Barhebwa-Mushamuka, F., Dauzère-Pérès, S., & Yugma, C. (2023). A global scheduling approach for cycle time control in complex manufacturing systems. International Journal of Production Research, 61(2), 559-579.

https://doi.org/10.1080/00207543.2021.2010828

A Global Scheduling Approach for Cycle Time Control in Complex Manufacturing Systems

Félicien Barhebwa-Mushamuka[1,2] Stéphane Dauzère-Pérès[2,3] Claude Yugma[2]

[1] IMT Atlantique, Departement of Automation, Production and Computer Sciences, Nantes, France
E-mail: felicien.barhebwa-mushamuka@imt-atlantique.fr, dauzere-peres@emse.fr, yugma@emse.fr

[2]Mines Saint-Etienne, Univ Clermont Auvergne
CNRS, UMR 6158 LIMOS
CMP, Department of Manufacturing Sciences and Logistics
Gardanne, France

[3]Department of Accounting and Operations Management
BI Norwegian Business School
Oslo, Norway

Abstract

This paper proposes a novel global scheduling approach for cycle time control strategy in large complex manufacturing systems with multiple workcenters, such as semiconductor manufacturing systems. The interaction between workcenters is taken into account by using global information such as release quantities, the Work-In-Process (WIP), cycle time targets and machine capacities. Local scheduling decisions in workcenters are steered by production targets, i.e., quantities of products to complete in each operation and each period on a scheduling horizon. These global production targets are determined by a mathematical model (global scheduling model), which optimizes the satisfaction of cycle time targets. One of the major innovations of the proposed model is that it relies on the temporal trace of the WIP. The mathematical model is coupled with a generic multi-method simulation model for evaluation purpose. Computational experiments conducted on industrial data show that our global scheduling approach efficiently controls the cycle times of products.

Keywords: Global scheduling, linear programming, cycle time control, semiconductor manufacturing

---

1. Introduction

Semiconductor manufacturing processes are probably the most complex manufacturing processes Mönch et al. (2012). In addition to common characteristics found in most manufacturing contexts, semiconductor manufacturing includes features that make production very complex such as re-entrant flows, hundreds of operations for each product leading to very long cycle times, hundreds of machines, different types of scheduling problems, etc. The manufacturing of Integrated Circuits includes two main processes: (1) The front end, or wafer fabrication, process which corresponds to the manufacturing of silicon chips on silicon wafers, and (2) The back end process which corresponds to the cutting and packaging of the chips and the final tests. The global scheduling problem studied in this paper typically pertains to front-end manufacturing. One of the most challenging problems in complex manufacturing systems is the consistency between the determination of high level goals/objectives and their operational implementation. To ensure this consistency, the use of policies and procedures, which aim to facilitate the operational control of activities in an organization, is a very common strategy in several industries (Lehmann (2016) and Simon and Schmidt (2015)). Production management decisions are generally grouped according to the time horizon on which they apply: Long term (strategic), medium term (tactical) and short term (operational). This decomposition allows the decision-making process to be simplified. The decisions taken at a higher level become constraints to satisfy or targets to meet at the lower levels. However, decisions at different levels are often made independently, and this can lead to inconsistent or unfeasible decisions (Dauzère-Pérès and Lasserre (2002)). In wafer fabrication, taking detailed scheduling decisions for the entire manufacturing facility (tens of thousands of operations on hundreds of machines) is so complex that the decision process is divided into two levels: A global level and a local level. The processing of products in the factory and to be released is simulated at the global level, on a relatively short-term horizon, to determine critical resources and to fix priorities of products at the various manufacturing stages. Then, resources or sets of resources are managed at the local level where the assignment of products to resources is determined as well as their production sequence. In this paper, the focus of the strategy implemented in the global scheduling approach is to control the cycle times of products by determining production targets for each product and each operation on a scheduling horizon. In this context, ensuring the consistency of decisions between both levels means that the production targets should be followed at the local level, with some degree of flexibility.

The paper is structured as follows. Section 2 introduces the problem under study and the scientific contributions. Section 3 reviews the literature on cycle time management in semiconductor manufacturing. Section 4 presents the global scheduling approach, and Section 5 how the approach is evaluated using a multi-method simulation model. Section 6 defines the global scheduling strategy to control cycle times, while Section 7 presents the global scheduling approach, and in particular the Linear Programming model used to model the global scheduling strategy. Computational results on industrial instances are discussed in Section 8. Finally, conclusions and perspectives are provided in Section 9.

2. Problem description and scientific contributions

In a wafer manufacturing facility, different products require hundreds of operations, with re-entrant flows, performed on hundreds of machines of different types that are grouped in workcenters. Each workcenter includes specific process characteristics, which increase the complexity of scheduling decisions such as batch process, parallel process, auxiliary resources, etc. Hence, determining detailed scheduling decisions for the entire facility is very difficult in semiconductor manufacturing. To cope with this complexity, commonly used approaches for scheduling decisions in workcenters are:

1. Real time scheduling using dispatching rules, i.e., every time a machine is available, a decision based on the selected rules is made to decide the next product to process. A review on dispatching rules can be found in Varadarajan and Sarin (2006).
2. Optimized scheduling algorithms dedicated to a workcenter, for instance scheduling on parallel machines with auxiliary resources in the photolithography workcenter (see Bitar et al. (2016)) or on batch machines in the diffusion workcenter (see Yugma et al. (2012), Jung et al. (2014) and Knopp et al. (2017)).

A general literature survey on scheduling in semiconductor manufacturing can be found in Mönch et al. (2011). The main downside of these approaches is that they are shortsighted. Independent scheduling decisions in each workcenter are bounded by the information available in the scope of the workcenter. Workcenters interact as products flow from one workcenter to others, but this interaction is not explicitly taken into account in the individual decisions of each workcenter. For instance, an upstream workcenter can send, in a short amount of time, large quantities of a given product to a downstream workcenter which has a limited number of machines that are qualified (also called eligible) to process this product. With a global view of the system, an unbalanced flow can be observed, which can deteriorate global key performance indicators, even though decisions taken locally in each workcenter are optimized.

This paper proposes a global scheduling approach whose goal is to steer scheduling decisions in workcenters by sending them production targets, i.e., quantities to be completed for each product, each operation and in each period of a scheduling horizon. These production targets are determined using different strategies depending on the objectives to optimize. The strategy explored in this paper aims at controlling cycle times, which is critical in semiconductor manufacturing times. Simple strategies are studied for instance in Barhebwa-Mushamuka et al. (2019a), Barhebwa-Mushamuka et al. (2019b) and Barhebwa-Mushamuka et al. (2021). To our knowledge, no paper has investigated the control of cycle times using an optimization model and the historical trace of the Work-In-Process (WIP). A common practice in the optimization models of the literature is to represent the WIP at each operation as a unique quantity without any additional information. The time this quantity has already been spent in the system and the time already spent in the operation are ignored in the optimization model. In this paper, we innovate by moving to a representation of the WIP of a product, not as a single quantity in an operation, but by multiple quantities depending on the time that the WIP of the product has already spent in the system. Hence, as detailed in Section 6.1, we use the historical trace of the WIP, which is essential in deciding which products are behind, on time, or ahead of their target cycle times. This enables our optimization model, presented in section 7, to accelerate the right quantities in the WIP of a product at an operation, and not the whole WIP as it is the case in the models of the literature.

3. Related literature

In this section, we review the literature related to the consistency of decisions in semiconductor manufacturing in Section 3.1, and the literature on cycle time management in semiconductor manufacturing in Section 3.2.

3.1. Consistency between decision levels in semiconductor manufacturing

Consistency of scheduling decisions in semiconductor manufacturing means ensuring that global objectives defined at factory level are followed at the local workcenter level. In the literature, hierarchical and iterative approaches are used to ensure the consistency between the tactical and operational decision levels. Priorities and production targets are used to ensure consistency at the operational decisions level between global scheduling decisions and local scheduling decisions.

- **Hierarchical approaches** (Hwang and Chang (2003), Liao et al. (1996), Tsakalis et al. (1997), El Adl et al. (1996), Vargas-Villamil et al. (2003)). In these approaches, information is exchanged only once. These approaches use an upper layer model (tactical level) which determines daily or weekly production targets and a lower layer model (operational level) which aims to reach these targets.

Targets are used as inputs to the lower layer model after being sliced into a very short detailed plan of three or six hours. Consistency is then ensured by additional constraints in the lower layer model in order to coordinate short-term actions to achieve the production objectives given by the higher level model.

- **Iterative approaches** (Hung and Leachman (1996), Kim et al. (2001),Bang and Kim (2010), Kim and Lee (2016)). In these approaches, the information is shared between the higher level model and the lower level model in each iteration. The iterative process is stopped when the plan provided by the higher level model is feasible in the lower level model. In some approaches, the decisions determined in the higher level model are the production quantities or the release quantities. These quantities are then evaluated in a lower level model, which is often a simulation model. In most cases, the higher level model is an Integer Linear Programming model or a Linear Programming model.

  Our approach differs from classical hierarchical and iterative approaches in the literature, which deal with the integration and communication between the tactical and operational decision levels. Higher level models are usually based on the future demand and on the resource capacity, while the approach proposed in this manuscript only deals with the operational level. Instead of the future demand as global information, our approach considers the lot release quantities, cycle time targets, resource capacity, Work-In-Process, etc.

- **Priority management approaches** (Bureau et al. (2007b), Bureau et al. (2007a), Vialletelle and France (2006), Sadeghi et al. (2016)) These approaches go beyond the communication framework between the tactical and operational decision levels. In these approaches, the operational decision level is structured into two levels: A global level relying on global information and taking decisions at the factory level, and a local level relying on local information and taking decisions at workcenter level. Lot priorities can be defined for various reasons such as the satisfaction of time constraints or customer priorities. Changing too often the priorities of lots in the WIP at the local level by speeding up late products or by slowing down early products is difficult to manage.

  The proposed approach in this paper enforces consistency in the operational decision level by switching from setting lot priorities to setting production targets. Production targets are the quantities of each product at each operation to complete in each period on a scheduling horizon. An adapted rule is required to ensure that these production targets are followed at the local level.

- **Production target management approaches** (Govind et al. (2008), Wu et al. (1998), Kao and Chang (2018)). A small number of studies in the literature use optimization methods to determine production targets at the operational level. In general, most studies that consider production targets at the operational level are essentially empirical or based on simple calculations.

  In this paper, we propose a complete global scheduling approach, i.e., the framework, principle, strategy, parameters and the evaluation environment. Our approach uses optimization models to determine production targets, which broaden the scope of the parameters to be used and offer the possibility to include several Work-In-Process management strategies in single-objective optimization models or in multi-objective optimization models. In this paper, a novel strategy to control cycle times is proposed.

3.2. Cycle time management in semiconductor manufacturing

In scheduling problems, most of the criteria are derived from the completion times of products, which constitute the main information to compute the cycle times of products, see e.g. Mati et al. (2011). The cycle time is one of the important key performance indicators in semiconductor manufacturing. The control of cycle times impacts several other metrics and key performance indicators such as throughput, yield and on time delivery. Controlling cycle times helps to reduce wafer risk contamination, yield loss and the inventory that should be maintained, see e.g. Lu et al. (1994). The cycle time includes the processing times, as well as the transport times and the time spent waiting in queue. Many factors influence the cycle times in semiconductor manufacturing, such as equipment availability, utilization, product mix, variability and hot products (high-priority products). In the semiconductor manufacturing literature, several studies focus on the understanding of cycle time and the way it can be improved. Bonal et al. (2001) provide a statistical method for cycle time management. The objective of the study was to ensure the quick detection of changes in the operation process that can affect the stability of the cycle time. In Sivakumar (2000), a discrete event simulation model for a semiconductor back-end manufacturing system is proposed to analyze the effect of controllable input parameters on the cycle time distribution and other output variables. In the same spirit, a simulation model is provided in Qi et al. (2002) to study the effect of some variables such as the job arrival distribution, the batch sizes, the downtime pattern and input control on the mean cycle time and average Work-In-Process. Chien et al. (2005) study how a learning curve approach can be used to determine empirical rules for cycle time improvement. Strategies based on the analysis of different problems related to the cycle time by using data from the manufacturing execution system are studied in Robinson and Chance (2000) and Ab Rahim et al. (2012). Kramer (1989) studies the improvement of cycle times with a focus on the breaking of the product cycle times into elements common to specific tools. The paper argues that the improvement of the cycle time of each element will lead to the improvement of the overall cycle time. For more studies on the understanding of the cycle time and the way it can be improved, see Nemoto et al. (2000) and Domaschke et al. (1998). The relationship between cycle times and other KPIs or parameters has also been investigated. A study based on the relationship between cycle time and yield in semiconductor wafer fabrication can be found in Wein (1992). Tirkel et al. (2009) investigate the relationship between cycle time and yield as affected by in-line metrology inspections of products. Leachman and Ding (2010) provide analytic formulas to quantify the revenue losses due to excursions not detected until the end-of-line testing as a function of the manufacturing cycle times, excursion probabilities and kill rates. The cycle time main challenges in semiconductor manufacturing are still based on how it can be predicted or estimated, controlled and reduced:

- Cycle time forecast and estimation are studied with the purpose to control and plan customer orders in tactical decisions, and further to manage some production factors such as the product releases and the WIP level in order to improve KPIs such as on-time delivery, throughput and yield. Different approaches are used for cycle time prediction and estimation: (1) Big data analytics (Wang and Zhang (2016) and Wang et al. (2020)), (2) Statistical methods, which include techniques such as probability distribution-based method and regression based method (Tai et al. (2012)), (3) Artificial intelligent techniques based on domain knowledge, machine learning and data mining (Tirkel (2011) and Hassoun (2013)), neural network (Chien et al. (2012)), and selective bayesian classifier based on a selection of minimal, most discriminative key-factor set for cycle time prediction (Meidan et al. (2011)), (4) Simulation for cycle time prediction (Chung and Huang (2002)), (5) Queueing model adapted for semiconductor manufacturing (Akhavan-Tabatabaei et al. (2009) and Shin et al. (2019)).

- Cycle time reduction refers to the strategy of decreasing the time a product spends in the factory from its release to its last operation. Shorter cycle times drive a better on time delivery, help to decrease Work-In-Process and ensure a better production quality (higher yield). Several strategies have been studied, essentially based on the management of factors that influence the cycle time. Variability is considered as one of the cycle time killers, see e.g. Robinson et al. (2002). Chen (2013) provides a three-step procedure for cycle time reduction: Identification of controllable factors that influence the product cycle time, investigation of the relationship between the controllable factors and the product cycle time and finally, based on this relationship, actions should be planned to shorten the product cycle time. In Hwang and Chang (2003), cycle time reduction is done by using a hierarchical approach based on two schedulers. A mid-term scheduler that maximizes the weighted production flow to ensure on time delivery, and a short-term schedule which slices the mid-term scheduling results into more detailed schedules. To reduce the cycle time, Kriett et al. (2017) address a planning problem, which determine how many lots have to be released during the next planning period and which target cycle times have to be assigned to each lot (including both the new releases and the initial WIP) such that both the cycle time and the deviation of the fab output from the master production schedule are minimized. Other factors that influence the cycle time have been used as a lever for cycle time reduction such as the batch sizes (Babbs and Gaskins (2007)), lot sizes (Zarifoglu et al. (2012) and Eberts et al. (2015)), Work-In-process management (Chien and Hu (2006)), queue time management and priority management. Equipment management, essentially the study of preventive maintenance segregation, is proposed in Rozen and Byrne (2016) with the goal to determine the optimum preventive maintenance policy that results in reduced fabrication cycle times. Leachman et al. (2002) provide a set of of methodologies and scheduling applications for managing the cycle time in semiconductor manufacturing called SLIM (Short cycle time and Low Inventory Manufacturing).

Due to the complexity of semiconductor manufacturing, some of the research in semiconductor manufacturing focus on the reduction of cycle times based on the activity of critical machines. This is the case in Swe et al. (2006) for cycle time reduction on cluster tools. Other works focus on a unique workcenter of the factory, such as Akcalt et al. (2001) for cycle time reduction in the photolitography area. In this paper, our global scheduling approach relies on a Linear Programming model to control cycle times by determining production targets. Production targets contribute to the smoothing of the WIP in the system and steer scheduling decisions at workcenter level. In addition, the control of cycle times is managed by controlling the competition of products on shared resources using these production targets. Cycle times are managed on the entire production line using blocks of operations (subdivision of product operations in sub-sequences of operations). In previous approaches of the literature, the release dates of products in the WIP were not considered. Our approach innovates by using both the release dates and the temporal trace of the WIP in the global scheduling approach. Temporally tracing the WIP is critical to differentiate quantities of the same product and at the same processing stage, but released at different times in the factory or that arrived at different times in the WIP of a workcenter.

4. Global scheduling approach

Front end semiconductor manufacturing is generally managed locally at the workcenter level with dispatching rules or dedicated scheduling algorithms. This local management has drawbacks such as a short-sighted view and may lead to an unbalanced Work-In-Process in the factory. To deal with this problem, global scheduling management is required. The global scheduling approach in this paper steers the local scheduling level using production targets, i.e. quantities of products to be completed for each operation and in each period on a scheduling horizon by using global information (Work-In-Process in the whole fab, lot releases, cycle time targets, resource capacity, etc.). Production targets are regularly updated on a rolling horizon to take into account the dynamics of the factory. The local scheduling level gets production targets as objectives to follow (using simple dispatching rules or dedicated scheduling algorithms) in term of quantities to produce and uses local information (waiting times of lots, processing times, current states of the machines, lots currently in queues, etc.). A key point in the global scheduling approach is the strategy which is considered to determine the production targets. The strategies, that depend on the criteria to optimize, are based on Work-In-Process management techniques and modeled using Linear Programming models written such as the ones in Barhebwa-Mushamuka et al. (2019a), Barhebwa-Mushamuka et al. (2019b) and Barhebwa-Mushamuka et al. (2021)

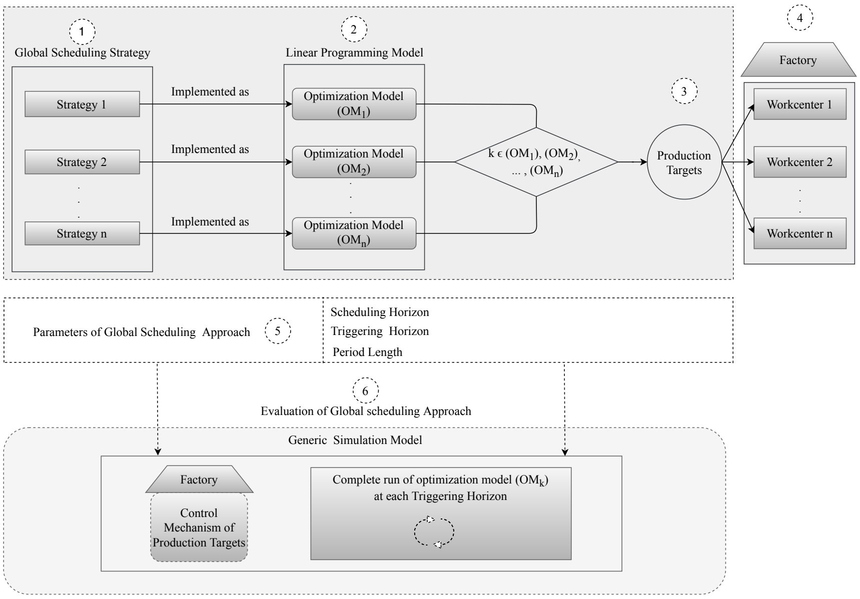

Figure 1: Framework of the global scheduling approach.

Figure 1 summarizes the global scheduling approach, which includes three main parts, and how it can be evaluated:

(1) **Global scheduling strategy.** A strategy is a Work-In-Process management policy which can be based on the operations of products, different resources, etc. with the objective to optimize a single or multiple objectives. In this paper, our objective is to control the cycle times of products.

(2) **Global scheduling model.** It implements the global scheduling strategy as a Linear Programming model, where the objective function is consistent with the strategy and the constraints bound the

actions of the strategy.

(3) **Production targets.** The production targets are outputs of the global scheduling model. By sending the quantities to be completed at the workcenter level, the global scheduling strategy is followed.

As for instance in Barhebwa-Mushamuka et al. (2019a), Barhebwa-Mushamuka et al. (2019b) and Barhebwa-Mushamuka et al. (2021), the global scheduling approach can be evaluated (6) using a simulation model which represents the local scheduling level (4). A control mechanism must be implemented in the simulation model to ensure that production targets are met at work center level. To evaluate the approach, various parameters are required such as the scheduling horizon, the length or duration of each period in the scheduling horizon and the horizon within which the global scheduling strategy is applied, called the triggering horizon (5).

In the global scheduling approach, a Linear Programming model is solved regularly on a rolling horizon. Thus, it is crucial to define the key parameters for the global scheduling approach (see Figure 2): (1) The duration of each period, a shift (8 hours) in our computational experiments, (2) The scheduling horizon, which is 99 periods (33 days) in our computational experiments and (3) The number of periods (called triggering horizon) before solving again the optimization model, 3 periods (1 day) in our computational experiments.

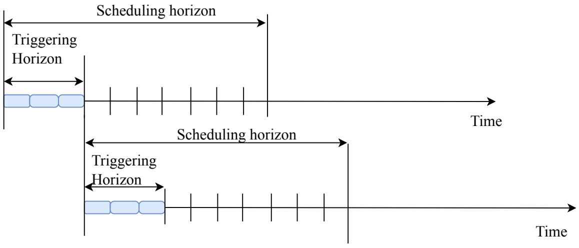

Figure 2: Scheduling horizon and triggering horizon in global scheduling approach.

The triggering horizon is important because the global scheduling model does not consider the detailed characteristics of the workcenters. Therefore, the model must be regularly solved to update the decisions by taking into account the events that occurred during the triggering horizon. The triggering horizon should not be too small to avoid changing decisions too often, or too long not to ignore some critical events. The scheduling horizon is important since it is used to predict the future behavior of the system affected by the scheduling decisions. A sufficiently long scheduling horizon helps the global scheduling model to mitigate the end of the horizon effects. As an illustration, consider a manufacturing system operating on 15 shifts per week (the plant is closed on Saturday and Sunday). Considering one hour as the period length, the scheduling horizon could be, for example, one week (15 shifts), i.e. the optimization model plans the production over one week of one-hour each period. The triggering horizon could be for example one day (3 shifts), i.e., only the scheduling decisions over the first three shifts are implemented, the remaining scheduling decisions will be rescheduled to take into account the evolution of the plant. The horizon between a call of the optimization model and its next call (the rescheduling point) is the triggering horizon. In our experiments, the triggering horizon and the scheduling horizon, but also the duration of each period, have been determined through intensive computational experiments that are not discussed in this paper. The choice of these global scheduling parameters mainly depends on the problem under study

5. Evaluation of the global scheduling approach

The multi-method simulation model used to evaluate the global scheduling approach is the one proposed in Sadeghi et al. (2016). It combines Discrete Event (DE) and Agent Based (AB) simulation methods and is coded with the AnyLogic simulation software (version 8.4). The notion of queues in Discrete-Event Simulation is used and the flexibility, behavior, and communication of agents in Agent Based simulation are used (Borshchev (2013)). The main types of agents are the products and the workcenters, while secondary agents are non-physical components such as operations and sequences of operations. The global scheduling optimization model is called on a rolling horizon by a simulation trigger event. After collecting dynamic parameters from the current status of the simulation model, such as the current WIP in workcenters, and static parameters such as future releases and aggregate resource capacities, the global scheduling optimization model is solved to determine production targets while the simulation model is paused. When the optimization is completed, the production targets determined by the global scheduling model are imposed as objectives at the workcenter level in terms of production quantities of each product to complete at each operation and in each period. Then, the simulation model resumes and tracks production quantities using a mechanism based on controller variables.

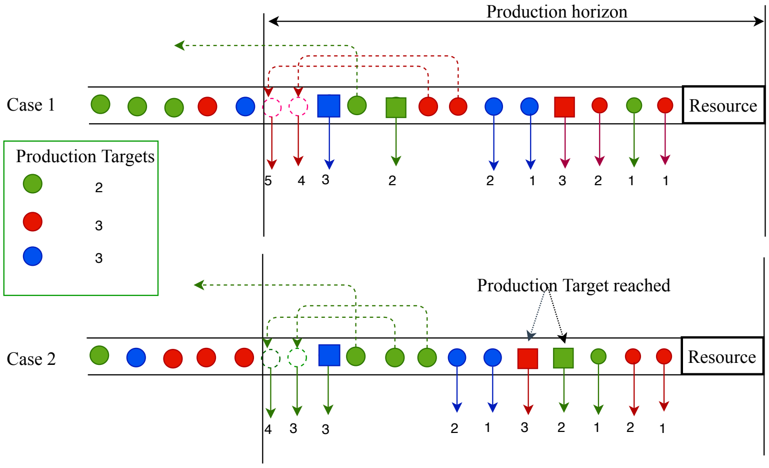

Figure 3: *Mechanism used in simulation model to track production targets.*

A controller variable is set up for each production target to indicate whether the production target is reached. If a product reaches its target in a period, then the controller variable ensure that the production of the product is temporally stopped to allow the products that have not reached their production targets to be processed. A product that has reached its production target can only be produced if all other products have reached their respective production targets or when the product is the only one in the queue of an available resource. Consider the example in Figure 3 with a processing time of one time unit on the resource, then the red product has an average cycle time of 5.8 time units with a throughput of 5 products in Case 1 (the first product is produced directly, the second product waits for two units of time, the third product waits for 3 units of time, the fourth product waits for 9 units of time and the fifth product waits for 10 units of time) while, in Case 2, the red product has an average cycle time of 2.6 time units with a throughput of 3 products (the first product is produced directly, the second product waits for 1 unit of time and the third product waits for 4 units of time). In the simulation model proposed in Sadeghi et al. (2016), we implement the input/output exchange/communication, which provides the communication strategy between our global scheduling optimization model for cycle time control that is integrated into a simulation model. The simulation model considers a granularity (process unit is one lot) that is different from the granularity of the global scheduling model which processes quantities. As the global scheduling optimization model is embedded in the simulation model, the interface of the exchange/communication is of great importance and is described in this section. This exchange/communication interface allows the simulation and optimization models to feed each other with input/output data. There are many studies in the literature on semiconductor manufacturing that combine simulation and optimization. However, it seems that no paper discusses the interface of the exchange/communication between optimization and simulation. A review of simulation optimization methods with application in semiconductor operational problems can be found in Ghasemi et al. (2018). Our simulation model has the following main components: (a) A meta-model, which describes the entities involved in the simulation model (products, lots, routes, machines, operations, etc.) and their relationships, (b) Parameters, that represent different adjustable characteristics of the simulation model (started lots per week, total number of lots to produce, warm-up time, priorities of products, etc.), and (c) A model structure, which specifies the interactions between entities. Our global scheduling model includes a Linear Programming model that materializes the global scheduling strategy, and parameters to evaluate the global scheduling strategy as described in Section 4. Therefore, the interface of exchange/communication, which allows the global scheduling model to communicate with the simulation model, includes the following elements:

1. **The coordination**, which represents the way both systems are synchronized. The procedure begins with the run of the simulation model until the triggering horizon is reached. Next, the global scheduling optimization model is called and, when the optimization is completed, the resulting production targets are sent as objectives at the workcenter level. Then, the simulation model resumes and tracks these production quantities.
2. **Interoperability**, which corresponds to the way the simulation model and the optimization model cooperate. The simulation model aims at satisfying the objectives sent by the global scheduling optimization model in terms of production targets. As already discussed, a mechanism based on controller variables is used. In addition, the interoperability ensures that the future product releases collected as static parameters are properly synchronized with the release scheme in the simulation model. Finally, the interoperability guarantees that, in the simulation model, the representation of the parameters of the global scheduling optimization model (scheduling horizon, triggering horizon and duration of each period in the scheduling horizon) are converted from periods to simulation time units.
3. **Input/output sharing data and aggregation**, which represents the way data (static and dynamic) are collected and aggregated. For example, machines are grouped in workcenters (each workcenter groups machines with the same capabilities), the capacities of the machines in the workcenter are summed to get the capacity of the workcenter, and the optimization model use quantities of lots in each operation instead of the individual lots in the operation that are considered in the optimization model.

6. Global scheduling strategy: Controlling cycle times

This section presents a strategy to control cycle times in front-end manufacturing using product cycle time targets and the temporal trace of the WIP. More precisely, the optimization model considers the number of periods products in the WIP have already been in the system. Product cycle time targets are given, and the strategy used in the global scheduling approach minimizes the tardiness on these cycle time targets. The cycle times are controlled throughout the entire production line. This is different for example from the work proposed in Bard et al. (2010), where there are no cycle time targets, but the daily target outputs are determined to minimize the sum of the deviations between the target outputs and the inventory of finished goods. Our strategy also differs from the one proposed in Hwang and Chang (2003), where the cycle time is not controlled but reduced. In Hwang and Chang (2003), the approach creates a communication tunnel between the tactical and operational levels by using production demand as input and not between different operational levels.

6.1. Temporal tracing of the Work-In-Process

A first originality of our approach is that the WIP is temporarily traced in the global scheduling model. As shown in Figure 4, instead of having just one parameter IW_{gl} for the initial WIP in operation l of product g, the parameter IW_{glr} is used for the initial WIP in operation l of product g released at period $-r$ in the past, where $IW_{gl} = \sum_{r=1}^{R} IW_{glr}$ and R is the number of release periods that must be considered in the past. The variables modeling the WIP of product g at operation l at the end of period p are also considering the release period. More precisely, Z^{P}_{glpr} corresponds to the quantity of product g at operation l released in period $-r$ in the past, and Z^{F}_{glpr} to the quantity of product g at operation l to be released in period r in the future. The optimization model that aims at satisfying given product cycle time targets is formalized in Section 7. A temporal tracing of the WIP is required to know which products in the WIP should actually be processed in a period. Indeed, if the information on the release periods of products is not available, it is impossible to know the time already spent in the system by products in the global scheduling model, and thus to ensure the satisfaction of cycle time targets.

When considering Figure 4, products in IW_{glr} should be processed before products in $IW_{glr'}$ if $r > r'$ since products in IW_{glr} have been released earlier. To our knowledge, no optimization models in the literature explicitly consider the temporal tracing of the WIP.

The approach for controlling cycle times starts by building blocks of operations, i.e. sub-sequences of operations of the products as in Bureau et al. (2007b). These blocks correspond to a logical separation of a route that allows intermediate controls in manufacturing. Each route of product g with $\mathscr{L}_g$ operations is divided into B_g blocks. Mathematically speaking, for a route of product g, blocks are formalized as follows as in Bureau et al. (2007b):

$block^1_g = \{OP^i_g | i = 1, ..., n_1\}$
$block^2_g = \{OP^i_g | i = n_1 + 1, ..., n_2\}$
...
$block^{B_g}_g = \{OP^i_g | i = n_{B_g - 1}, ..., L_g\}$

The i^{th} block of the route of product g has b_i operations spanning from operation $OP^{m_{i+1}}_g$ to operation $OP^{n_i}_g$ with $m < n$. The goal is then to control the cycle times of products by controlling the completion times of products in blocks, which are determined from their release dates. Hence, another important aspect of our approach is to establish a cycle time target for each block of operations (expressed as parameter T_{gl} for operation l of product g), which is derived from the cycle time target of the product.

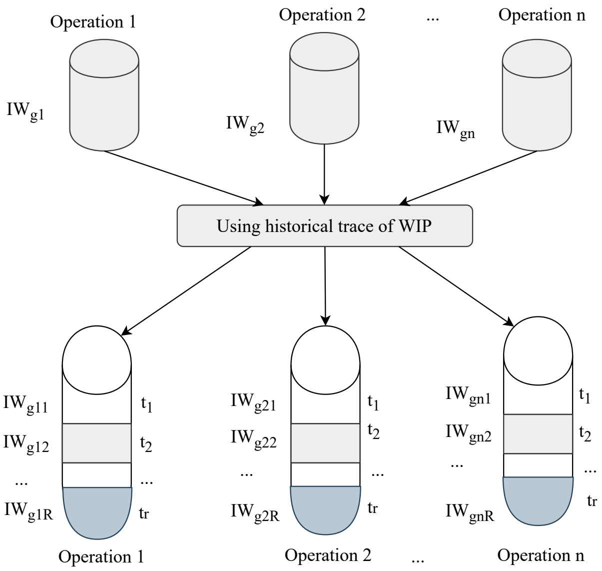

Figure 4: *Operations with historical trace of initial WIP.*

The objective of the global scheduling model is to prioritize products in a block that are behind their cycle time targets.

6.2. *Product cycle time targets and blocks of operations*

To control the cycle time of a product, its cycle time target is distributed over product operations. The number of operations is divided into blocks of operations in such a way that the last operation in the last block of operations should end at the cycle time target of the product. Blocks are defined using two different methods:

- A *naive method*, where the operations of a product is divided into blocks (sub-sequence of operations) with the same number of operations in each block, and the cycle time target is the same in each block. Figure 5 illustrates the process of defining blocks of operations. For a product with a cycle time target of 40 days and 5 blocks, blocks of operations are defined in such a way that operations in the first block should end at 20% of the cycle time target of the product, operations in the second block should end at 40% of the cycle time target of the product, etc.

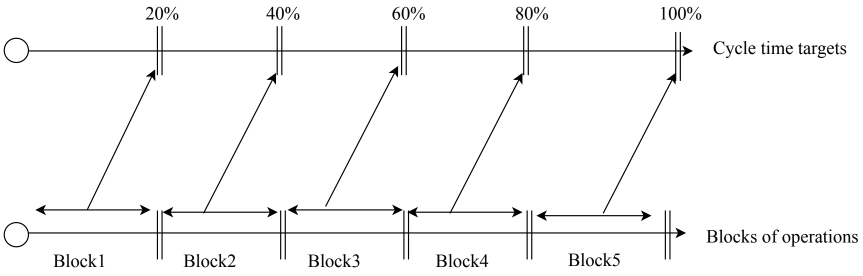

Figure 5: *Example of the definition of cycle time targets for blocks with the naive method.*

- A method based on simulation (*simulation-based method*), where the operations of product are divided into blocks (sub-sequence of operations) with the same number of operations, but with different cycle time targets. The time duration or the cycle time target of a block is determined based on the time each product spends in that block in the simulation. After the warm-up time, products are traced in the simulation, and the times they spent in each block are collected. These times provide the percentage of the cycle time target of a product in each block, and are used to determine the cycle time target of each block. Based on the example in Figure 6, for a product with a cycle time target of 40 days and 5 blocks of operations, if the information collected from the simulation indicates that the product spent on average 2 days, 14 days, 4 days, 16 days and 4 days in blocks 1, 2, 3, 4 and 5 respectively, then operations in the first block should end at 5% of the cycle time target of the product, operations in the second block should end at 40% of the cycle time target of the product, etc.

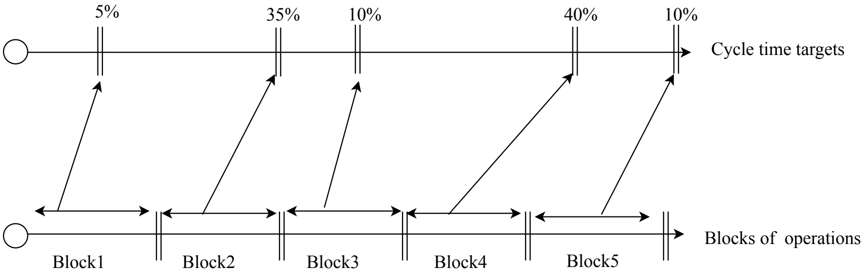

Figure 6: *Example of definition of cycle time targets for blocks with the simulation-based method.*

All cycle time targets of blocks are converted into number of periods and are used in the optimization model. Our computational experiments are conducted and compared with 12 blocks and 50 blocks of operations. Experiments with 12 blocks and 50 blocks were considered to show the sensitivity of the results to this parameter. Values between 12 and 50 were tested, but since computational times for 50 blocks are good and the results did not improve for more than 50 blocks, in this paper we present only the results for 12 blocks and 50 blocks.

6.3. Procedure for executing the strategy

As shown in Figure 7, the procedure starts with initial simulation (pre-simulation) to determine and set parameters such as the blocks, the block cycle time targets, the triggering horizon and the scheduling horizon. Then, the simulation process starts. As long as the simulation elapsed time since the beginning of the simulation (if the first call of the optimization model has not yet taken place) or since the last call of the optimization model is not equal to the triggering horizon, the simulation continues. Otherwise, the simulation is paused, static and dynamic data are collected to feed the optimization. The optimization model is launched automatically and at the end of the optimization, the production targets are returned to the simulation model. If it is not yet the end of the simulation horizon, the simulation resumes, otherwise the simulation is stopped.

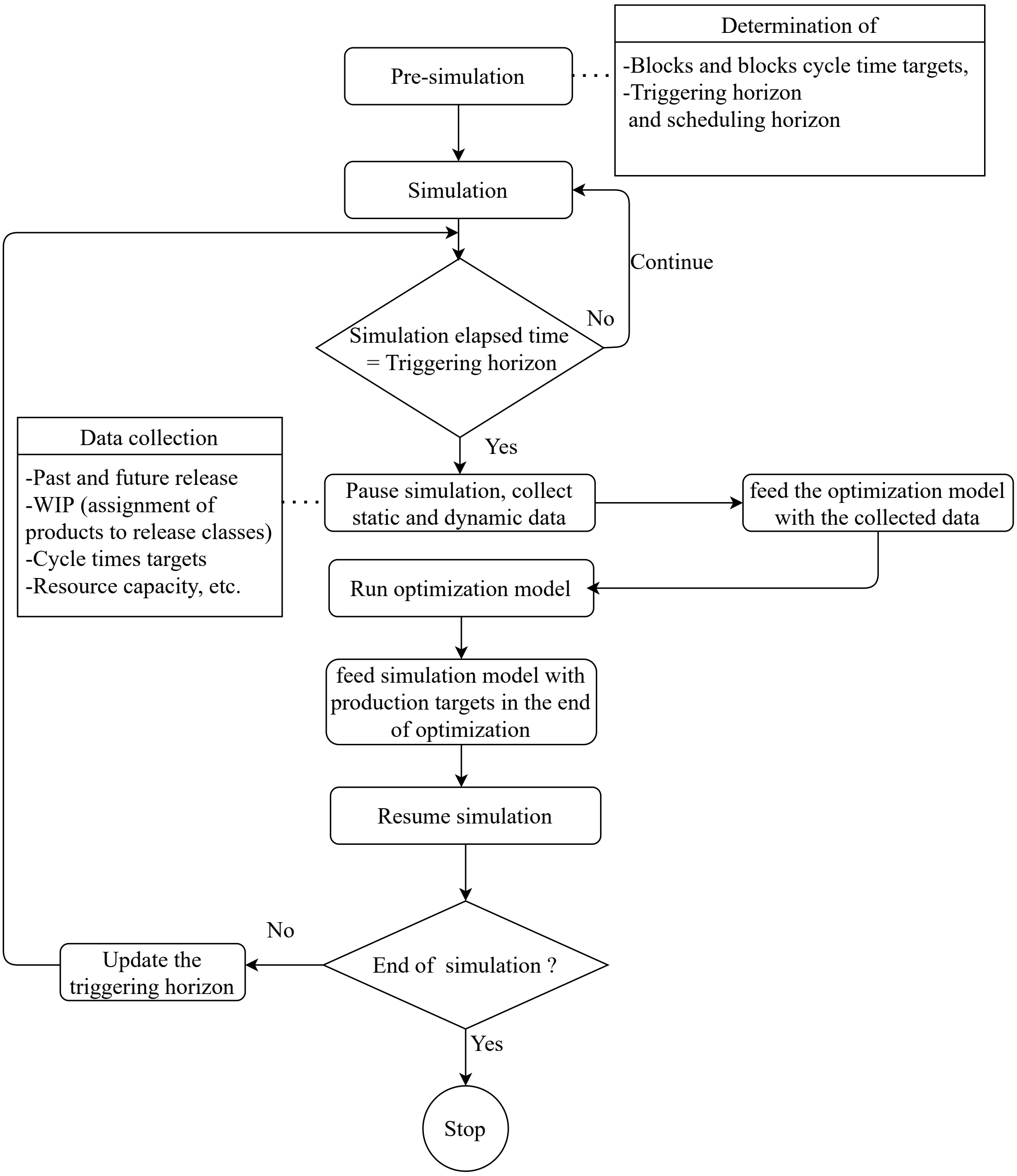

Figure 7: *Strategy execution procedure.*

7. Global scheduling model for cycle times control

This section describes the global scheduling model for controlling product cycle times through the Linear Programming (LP) model presented in Section 7.1. The approach is driven by the objective function of the LP model that considers two novel levers, the temporal trace of the WIP and the release periods of products. Section 7.2 shows how past release periods of products are aggregated into classes to improve the tractability of the LP model.

7.1. Global scheduling optimization model: Linear Program

The following parameters and decisions variables are necessary for the global scheduling optimization model.

Parameters:

T_{gl}: Cycle time target of operation l of product g, which is derived from the cycle time target of the block of the operation,
R: Number of periods considered in the past,
IW_{glr}: Initial WIP in operation l of product g released in period $-r$,
Q_{gp}: Release quantity of product g in period p,
α_{gl}: Unit processing time of product g at operation l,
C_{kp}: Capacity of workcenter k in period p,
$\mathscr{G}$: Set of all products, where each product has a sequence of operations required to complete the product,
$\mathscr{L}_g$: Set of operations of product $g \in \mathscr{G}$,
$\mathscr{K}$: Set of all workcenters,
$\mathscr{LK}(k)$: Set of operations of products that must be processed in workcenter k, i.e $(g,l) \in \mathscr{LK}(k)$ means that operation l of product g must be processed in workcenter k,
H: Number of periods in the planning horizon.

Decision variables:

X^P_{glpr}: Quantity of product g released at period $-r$ in the past, completing operation l in period p, where $r = 1,\dots,R$,
X^F_{glpr}: Quantity of product g to be released at period r, completing operation l in period p, where $r = 1,\dots,p$,
Y_{glp}: Total quantity of product g to complete in operation l in period p, i.e. the production target,
Z^P_{glpr}: WIP of product g at operation l at the end of period p released at period $-r$ in the past, where $r = 1,\dots,R$,
Z^F_{glpr}: WIP of product g at operation l at the end of period p to be released at period r, where $r = 1,\dots,p$.

The global scheduling optimization model is formalized below.

$$Min \sum_{g\in\mathscr{G}} \sum_{l\in\mathscr{L}_g} \sum_{p=1}^{H} \left(\sum_{r=1}^{R} \max(0, p+r-T_{gl}) Z^P_{glpr} + \sum_{r=1}^{p} \max(0, p-r-T_{gl}) Z^F_{glpr} \right) \tag{1}$$

Subject to :

$$Z^P_{g11r} = IW_{g1r} - X^P_{g11r} \quad \forall g \in \mathscr{G},\ r = 1,\dots,R \tag{2}$$

$$Z^F_{g111} = Q_{g1} - X^F_{g111} \quad \forall g \in \mathscr{G} \tag{3}$$

$$Z^P_{gl1r} = IW_{glr} + X^P_{g(l-1)1r} - X^P_{gl1r} \quad \forall g \in \mathscr{G},\ \forall l \in \mathscr{L}_g,\ l \geq 2,\ r = 1,\ldots,R \tag{4}$$

$$Z^F_{gl11} = X^F_{g(l-1)11} - X^F_{gl11} \quad \forall g \in \mathscr{G},\ \forall l \in \mathscr{L}_g,\ l \geq 2 \tag{5}$$

$$Z^P_{g1pr} = Z^P_{g1(p-1)r} - X^P_{g1pr} \quad \forall g \in \mathscr{G},\ p = 2,\ldots,\ H,\ r = 1,\ldots,R \tag{6}$$

$$Z^F_{g1pr} = Z^F_{g1(p-1)r} - X^F_{g1pr} \quad \forall g \in \mathscr{G},\ p = 2,\ldots,\ H,\ r = 1,\ldots,p-1 \tag{7}$$

$$Z^F_{g1pp} = Q_{gp} - X^F_{g1pp} \quad \forall g \in \mathscr{G},\ p = 2,\ldots,\ H \tag{8}$$

$$Z^P_{glpr} = Z^P_{gl(p-1)r} + X^P_{g(l-1)pr} - X^P_{glpr} \quad \forall g \in \mathscr{G},\ \forall l \in \mathscr{L}_g,\ l \geq 2,\ p = 2,\ldots,\ H,\ r = 1,\ldots,R \tag{9}$$

$$Z^F_{glpr} = Z^F_{gl(p-1)r} + X^F_{g(l-1)pr} - X^F_{glpr} \quad \forall g \in \mathscr{G},\ \forall l \in \mathscr{L}_g,\ l \geq 2,\ p = 2,\ldots,\ H,\ r = 1,\ldots,p-1 \tag{10}$$

$$Z^F_{glpp} = X^F_{g(l-1)pp} - X^F_{glpp} \quad \forall g \in \mathscr{G},\ \forall l \in \mathscr{L}_g,\ l \geq 2,\ p = 2,\ldots,\ H \tag{11}$$

$$\sum_{r=1}^{R} X^P_{glpr} + \sum_{r=1}^{p} X^F_{glpr} = Y_{glp} \quad \forall g \in \mathscr{G},\ \forall l \in \mathscr{L}_g,\ p = 1,\ldots,\ H \tag{12}$$

$$\sum_{g \in \mathscr{G}} \sum_{l \in \mathscr{L}_g;\ (g,l) \in \mathscr{L}\mathscr{K}(k)} \alpha_{gl} Y_{glp} \leq C_{kp} \quad \forall k \in K,\ p = 1,\ldots,\ H \tag{13}$$

$$Z^P_{glpr},\ X^P_{glpr} \geq 0 \quad \forall g \in \mathscr{G},\ \forall l \in L(g),\ p = 1,\ldots,\ H,\ r = 1,\ldots,R \tag{14}$$

$$Z^F_{glpr},\ X^F_{glpr} \geq 0 \quad \forall g \in \mathscr{G},\ \forall l \in L(g),\ p = 1,\ldots,\ H,\ r = 1,\ldots,H \tag{15}$$

$$Y_{glp} \geq 0 \quad \forall g \in \mathscr{G},\ \forall l \in L(g),\ p = 1,\ldots,\ H \tag{16}$$

The objective function (1) aims at satisfying the cycle time target of operations in blocks, by prioritizing the reduction of the WIP at operations with products that are late the most. The lateness is equal to $\max(0, p+r-T_{gl})$ for products released in past periods ($r = 1,\ldots,R$) and to $\max(0, p-r-T_{gl})$ for products released in future periods ($r = 1,\ldots,p$, for $p = 1,\ldots,H$). Hence, late products are pushed forward to their following operations. Constraints (2) and (3), resp. Constraints (4) and (5), determine the remaining WIP at the end of the first period in the first operation of each product, resp. in the following operations of each product. Constraints (2) and (4) correspond to the WIP for products released in past periods, while Constraints (3) and (5) correspond to the WIP for products released in the first period. Constraints (6), (7) and (8), resp. Constraints (9), (10) and (11), determine the remaining WIP in the first operation, resp. in each operation except the first one, of each product at the end of each period except the first one. Constraints (6) and (9) correspond to the WIP for products released in past periods, while Constraints (7), (8), (10) and (11) correspond to the WIP for products released in previous periods in the horizon except the first one. Constraints (12) ensure that Y_{glp}, the total quantity of product g that completes operation l in period p, is equal to the sum of the quantities of product g released in past periods and of the quantities of product g released in previous periods in the horizon. Constraints (13) model the resource capacity constraints. Finally, Constraints (14) through (16) are the non-negativity constraints.

Note that the initial WIP from products released in the past (considered in Constraints (2) and (4)), the products released before p (considered in Constraints (3) and (8)) or the WIP in previous periods (considered in Constraints (6), (7), (9) and (10)) are not considered in Constraints (5) and (11), since the products released in p are only entering through the first operation in Constraints (8).

7.2. Aggregating into classes of release periods

As shown in the previous section, past release periods of each product are considered in the global scheduling model to ensure the temporal tracing of the WIP and to control the product cycle times. However, because cycle times are very long in semiconductor manufacturing, up to 3 months, the number R of release periods considered in the past can be very large, leading to a very large number of variables X^P_{glpr} and Z^P_{glpr} in the linear programming model. This is why, instead of modeling each release period in the past and to make the model tractable, we aggregate past periods into N classes of A consecutive past periods, where $R = AN$.

N should be chosen not too large (too few past periods in each class), to avoid having a very large scheduling model, and not too small (too many past periods in each class), to avoid having in the same class products released at very different periods. There are several ways to build classes of past periods, where the extreme cases correspond either to using each period as a class ($N = R$) or to using a single class ($N = 1$). In the linear programming model, R is replaced by N, and the objective function (1) needs to be adjusted accordingly as follows:

$$Min \sum_{g\in\mathscr{G}} \sum_{l\in\mathscr{L}_g} \sum_{p=1}^{H} \left(\sum_{r=1}^{N} \max(0, p+(r)A-T_{gl}) Z^P_{glpr} + \sum_{r=1}^{p} \max(0, p-r-T_{gl}) Z^F_{glpr} \right) \tag{17}$$

In the computational experiments of Section 8, 12 ($N = 12$) classes of 7 days ($A = 7$) are used, i.e. 84 days, which is much larger than the average cycle time of most products in our experiments. Hence, products in the system and released in the past week are aggregated into the first release class, products in the system and released two weeks ago are aggregated into the second release class, etc.

8. Computational experiments

The design of the experiments is first detailed in Section 8.1. Then, the numerical results obtained without the Section global scheduling model are presented in 8.2. Section 8.3 shows the results obtained with the naive method to determine cycle time targets of blocks for instance 1 in Section 8.3.1, for instance 2 in Section 8.3.2 and when the satisfaction of the cycle time target of product 6 is emphasized in Section 8.3.3. The limits of the naive method is emphasized in Section 8.3. Finally, Section 8.4 presents the results obtained with the simulation-based method to determine cycle time targets of blocks for instance 1 in Section 8.4.1, for instance 2 in Section 8.4.2 and when cycle time targets of products are reduced in Section 8.4.3.

8.1. Design of the experiments

Industrial data were used for our experiments from a factory with about 600 machines distributed in about 300 workcenters. In our study, two instances with 5 products each are used:

1. The first instance includes products numbered 1 through 5, which have between 104 and 315 operations. One unit of product is released every 280 minutes, 360 minutes, 480 minutes, 480 minutes and 480 minutes for products 1, 2, 3, 4 and 5, respectively.
2. The second instance includes products numbered 6 through 10, which have between 153 and 221 operations. One unit of product is released every 460 minutes, 460 minutes, 480 minutes, 480 minutes and 480 minutes for products 6, 7, 8, 9 and 10, respectively.

The global scheduling model and the data driven generic multi-method simulation model developed in Sadeghi et al. (2016) were implemented using the Anylogic software (version 8.4) which interacts with the standard solver IBM ILOG CPLEX (version 12.6). Experiments were performed on a computer with windows 10 as operating system, processor Intel(R)Xeon(R) with 3.50 GHz and 32 Go of RAM. The release rates of the products in each instance have been defined using the simulation model without the global scheduling model. To avoid starting with an empty factory, six months of warm-up time are used, which are excluded when collecting statistical data. The simulation is then ran for 6 months after the warm-up time (when the steady-state is reached). The period duration in the global scheduling model corresponds to one shift (8 hours), and the global scheduling model is called in the simulation model every three periods (triggering horizon of 24 hours), i.e., more than 300 times in total. The scheduling horizon in the global scheduling model is fixed to 33 days ($H = 99$, i.e. 792 hours). The global scheduling model itself is a Linear Program, which is solved using CPlex. In our experiments, the resolution of each linear program never exceeds 45 seconds.

8.2. *Without global scheduling model*

This section presents the results obtained with the simulation model without the global scheduling model, i.e. where only First-In-First-Out dispatching rules are used. Table 1 shows the results for the first instance. For each product, the average cycle times, the release quantities and the completed quantities (throughput), both in number of products, are given. Note that the release quantities correspond to the number of products released after the warm-up period, and the completed quantities are equal to the number of products completed among the release quantities and that are used to compute the average cycle times. As expected, the average cycle time of a product usually decreases when the number of completed products increases.

The average cycle times of products 1, 2 and 3 are rather close, with a value of about 50 days, while products 4 and 5 are faster with a value of about 40 days.

	Products				
	1	**2**	**3**	**4**	**5**
Average Cycle Times (days)	**48.9**	**50.3**	**51.1**	**40.1**	**39.8**
Release Quantities	926	721	541	541	541
Completed Quantities	617	479	356	394	394

Table 1: Simulation without global scheduling approach, instance 1

Table 2 shows the results for the second instance. The average cycle times of Products 6 and 9 are the fastest, with a value of about 32 days, while Products 7, 8 and 10 are slower, in particular product 10 with a value of about 78 days.

In the experiments presented in Sections 8.3 and 8.4, the average cycle times in Tables 1 and 2 will be used as initial cycle time targets in the global scheduling model.

8.3. *Naive method to determine cycle time targets of blocks*

8.3.1. *Instance 1*

Table 3, resp. Table 4, shows the results for instance 1 obtained when the global scheduling approach is applied, i.e. when the simulation model is coupled with the global scheduling model, with 12 blocks of

	Products				
	6	**7**	**8**	**9**	**10**
Average Cycle Times (days)	**31.8**	**59.1**	**46.1**	**32.2**	**77.7**
Release Quantities	563	563	541	541	541
Completed Quantities	441	341	374	423	265

Table 2: Simulation without global scheduling approach, instance 2

	Products				
	1	**2**	**3**	**4**	**5**
Cycle Time Targets (days)	48.0	50.0	51.0	40.0	39.0
Average Cycle Times (days)	42.7	40.8	49.2	36.8	*43.1*
Cycle Time Gaps (%)	-11.0%	-18.4%	-3.5%	-8.0%	*10.5%*
Release Quantities	926	721	541	541	541
Completed Quantities	703	546	373	411	388

Table 3: Naive method, Simulation with global scheduling approach, instance 1 with 12 blocks of operations

operations, resp. 50 blocks of operations. The first row provides the cycle time targets, defined with the results obtained in Section 8.2, which are used to derive the cycle time targets of blocks with the naive method. The second row presents the average cycle times obtained with the global scheduling approach, and the third row the gaps between the cycle time targets and the average cycle times. The last two rows provide the release quantities and the completed quantities, both in number of products.

	Products				
	1	**2**	**3**	**4**	**5**
Cycle Time Targets (days)	48.0	50.0	51.0	40.0	39.0
Average Cycle Times (days)	39.3	43.7	39.8	*42.5*	*41.0*
Cycle Time Gaps (%)	-18.1%	-12.6%	-22.3%	*6.0%*	*5.1%*
Release Quantities	926	721	541	541	541
Completed Quantities	695	532	402	386	383

Table 4: Naive method, simulation with global scheduling approach, instance 1 with 50 blocks of operations

In Table 3, the average cycle times of products 1 to 4 are smaller than their cycle time targets, but the average cycle time of product 5 is about 4 days larger than its cycle time target. The results in Table 4 show that increasing the number of blocks to 50 helps to improve the results since, although two products have an average cycle time which is larger than their cycle time target, the largest difference is reduced to 2.5 days. However, the negative cycle time gaps of some products are very large, up to -22.3% for product 3 with 50 blocks of operations, which is not wanted when other products have positive cycle time gaps.

To reduce the average cycle time of product 5, its cycle time target has been reduced to 15 days and the numerical results with the global scheduling approach and 12 blocks of operations can be found in Table 5. Note that the cycle times targets of the other products have not been changed. Table 5 shows

	Products				
	1	**2**	**3**	**4**	**5**
Cycle Time Targets (days)	48.0	50.0	51.0	40.0	**15.0**
Average Cycle Times (days)	42.0	40.7	49.7	37.6	**39.5**
Cycle Time Gaps (%)	-12.5%	-18.6%	-2.5%	-6.0%	163.3%
Release Quantities	926	721	541	541	541
Completed Quantities	698	546	364	408	404

Table 5: Naive method, simulation with global scheduling approach, instance 1 with 12 blocks of operations and reduction of cycle time target of product 5

that the average cycle times of products 1 to 4 are still smaller than their cycle time targets, and that the average cycle time of product 5 is now very close to its cycle time target in Table 3 and smaller than its average cycle time in Table 1. However, the cycle time target of product 5 had to be drastically reduced to obtain this result.

8.3.2. Instance 2

	Products				
	6	**7**	**8**	**9**	**10**
Cycle Time Targets (days)	31.0	59.0	46.0	32.0	77.0
Average Cycle Times (days)	*44.0*	54.5	37.5	26.0	61.8
Cycle Time Gaps (%)	*41.9%*	-7.6%	-18.5%	-18.7%	-19.7%
Release Quantities	563	563	541	541	541
Completed Quantities	391	324	402	459	355

Table 6: Naive method, simulation with global scheduling approach, instance 2 with 12 blocks of operations

Table 6, resp. Table 7, shows the results for instance 2 obtained with the simulation model coupled with the global scheduling model and with 12 blocks of operations, resp. 50 blocks of operations. The average cycle times of products 7 to 10 are smaller than their cycle time targets, and significantly smaller for products 8, 9 and 10. However, the average cycle time of product 6 is much larger (13 days) than its cycle time target. As for instance 1, increasing the number of blocks to 50 helps to reduce the maximum cycle time gaps, as shown in Table 7 since, although two products have now an average cycle time which is larger than their cycle time target, the largest difference is reduced to 7.5 days, which is still quite large. Also as in for instance 1, the negative cycle time gaps of some products are very large, up to -42.9% for product 10 with 50 blocks of operations.

Because of the re-entrant flows and shared resources, trying to satisfy the cycle times of products 7 to 10 leads to a significant slowdown of product 6. Hence, and as in the previous section, the cycle time target of product 6 is decreased to 15 days, while the cycle time targets of the other products remain the same, and the global scheduling approach with 12 blocks of operations is applied again. The associated numerical results are given in Table 8. They show that the average cycle times of products 7 to 10 remain smaller than their cycle time targets, and that the average cycle time of product 6 has only been slightly reduced and remains much larger than its average cycle time in Table 1.

	Products				
	6	**7**	**8**	**9**	**10**
Cycle Time Targets (days)	31.0	59.0	46.0	32.0	77.0
Average Cycle Times (days)	*36.3*	*66.5*	32.4	23.2	44.1
Cycle Time Gaps (%)	*17.1%*	*12.7%*	-13.6%	-27.5%	-42.9%
Release Quantities	563	563	541	541	541
Completed Quantities	427	301	416	450	409

Table 7: Naive method, simulation with global scheduling approach, instance 2 with 50 blocks of operations

	Products				
	6	**7**	**8**	**9**	**10**
Cycle Time Targets (days)	**15.0**	59.0	46.0	32.0	77.0
Average Cycle Times (days)	**42.4**	55.2	37.9	26.7	62.3
Cycle Time Gaps (%)	182.6%	-6.4%	-17.6%	-16.6%	-19.1%
Release Quantities	563	563	541	541	541
Completed Quantities	395	336	400	455	337

Table 8: Naive method, simulation with global scheduling approach, instance 2 with with 12 blocks of operations and reduction of cycle time target of product 6

8.3.3. Reducing the cycle time of product 6

For the average cycle time of product 6 to reach the cycle time target of 31 days, the first step is to understand how the flows of other products impact the flow of product 6. The cycle time of product 6 can be decreased by slowing down other products, i.e. increasing their cycle times targets. Four scenarios are thus considered as shown in Table 9, where each of the four last products is alternatively slowed down by increasing its cycle time target to 150 days, and the cycle time target of product 6 is set to 31 days again.

		Products				
		6	**7**	**8**	**9**	**10**
Scenario 1	**Cycle time targets (days)**	31.0	**150.0**	46.0	32.0	61.0
	Average Cycle Times (days)	35.3	100.7	46.8	25.2	49.3
Scenario 2	**Cycle time targets (days)**	31.0	59.0	**150.0**	32.0	61.0
	Average cycle times (days)	42.2	52.5	56.2	26.2	56.0
Scenario 3	**Cycle time targets (days)**	31.0	59.0	46.0	**150.0**	61.0
	Average cycle times (days)	34.3	46.0	37.1	70.1	49.7
Scenario 4	**Cycle time targets (days)**	31.0	59.0	46.0	32.0	**150.0**
	Average cycle times (days)	39.5	47.7	35.2	26.4	109.5

Table 9: Impact of slowing down a single product on average cycle times of product 6

The results in Table 9 show how the flow of each product impacts the cycle time of product 6. Increasing the cycle time target of each product helps to reduce the average cycle time of product 6 from

	Products				
	6	**7**	**8**	**9**	**10**
Cycle Time Targets (days)	**31.0**	59.0	46.0	**250.0**	61.0
Average Cycle Times (days)	**26.5**	43.6	36.5	**86.9**	48.5
Cycle Time Gaps (%)	-14.5%	-26.1%	-20.6%	-65.2%	-20.5%
Release Quantities	563	563	541	541	541
Completed Quantities	434	386	406	279	387

Table 10: Naive method, simulation with global scheduling approach, slowing down product 9

its initial value or 44 days (see Table 6). However, the impact of products 7 and 9 is more significant and rather close. This is because products 6, 7 and 9 are often competing for the same machines. Table 10 shows how the average cycle time of product 6 can be further reduced by increasing even more the cycle time target of product 9 from 150 days to 250 days. The average cycle time of product 6 is now equal to 26.5 days, i.e. it is finally lower than the cycle time target of 31 days.

		Products				
		6	**7**	**8**	**9**	**10**
Scenario 5	**Cycle time targets (days)**	31.0	**90.0**	46.0	**90.0**	61.0
	Average cycle times (days)	34.5	59.0	45.5	50.5	52.1
Scenario 6	**Cycle time targets (days)**	31.0	59.0	46.0	**90.0**	**90.0**
	Average cycle times (days)	36.0	45.2	37.4	49.9	73.3
Scenario 7	**Cycle time targets (days)**	31.0	**90.0**	46.0	32.0	**90.0**
	Average cycle times (days)	37.6	59.7	41.3	27.1	73.5
Scenario 8	**Cycle time targets (days)**	31.0	**90.0**	46.0	**90.0**	**90.0**
	Average cycle times (days)	**30.4**	59.5	41.9	47.1	72.8

Table 11: Impact of slowing down multiple products on average cycle times of product 6

The issue with the results in Table 10 is that product 9 is significantly slowed down. An alternative is to slow down multiple products simultaneously and less drastically than in Tables 9 and 10. Four new scenarios are thus considered as shown in Table 11, where two or three products are slowed down by increasing their cycle time target to 90 days, instead of 150 days in Table 9 and 250 days in Table 10. More precisely, the cycle time targets of products 7 and 9 are increased in scenario 5, of products 9 and 10 in scenario 6, of products 7 and 10 in scenario 7 and of products 7, 9 and 10 in scenario 8. The cycle time target of product 6 remains equal to 31 days.

The results in Table 11 show that the average cycle time of product 6 is always significantly reduced from its initial value or 44 days (see Table 6) when the cycle time targets of two products are reduced. However, it is when the cycle time targets of three products are reduced (scenario 8) that the cycle time target of product 6 is finally satisfied.

8.4. Simulation-based method to determine cycle time targets of blocks

The analysis conducted in Section 8.3 shows that the naive method to determine cycle time targets of blocks is limited, and makes it difficult for the global scheduling approach to ensure that the cycle times of some products are satisfied. The results in this section show that the simulation-based method to determine cycle time targets of block helps to answer these limits. Let us recall that, using the simulation-based method, blocks include the same number of operations but have different cycle time targets.

8.4.1. Instance 1

	Products				
	1	**2**	**3**	**4**	**5**
Cycle Time Targets (days)	48.0	50.0	51.0	40.0	39.0
Average Cycle Times (days)	*48.3*	40.3	*51.6*	*40.1*	*42.5*
Cycle Time Gaps (%)	*0.6%*	-19.4%	*1.7%*	*0.3%*	*8.9%*
Release Quantities	926	721	541	541	541
Completed Quantities	675	477	371	386	386

Table 12: Simulation-based method, simulation with global scheduling approach, instance 1 with 12 blocks of operations

Table 12, resp. Table 13, shows the results for instance 1 obtained with the simulation model coupled with the global scheduling model with 12 blocks of operations, resp. 50 blocks of operations. The same cycle time targets for products than in Section 8.3 are used. In Table 12, most products have their average cycle times that are very close to their cycle time targets, except for product 2 with an average cycle time which is 19.4% lower and product 5 with an average cycle time which is 8.9% larger. Using 50 blocks of operations leads to very good results as shown in Table 13, where all products have an average cycle time which is lower than their cycle time target.

	Products				
	1	**2**	**3**	**4**	**5**
Cycle Time Targets (days)	48.0	50.0	51.0	40.0	39.0
Average Cycle Times (days)	40.7	43.5	47.3	40.0	35.7
Cycle Time Gaps (%)	-15.2%	-13.0%	-7.3%	0.0%	-8.5%
Release Quantities	926	721	541	541	541
Completed Quantities	705	524	377	389	405

Table 13: Simulation-based method, simulation with global scheduling approach, instance 1 with 50 blocks of operations

Through the use of optimized production targets and the simple controller variables used in the simulation model, cycle times are under control and are even all improved compared to the simulation model without the global scheduling approach.

8.4.2. Instance 2

Table 14, resp. Table 15, presents the results for instance 2 obtained with the global scheduling approach with 12 blocks of operations, resp. 50 blocks of operations. The results are worse than in instance 1 for 12 blocks of operations, with three products that have a positive cycle time gap and a maximum cycle time gap of 11.3%. Again, the improvements when using 50 blocks of operations are significant, as all products have an average cycle time which is lower than the corresponding cycle time target as shown in Table 15.

	Products				
	6	**7**	**8**	**9**	**10**
Cycle Time Targets (days)	31.0	59.0	46.0	32.0	77.0
Average Cycle Times (days)	*34.5*	*60.2*	37.9	*34.7*	73.7
Cycle Time Gaps (%)	*11.3%*	*2.0%*	-17.6%	*8.4%*	-4.3%
Release Quantities	563	563	541	541	541
Completed Quantities	405	314	412	396	306

Table 14: Simulation-based method, simulation with global scheduling approach, instance 2 with 12 blocks of operations

	Products				
	6	**7**	**8**	**9**	**10**
Cycle Time Targets (days)	31.0	59.0	46.0	32.0	77.0
Average Cycle Times (days)	25.0	57.3	40.0	30.1	77.0
Cycle Time Gaps (%)	-19.4%	-2.9%	-13.0%	-5.9%	0.0%
Release Quantities	563	563	541	541	541
Completed Quantities	471	337	387	425	281

Table 15: Simulation-based method, simulation with global scheduling approach, instance 2 with 50 blocks of operations

8.4.3. Reducing the cycle time targets of products

This section aims at illustrating that our global scheduling approach helps to control cycle times by reducing the cycle time targets of different products in instances 1 and 2. Due to the quality of the results obtained in Sections 8.4.1 and 8.4.2, 50 blocks of operations are considered in the remaining experiments.

	Products				
	1	**2**	**3**	**4**	**5**
Cycle Time Targets (days)	48.0	50.0	**42.0**	40.0	39.0
Average Cycle Times (days)	42.0	42.9	**43.3**	41.5	38.3
Cycle Time Gaps (%)	-7.9%	-18.2%	3.0%	3.7%	-1.8%
Release Quantities	926	721	541	541	541
Completed Quantities	676	524	383	391	403

Table 16: Simulation-based method, simulation with global scheduling approach, instance 1, reducing the cycle time target of product 3

First, the cycle time target of product 3 in instance 1, whose average cycle time is equal to 47.3 days in Table 13, is decreased from 51 days to 42 days. The results in Table 16 show that the average cycle time decreases from 47.3 to 43.3 days, only 3% above the target cycle time, and the average cycle times of the other products remain under control since the largest cycle time gap is equal to 3.7%.

The cycle time target of product 4 is decreased from 40 to 35 days in Table 17, and its average cycle time decreases from 40.0 days in Table 13 to 32.3 days, again with a limited impact on the satisfaction of other cycle time targets, since the largest cycle time gap is equal to 2.8% for product 5.

	Products				
	1	**2**	**3**	**4**	**5**
Cycle Time Targets (days)	48.0	50.0	51.0	**35.0**	39.0
Average Cycle Times (days)	41.0	43.8	50.6	**32.3**	40.1
Cycle Time Gaps (%)	-14.5%	-12.40%	-0.8%	-7.7%	2.8%
Release Quantities	926	721	541	541	541
Completed Quantities	679	526	368	416	390

Table 17: Simulation-based method, simulation with global scheduling approach, instance 1, reducing the cycle time target of product 4

	Products				
	6	**7**	**8**	**9**	**10**
Cycle Time Targets (days)	31.0	59.0	46.0	**27.0**	77.0
Average Cycle Times (days)	25.3	57.0	42.6	**26.8**	78.1
Cycle Time Gaps (%)	-18.4%	-3.4%	-7.4%	-0.7%	1.4%
Release Quantities	563	563	541	541	541
Completed Quantities	472	359	378	423	273

Table 18: Simulation-based method, simulation with global scheduling approach, instance 2, reducing the cycle time target of product 9

Considering now instance 2, the cycle time target of product 9 is decreased from 32 to 27 days in Table 18. The resulting average cycle time of product 9 decreases from 30.1 days in Table 15 to 26.8 days, with a very small maximum cycle time gap of 1.4% for product 10.

	Products				
	6	**7**	**8**	**9**	**10**
Cycle Time Targets (days)	31.0	59.0	46.0	32.0	**70.0**
Average Cycle Times (days)	27.1	61.3	46.2	31.6	**70.0**
Cycle Time Gaps (%)	-12.6%	3.9%	0.4%	-1.3%	0.0%
Release Quantities	563	563	541	541	541
Completed Quantities	466	304	360	421	290

Table 19: Simulation-based method, simulation with global scheduling approach, instance 2, reducing the cycle time target of product 10

In our last experiment, the cycle time target of product 10 is decreased to from 77 to 70 days. Table 19 shows that the average cycle time of product 10 exactly reaches its cycle time target and, as importantly, the other products remain under control with a cycle time gap always smaller than 3.9%.

9. Conclusions and perspectives

Controlling cycle times is very challenging in complex manufacturing systems such as semiconductor manufacturing. This paper proposes a global scheduling approach and a strategy that aims at meeting product cycle time targets. The approach uses global information at fab level and determines production targets to be followed by workcenters. These targets are production quantities to complete for each product at each operation and in each period on a scheduling horizon. By using the historical trace of the Work-In-Process and temporal tracing of the WIP, the global scheduling model minimizes the gap between the planned cycle times and the cycle time targets of products throughout their process. Two methods to determine cycle time targets in blocks of product routes are presented and compared. Numerical results on industrial data show that the global scheduling approach is effective in steering the manufacturing system to control the cycle times.

We hope the strategy proposed in this paper opens up a new way of explicitly controlling cycle times in complex manufacturing systems that other researchers will exploit. We are exploring various research directions. First, we are investigating how to combine the cycle time control strategy of this paper with other strategies, such as the ones in Barhebwa-Mushamuka et al. (2019a), in a multi-objective approach. We are also working on the development of other solution approaches to solve more complex models with additional constraints and a larger number of products.

Acknowledgments

This project has received funding from the Electronic Component Systems for European Leadership Joint Undertaking under grant agreement No 737459 (project Productive4.0). This Joint Undertaking receives support from the European Unions Horizon 2020 research and innovation program and Germany, Austria, France, Czech Republic, Netherlands, Belgium, Spain, Greece, Sweden, Italy, Ireland, Poland, Hungary, Portugal, Denmark, Finland, Luxembourg, Norway, Turkey.

Data availability

The data used in this article is industrial data currently used in inc's manufacturing system. Due to the confidentiality of this data, it will not be shared publicly.

References

Ab Rahim, S.R., Ahmad, I., Chik, M.A., 2012. Technique to improve visibility for cycle time improvement in semiconductor manufacturing, in: 2012 10th IEEE International Conference on Semiconductor Electronics (ICSE), IEEE. pp. 627–630.

Akcalt, E., Nemoto, K., Uzsoy, R., 2001. Cycle-time improvements for photolithography process in semiconductor manufacturing. IEEE Transactions on Semiconductor Manufacturing 14, 48–56.

Akhavan-Tabatabaei, R., Ding, S., Shanthikumar, J.G., 2009. A method for cycle time estimation of semiconductor manufacturing toolsets with correlations, in: Winter Simulation Conference, pp. 1719–1729.

Babbs, D., Gaskins, R., 2007. Effectiveness of small batch size on cycle time reduction in a conventional 300mm factory, in: 2007 IEEE/SEMI Advanced Semiconductor Manufacturing Conference, IEEE. pp. 105–110.

Bang, J.Y., Kim, Y.D., 2010. Hierarchical production planning for semiconductor wafer fabrication based on linear programming and discrete-event simulation. IEEE Transactions on Automation Science and Engineering 7, 326–336.

Bard, J.F., Deng, Y., Chacon, R., Stuber, J., 2010. Midterm planning to minimize deviations from daily target outputs in semiconductor manufacturing. IEEE transactions on semiconductor manufacturing 23, 456–467.

Barhebwa-Mushamuka, F., Dauzère-Pérès, S., Yugma, C., 2019a. Multi-objective optimization for work-in-process balancing and throughput maximization in global fab scheduling, in: 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), IEEE. pp. 697–702.

Barhebwa-Mushamuka, F., Dauzère-Pèrès, S., Yugma, C., 2019b. Work-in-process balancing control in global fab scheduling for semiconductor manufacturing, in: 2019 Winter Simulation Conference (WSC), IEEE. pp. 2257–2268.

Barhebwa-Mushamuka, F., Dauzère-Pérès, S., Yugma, C., 2021. Push and time at operation strategies for cycle time minimization in global fab scheduling for semiconductor manufacturing, in: 2021 IEEE 17th International Conference on Automation Science and Engineering (CASE), IEEE. pp. 1309–1314.

Bitar, A., Dauzère-Pérès, S., Yugma, C., Roussel, R., 2016. A memetic algorithm to solve an unrelated parallel machine scheduling problem with auxiliary resources in semiconductor manufacturing. Journal of Scheduling 19, 367–376.

Bonal, J., Fernandez, M., Maire-Richard, O., Aparicio, S., Oliva, R., Gonzalez, S.G.B., Rodriguez, L., Rosendo, M., Villacieros, J., Becerro, J., 2001. A statistical approach to cycle time management, in: 2001 IEEE/SEMI Advanced Semiconductor Manufacturing Conference, IEEE. pp. 11–15.

Borshchev, A., 2013. The big book of simulation modeling: multimethod modeling with AnyLogic 6. AnyLogic North America.

Bureau, M., Dauzère-Pérès, S., Yugma, C., Vermarien, L., 2007a. An approach for simulating consistent global and local scheduling, in: 2007 IEEE/SEMI Advanced Semiconductor Manufacturing Conference, IEEE. pp. 96–99.

Bureau, M., Dauzère-Pérès, S., Yugma, C., Vermariën, L., Maria, J.B., 2007b. Simulation results and formalism for global-local scheduling in semiconductor manufacturing facilities, in: Henderson, S.G., Biller, B., Hsieh, M.H., Shortle, J., Tew, J.D., Barton, R.R. (Eds.), Proceedings of the 2007 Winter Simulation Conference, Institute of Electrical and Electronics Engineers, Inc., Washington, DC. pp. 1768–1773.

Chen, T., 2013. A systematic cycle time reduction procedure for enhancing the competitiveness and sustainability of a semiconductor manufacturer. Sustainability 5, 4637–4652.

Chien, C.C., Meng, C., Chen, K.L., Hung, K.T., 2005. Cycle time learning curve in semiconductor foundry industry, in: ISSM 2005, IEEE International Symposium on Semiconductor Manufacturing, 2005., IEEE. pp. 359–360.

Chien, C.F., Hsu, C.Y., Hsiao, C.W., 2012. Manufacturing intelligence to forecast and reduce semiconductor cycle time. Journal of Intelligent Manufacturing 23, 2281–2294.

Chien, C.F., Hu, C.H., 2006. Segmented wip control for cycle time reduction, in: 2006 IEEE International Symposium on Semiconductor Manufacturing, IEEE. pp. 265–268.

Chung, S.H., Huang, H.W., 2002. Cycle time estimation for wafer fab with engineering lots. Iie Transactions 34, 105–118.

Dauzère-Pérès, S., Lasserre, J.B., 2002. On the importance of sequencing decisions in production planning and scheduling. International transactions in operational research 9, 779–793.

Domaschke, J., Brown, S., Robinson, J., Leibl, F., 1998. Effective implementation of cycle time reduction strategies for semiconductor back-end manufacturing, in: 1998 Winter Simulation Conference. Proceedings, IEEE. pp. 985–992.

Eberts, D., Keil, S., Peipp, F., Lasch, R., 2015. Shortening of cycle time in semiconductor manufacturing via meaningful lot sizes, in: 2015 26th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), IEEE. pp. 34–41.

El Adl, M., Rodriguez, A.A., Tsakalis, K.S., 1996. Hierarchical modeling and control of re-entrant semiconductor manufacturing facilities, in: Decision and Control, 1996., Proceedings of the 35th IEEE Conference on, IEEE. pp. 1736–1742.

Ghasemi, A., Heavey, C., Laipple, G., 2018. A review of simulation-optimization methods with applications to semiconductor operational problems, in: 2018 Winter Simulation Conference (WSC), IEEE. pp. 3672–3683.

Govind, N., Bullock, E.W., He, L., Iyer, B., Krishna, M., Lockwood, C.S., 2008. Operations management in automated semiconductor manufacturing with integrated targeting, near real-time scheduling, and dispatching. IEEE Transactions on Semiconductor Manufacturing 21, 363–370.

Hassoun, M., 2013. On improving the predictability of cycle time in an nvm fab by correct segmentation of the process. IEEE Transactions on Semiconductor Manufacturing 26, 613–618.

Hung, Y.F., Leachman, R.C., 1996. A production planning methodology for semiconductor manufacturing based on iterative simulation and linear programming calculations. IEEE Transactions on Semiconductor manufacturing 9, 257–269.

Hwang, T.K., Chang, S.C., 2003. Design of a lagrangian relaxation-based hierarchical production scheduling environment for semiconductor wafer fabrication. IEEE Transactions on Robotics and Automation 19, 566–578.

Jung, C., Pabst, D., Ham, M., Stehli, M., Rothe, M., 2014. An effective problem decomposition method for scheduling of diffusion processes based on mixed integer linear programming. IEEE Transactions on Semiconductor Manufacturing 27, 357–363.

Kao, Y.T., Chang, S.C., 2018. Setting daily production targets with novel approximation of target tracking operations for semiconductor manufacturing. Journal of Manufacturing Systems 49, 107–120.

Kim, S.H., Lee, Y.H., 2016. Synchronized production planning and scheduling in semiconductor fabrication. Computers & Industrial Engineering 96, 72–85.

Kim, Y.D., Kim, J.G., Choi, B., Kim, H.U., 2001. Production scheduling in a semiconductor wafer fabrication facility producing multiple product types with distinct due dates. IEEE Transactions on Robotics and Automation 17, 589–598.

Knopp, S., Dauzère-Pérès, S., Yugma, C., 2017. A batch-oblivious approach for complex job-shop scheduling problems. European Journal of Operational Research 263, 50–61.

Kramer, S.S., 1989. Total cycle time management by operational elements, in: IEEE/SEMI International Semiconductor Manufacturing Science Symposium, IEEE. pp. 17–20.

Kriett, P.O., Eirich, S., Grunow, M., 2017. Cycle time-oriented mid-term production planning for semiconductor wafer fabrication. International Journal of Production Research 55, 4662–4679.

Leachman, R.C., Ding, S., 2010. Excursion yield loss and cycle time reduction in semiconductor manufacturing. IEEE Transactions on Automation science and engineering 8, 112–117.

Leachman, R.C., Kang, J., Lin, V., 2002. Slim: Short cycle time and low inventory in manufacturing at samsung electronics. Interfaces 32, 61–77.

Lehmann, C.F., 2016. Strategy and business process management: Techniques for improving execution, adaptability, and consistency. Auerbach Publications.

Liao, D.Y., Chang, S.C., Pei, K.W., Chang, C.M., 1996. Daily scheduling for r&d semiconductor fabrication. IEEE transactions on Semiconductor Manufacturing 9, 550–561.

Lu, S.C., Ramaswamy, D., Kumar, P., 1994. Efficient scheduling policies to reduce mean and variance of cycle-time in semiconductor manufacturing plants. IEEE Transactions on Semiconductor Manufacturing 7, 374–388.

Mati, Y., Dauzère-Pérès, S., Lahlou, C., 2011. A general approach for optimizing regular criteria in the job-shop scheduling problem. European Journal of Operational Research 212, 33–42.

Meidan, Y., Lerner, B., Rabinowitz, G., Hassoun, M., 2011. Cycle-time key factor identification and prediction in semiconductor manufacturing using machine learning and data mining. IEEE Transactions on Semiconductor Manufacturing 24, 237–248.

Mönch, L., Fowler, J.W., Dauzère-Pérès, S., Mason, S.J., Rose, O., 2011. A survey of problems, solution techniques, and future challenges in scheduling semiconductor manufacturing operations. Journal of scheduling 14, 583–599.

Mönch, L., Fowler, J.W., Mason, S.J., 2012. Production planning and control for semiconductor wafer fabrication facilities: modeling, analysis, and systems. volume 52. Springer Science & Business Media.

Nemoto, K., Akcali, E., Uzsoy, R.M., 2000. Quantifying the benefits of cycle time reduction in semiconductor wafer fabrication. IEEE Transactions on Electronics Packaging Manufacturing 23, 39–47.

Qi, C., Tang, T.K., Sivakumar, A.L., 2002. Modeling methodology: simulation based cause and effect analysis of cycle time and wip in semiconductor wafer fabrication, in: Proceedings of the 34th conference on Winter simulation: exploring new frontiers, Winter Simulation Conference. pp. 1423–1430.

Robinson, J., Chance, F., 2000. Wafer fab cycle time management using mes data, in: Proceedings of the 2000 Modeling and Analysis for Semiconductor Manufacturing Conference (MASM 2000), Tempe, AZ.

Robinson, J.K., et al., 2002. Understanding and improving wafer fab cycle times. Semiconductor FabTech 17.

Rozen, K., Byrne, N.M., 2016. Using simulation to improve semiconductor factory cycle time by segregation of preventive maintenance activities, in: Proceedings of the 2016 Winter Simulation Conference, IEEE Press. pp. 2676–2684.

Sadeghi, R., Dauzere-Pérès, S., Yugma, C., 2016. A multi-method simulation modelling for semiconductor manufacturing. IFAC-PapersOnLine 49, 727–732.

Shin, J., Grosbard, D., Morrison, J.R., Kalir, A., 2019. Decomposition without aggregation for performance approximation in queueing network models of semiconductor manufacturing. International Journal of Production Research 57, 7032–7045.

Simon, D., Schmidt, C., 2015. Business architecture management: Architecting the business for consistency and alignment. Springer.

Sivakumar, A.I., 2000. Simulation based cause and effect analysis of cycle time distribution in semiconductor backend, in: Proceedings of the 32nd conference on Winter simulation, Society for Computer Simulation International. pp. 1464–1471.

Swe, A.N., Gupta, A.K., Sivakumar, A.I., Lendermann, P., 2006. Cycle time reduction at cluster tool in semiconductor wafer fabrication, in: 2006 8th Electronics Packaging Technology Conference, IEEE. pp. 671–677.

Tai, Y., Pearn, W., Lee, J., 2012. Cycle time estimation for semiconductor final testing processes with weibull-distributed waiting time. International Journal of Production Research 50, 581–592.

Tirkel, I., 2011. Cycle time prediction in wafer fabrication line by applying data mining methods, in: 2011 IEEE/SEMI Advanced Semiconductor Manufacturing Conference, IEEE. pp. 1–5.

Tirkel, I., Reshef, N., Rabinowitz, G., 2009. In-line inspection impact on cycle time and yield. IEEE Transactions on Semiconductor Manufacturing 22, 491–498.

Tsakalis, K.S., Flores-Godoy, J.J., Rodriguez, A.A., 1997. Hierarchical modeling and control for reentrant semiconductor fabrication lines: a mini-fab benchmark, in: Emerging Technologies and Factory Automation Proceedings, 1997. ETFA'97., 1997 6th International Conference on, IEEE. pp. 508–513.

Varadarajan, A., Sarin, S.C., 2006. A survey of dispatching rules for operational control in wafer fabrication. IFAC Proceedings Volumes 39, 715–726.

Vargas-Villamil, F.D., Rivera, D.E., Kempf, K.G., 2003. A hierarchical approach to production control of reentrant semiconductor manufacturing lines. IEEE Transactions on control systems technology 11, 578–587.

Vialletelle, P., France, G., 2006. An overview of an original wip management framework at a high volume/high mix facility. IFAC Proceedings Volumes 39, 89–92.

Wang, J., Zhang, J., 2016. Big data analytics for forecasting cycle time in semiconductor wafer fabrication system. International Journal of Production Research 54, 7231–7244.

Wang, J., Zheng, P., Zhang, J., 2020. Big data analytics for cycle time related feature selection in the semiconductor wafer fabrication system. Computers & Industrial Engineering , 106362.

Wein, L.M., 1992. On the relationship between yield and cycle time in semiconductor wafer fabrication. IEEE transactions on semiconductor manufacturing 5, 156–158.

Wu, G.L., Wei, K., Tsai, C.Y., Chang, S.C., Wang, N.J., Tsai, R.L., Liu, H.P., 1998. Tss: a daily production target setting system for fabs, in: 1998 Semiconductor Manufacturing Technology Workshop (Cat. No. 98EX133), IEEE. pp. 86–98.

Yugma, C., Dauzère-Pérès, S., Artigues, C., Derreumaux, A., Sibille, O., 2012. A batching and scheduling algorithm for the diffusion area in semiconductor manufacturing. International Journal of Production Research 50, 2118–2132.

Zarifoglu, E., Hasenbein, J.J., Kutanoglu, E., 2012. Lot size management in the semiconductor industry: Queueing analysis for cycle time optimization. IEEE Transactions on Semiconductor Manufacturing 26, 92–99.