, 87–93. IEEE.

Huang, C.-J., K.-H. Chang, and J. T. Lin. 2012. "Optimal Vehicle Allocation for an Automated Materials Handling System Using Simulation Optimization". *International Journal of Production Research* 50 (20): 5734–46.

Huang, C.-W., and J. T. Lin. 2016. "A Pre-Dispatching Vehicle Method for a Diffusion Area in a 300 mm Wafer Fab". *Journal of Industrial and Production Engineering* 33 (8): 579–91.

Jimenez, J. A., A. Grosser, C. Adikaram, V. Davila, R. Wright, and M. Bell. 2010. "AMHS Factors Enabling Small Wafer Lot Manufacturing in Semiconductor Wafer Fabs". In *Proceedings of the 2010 Winter Simulation Conference (WSC)*, edited by B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yücesan, 2575–85. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Kiba, J. E., S. Dauzère-Pérès, C. Yugma, S. Ga Charpak, and G. Lamiable. 2010. "Comparing Transport Policies in a Full-Scale 300mm Wafer Manufacturing Facility". In *Proceedings of the 11th International Material Handling Research Colloquium,* 2010.

Lin, J. T., F.-K. Wang, and C. K. Wu. 2003. "Connecting Transport AMHS in a Wafer Fab." *International Journal of Production Research* 41 (3): 529–44.

Lin, J. T., F. K. Wang, and J. R. Young. 2004. "Virtual Vehicle in the Connecting Transport Automated Material-Handling System (AMHS)". *International Journal of Production Research* 42 (13): 2599–2610.

Lin, J. T., and C.-J. Huang. 2013. "Simulation-Based Evolution Algorithm for Automated Material Handling System in a Semiconductor Fabrication Plant". In *Proceedings of International Asia Conference on Industrial Engineering and Management Innovation (IEMI2013)*, 1035–46. Springer.

Lin, J. T., C.-H. Wu, and C.-W. Huang. 2013. "Dynamic Vehicle Allocation Control for Automated Material Handling System in Semiconductor Manufacturing". *Computers & Operations Research* 40 (10): 2329–39.

Lin, J. T., and C.-J. Huang. 2014. "A Simulation-Based Optimization Approach for a Semiconductor Photobay with Automated Material Handling System". *Simulation Modelling Practice and Theory*, 76–100.

Rothe, J., G. Gaxiola, L. Marshall, T. Asakawa, K. Yamagata, and M. Yamamoto. 2015. "Novel Approaches to Optimizing Carrier Logistics in Semiconductor Manufacturing." *IEEE Transactions on Semiconductor Manufacturing* 28 (4): 494–501.

Schmaler, R., C. Hammel, T. Schmidt, and M. Schöps. 2016. "Dispatching Rules Considering Transport-Related Restrictions During Failure Scenarios—A Use Case: FA: Factory automation". In *SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, 94–99. IEEE.

Sivanandam, S. N., and S. N. Deepa. 2007. *Introduction to Genetic Algorithms*. Springer Science & Business Media.

Wang, L.-C., A. Wang, and C.-Y. Chueh. 2016. "Development of a Capacity Analysis and Planning Simulation Model for Semiconductor Fabrication". *The International Journal of Advanced Manufacturing Technology*, 1–16.

AUTHOR BIOGRAPHIES

ROBERT SCHMALER is a PH.D. student at Technische Universität Dresden and member of the scientific staff at the Chair of Logistics Engineering. He received his Diploma in Economics and Engineering from Technische Universität Dresden. His research interests include vehicle allocation and AMHS system optimization at semiconductor fabs. His email address is robert.schmaler@tu-dresden.de.

THORSTEN SCHMIDT heads the Chair of Logistics Engineering at Technische Universität Dresden. He holds degrees in Mechanical Engineering from Technische Universität Dortmund and Industrial Engineering from the Georgia Institute of Technology. His research interests focus on machinery and design of facility logistics and production systems. His email address is thorsten.schmidt@tu-dresden.de.

MATTHIAS SCHOEPS is a Member Technical Staff in Factory Automation. He received his Diploma in the field of communication technology of electrical engineering at the Dresden University of Applied Sciences and is in charge of AMHS system performance analysis as well as simulation of transport systems at GLOBALFOUNDRIES Fab1 and is responsible for coordination with Fab dispatching and control groups. His email address is matthias.schoeps@globalfoundries.com.

JOERG LUEBKE is Senior Member of Technical Staff with GLOBALFOUNDRIES in Dresden, Germany. He is responsible for the AMHS Control Systems of Fab1. He holds a degree as Diplomingenieur for Electrical Engineering with a major in Automation and Control from the Technische Universität Dresden. His email address is joerg.luebke@globalfoundries.com.

RALF HUPFER is Principal Engineer Factory Automation with GLOBALFOUNDRIES in Dresden, Germany. He is currently in charge of Material Control Systems (MCS) software components, holding a degree as Diplomingenieur for Electrical Engineering with a major in Automation and Control from the Technische Universität Dresden. His email address is ralf.hupfer@globalfoundries.com.

NIKOLAS SCHLAUS is an intern at GLOBALFOUNDRIES Fab1 and supports the Factory Automation staff especially in AMHS system analysis and simulation. His email address is nikolas.schlaus@globalfoundries.com.