Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

QUALIFICATION MANAGEMENT TO REDUCE WORKLOAD VARIABILITY IN SEMICONDUCTOR MANUFACTURING

Mehdi Rowshannahad
Stéphane Dauzère-Pérès

Ecole des Mines de Saint-Etienne – CMP
Department of Manufacturing Sciences and Logistics
CNRS UMR 6158 LIMOS
13541 Gardanne, FRANCE

Bernard Cassini

Soitec
Parc Technologique des Fontaines
Chemin des Franques
38190 Bernin, FRANCE

ABSTRACT

Variability is an inherent component of all production systems. To prevent variability propagation through the whole production line, variability must be constantly monitored, especially for bottleneck toolsets. In this paper, we propose measures to evaluate workload variability for a toolset configuration. Using industrial data, we show how making the toolset configuration more flexible by qualifying products on machines decreases variability. By quantifying the toolset workload variability, our variability measures makes it possible to estimate the variability reduction associated to each new qualification. The industrial results show significant workload variability reduction and capacity improvement.

1 INTRODUCTION

The dynamic business environment in semiconductor manufacturing industry calls for an increasing product-mix flexibility. Besides, production systems are restricted by machine capabilities. The ultimate goal of every manufacturing system is to satisfy customer demand while making the best possible use of the production facilities.

Variability has *stochastic* and *deterministic* sources. The stochastic variability sources are uncontrollable, and the most well-known are demand, machine breakdown, rework, operator delay, etc. On the other hand, deterministic variability sources are controllable. They include machine eligibility, batches, setups, re-entrant flows, etc. The controllable deterministic variability source considered in this paper is *product-to-machine eligibility*, called *qualification* in this paper.

Production variability influences production planning and scheduling, capacity planning, inventory management, equipment and labor cost, etc. (Kim and Alden 1997). Even little variability in bottleneck workcenters can cause high variability in the whole production line. Therefore, production variability reduction at bottleneck workcenters is crucial to prevent variability propagation to the whole production line. Production variability leads to loss of capacity. Therefore, due to expensive equipment cost, mastering variability is critical in semiconductor manufacturing. Production variability decreases as manufacturing systems become more flexible. The flexibility of a manufacturing system is determined based on to which extent a product can be allocated to a machine. Hence, capacity allocation determines the flexibility of a production system (Muriel, Somasundaram, and Zhang 2006).

This study is a continuation of the investigations of (Johnzén, Dauzère-Pérès, and Vialletelle 2011) in which flexibility measures are proposed to quantify the flexibility gain related to performing qualifications. However, variability is not explicitly considered in these measures. This is why, in this paper, we aim at reducing production variability by defining variability measures. The workload variations in a single workcenter in one period are considered. We want to show how variability is reduced when the flexibility of the toolset is increased by performing new qualifications. The variability measures proposed can be used to measure the impact of batch or single wafer processing. However additional constraints must be added to the optimization model. (Rowshannahad and Dauzère-Pérès 2013) specifically address the problem of batch processing. Through experiments with industrial data, we illustrate that production variability and manufacturing flexibility are two sides of the same coin.

This paper is organized as follows. Section 2 briefly reviews related studies and literature in the area of production variability and manufacturing flexibility. Section 3 details the framework of the study. The measures that we propose to evaluate the workload variability are presented in Section 4. In Section 5, numerical experiments on industrial data are discussed. Finally, Section 6 provides conclusions and proposes future work.

2 LITERATURE REVIEW

The present study is mainly related to two domains: *Production variability* and *manufacturing flexibility*.

Stochastic modeling has been widely used to measure the production variability of production lines (He, Wu, and Li 2007). Many studies have been done on the measurement of the variability of the production flows in a fluid modeling network (Ciprut, Hongler, and Salama 1999). In our study, we only consider one workcenter in one period and evaluate the variability using an optimization model.

The benefits of process flexibility in capacity utilization and sales increase in supply chains is extensively studied in (Jordan and Graves 1995). (Graves and Tomlin 2003) define a flexibility measure for supply chain systems. At plant level, manufacturing flexibility between plants and products, with various capacity limitations and demands is studied in (Boyer and Leong 1996). At workcenter level, (Johnzén, Dauzère-Pérès, and Vialletelle 2011) propose flexibility measures to evaluate the manufacturing flexibility of a given workcenter configuration according to the production volume or production time. The industrial implementation and consideration of special cases for new and alternate recipes and their impact on toolset capacity is further studied in (Rowshannahad, Dauzère-Pérès, and Cassini 2013).

The relationship between manufacturing flexibility and production variability is studied in (Muriel, Somasundaram, and Zhang 2006). The study is conducted on a multi-plant multi-product make-to-order manufacturing supply chain. Based on an optimization-based simulation model, it is shown that, by increasing the manufacturing flexibility, the production variability is reduced.

Production throughput variability is calculated for a single workstation with deterministic process times and random downtimes in (Kim and Alden 1997). Probability density function and variance of time to produce are developed for a fixed lot size. Finally, it is shown how the proposed probability density function can be used in discrete event simulation to generate a cycle time distribution of a lot size of one.

(Delp et al. 2006) define a complete X-factor contribution measure to identify the capacity constraining machine using raw process time, utilization, availability, variability of the processing time and arrival rate of the lots. This new measure is used to reduce the mean cycle time and cycle time variability.

3 TOOLSET WORKLOAD VARIABILITY AND MANUFACTURING FLEXIBILITY

In semiconductor manufacturing, the production system consists of workcenters called *toolsets*. A toolset is a collection of similar but not necessarily identical machines. The process performed on products visiting a toolset is called a *recipe*. In other words, a recipe is the process instructions to be performed on a product. The machines of a toolset are usually not capable of performing all recipes. Before the execution of a recipe on a machine, the machine must be *qualified* for that recipe. Increasing the number of qualifications adds flexibility to the manufacturing system. However, all recipes cannot or should not be qualified on all machines since recipe-to-machine qualification is costly and time-consuming. In this study, we consider a recipe to be either *(already) qualified* or *qualifiable* on a machine. By qualifying a *qualifiable* recipe, the toolset manufacturing flexibility increases. Figure 1 shows three configurations of a toolset. Figure 1a shows dedicated machines to recipes where no flexibility exists in the manufacturing system. In Figure 1b, each recipe is qualified on at least two machines making the manufacturing system partially flexible. Figure 1c depicts a totally flexible toolset where all recipes are qualified on all machines. In the next section, we define variability measures to evaluate to which extent the increase of the manufacturing flexibility contributes to workload variability reduction.

Increasing manufacturing flexibility implies at least one new qualification, enabling the qualified machine to process another recipe by creating a new link between the qualified recipe-machine couple.

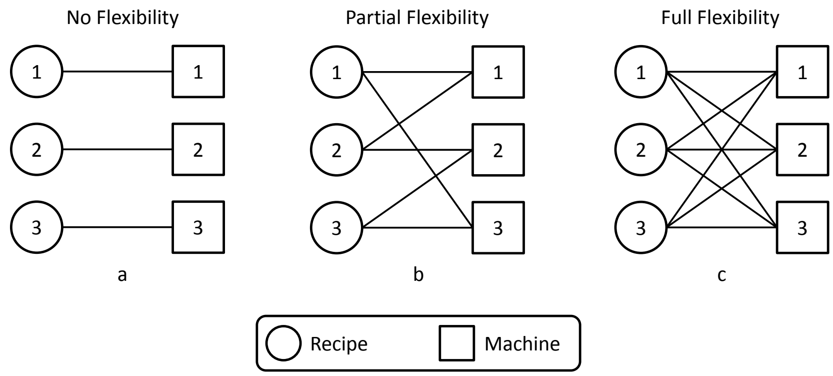

Figure 1: No Flexibility (a), Partial Flexibility (b) and Full Flexibility (c) Recipe-to-Machine Configurations (inspired from (Graves and Tomlin 2003) and (Muriel, Somasundaram, and Zhang 2006)).

As already specified in the introduction, the notion of *manufacturing flexibility* in this study refers to (Johnzén, Dauzère-Pérès, and Vialletelle 2011), where *flexibility measures* are defined to evaluate the flexibility gain associated to performing qualifications. One of the main objectives of their measures is to estimate the impact of qualifications on workload balancing. We have extended and implemented these measures in industry (see (Rowshannahad, Dauzère-Pérès, and Cassini 2013) and (Rowshannahad and Dauzère-Pérès 2013)). They are used to evaluate new qualifications in the daily fab operations. However, by only considering these flexibility measures, the impact of a qualification on the toolset workload variability is not clear to decision makers. This study aims at throwing light on the other side of qualification management via variability measurement. Moreover, compared to (Johnzén, Dauzère-Pérès, and Vialletelle 2011), we also consider machine capacity restrictions.

4 VARIABILITY MEASURES

In order to obtain a unique variability value, the optimal workload balance of the toolset according to its configuration must be calculated. By minimizing the proposed measures, which are inspired from statistical moments, the optimal workload balance of a toolset must be calculated. The notations used in this paper are listed below.

Parameters:

R	Total number of recipes,
M	Total number of machines in the toolset,
WIP_r	Total production volume of recipe r,
$TP_{r,m}$	Throughput rate of recipe r on machine m,
$Capa_m$	Capacity of each machine m,
$Q_{r,m}$	$\begin{cases} 1 \text{ if recipe } r \text{ is qualified on machine } m, \\ 0 \text{ otherwise,} \end{cases}$
γ	Workload balancing exponent ($\gamma \geq 1$).

Variables:

$WIP_{r,m}$ Production volume of recipe r allocated to machine m,

C_m Total production time on machine m ($C_m = \sum_{r=1}^{R} \frac{WIP_{r,m}}{TP_{r,m}}$).

The variability measures introduced in the next sections (Var^{Time}_{Uncapa} in Section 4.1 and Var^{Time}_{Capa} in Section 4.2) are used as the objective function ($Var^{\bullet}_{\bullet}$) of the mathematical model below. The only set of constraints (1b) of the model guarantees that the production volume of recipe r is only allocated to machines that are qualified for r, i.e. machines m such that $Q_{r,m} = 1$. By minimizing the selected variability measure subject to the set of constraints, the optimal toolset workload balance is obtained.

$$\min \quad Var^{\bullet}_{\bullet} \tag{1a}$$

$$\text{Subject to} \quad \sum_{m=1|Q_{r,m}=1}^{M} WIP_{r,m} = WIP_r \qquad \forall r \tag{1b}$$

$$WIP_{r,m} \geq 0 \qquad \forall r, m$$

This model can be solved by adapting the Active Set method described in (Johnzén 2009) and (Griva, Nash, and Sofer 2008). Detailing this method is out of the scope of this paper.

4.1 Uncapacitated Time Variability Measure (Var^{Time}_{Uncapa})

For the current toolset qualification configuration, by minimizing the sum of the total process time of each machine (2), the workload variability is calculated. Then, for each qualifiable recipe-to-machine couple, we re-calculate the variability (Var^{Time}_{Uncapa}) by virtually qualifying the qualifiable couple. In order to evaluate the variability reduction associated to each new qualification, the variability of each new configuration is subtracted from the variability of the initial configuration.

$$Var^{Time}_{Uncapa} = \sum_{m=1}^{M} (C_m)^{\gamma} \tag{2}$$

By increasing the workload balancing exponent (γ), the load of high speed machines is shifted to slower machines where qualification is allowed. Increasing γ creates a smoother workload distribution on the toolset.

4.2 Capacitated Time Variability Measure (Var^{Time}_{Capa})

Machine failures, operator unavailability, scheduled and unscheduled maintenance are sources of variability which affect the uptime of machines. The uptimes of each machine in the same toolset can be different and are considered to be deterministic in this paper. Var^{Time}_{Capa} (3) evaluates the workload variability of a toolset while considering the capacity of machines.

$$Var^{Time}_{Capa} = \sum_{m=1}^{M} (C_m - Capa_m)^{\gamma} \tag{3}$$

In order to calculate the variability reduction of each new qualification, as explained in Section 4.1, the variability of the current qualification configuration and the configuration after each new qualification must be calculated.

By increasing the workload balancing exponent (γ), the model tries to fit better the toolset workload to the available capacity by shifting workload from overloaded machines to less loaded machines.

5 INDUSTRIAL EXPERIMENTS

The *Capacitated Time Variability Measure* (Var^{Time}_{Capa}) (3) is used as the objective function of the mathematical model to conduct experiments on industrial data of a Thermal Treatment Toolset in an "SOI" (Silicon-On-Insulator) production line. The Thermal Treatment Toolset, which consists of non-homogeneous furnaces, is a bottleneck toolset. Each standard SOI product must at least visit three times this toolset. For the industrial experiments, the value of γ is set to 4. Careful observations of the shop floor workload allocation shows that any value between 4 to 6 suits the model for practical purposes.

First, we consider a data set for one period and study the impact of performing a single qualification on the percentage of the workload variability reduction. Other independent performance indicators used to interpret the workload balance are: The overall toolset workload variation percentage (4), overload (5) and unused capacity (underload) (6) variation percentages. The variation comparison for each performance indicator is simply calculated as shown in (7) for overload (OL) variation comparison. Using the static workload balance diagram, we show the impact of one new qualification on the workload variability. Finally, we discuss the impact of new qualifications on the reduction of the production variability for some industrial instances taken from the daily fab operations.

$$Workload\ Sum = \sum_{m=1}^{M} C_m \tag{4}$$

$$Overload\ Sum = \sum_{m=1|C_m \geq Capa_m}^{M} (C_m - Capa_m) \tag{5}$$

$$Unused\ Capacity\ Sum = \sum_{m=1|C_m \leq Capa_m}^{M} (C_m - Capa_m) \tag{6}$$

$$OL\ Comparison = \frac{(OL_{New\ Config.} - OL_{Current\ Config.})}{OL_{Current\ Config.}} \times 100 \tag{7}$$

First, we consider an industrial instance of a toolset consisting of 22 machines and 37 recipes for a single period. Diagram 2 depicts the current toolset workload balance. The vertical lines correspond to machine capacities. Each horizontal bar represents the workload of a machine. Bars above the capacity lines show the overloaded machines, and the opposite for underloaded machines.

By calculating the toolset variability associated to each new qualification (creating additional manufacturing flexibility), the qualification which reduces the most workload variability, is chosen. The workload balancing diagram for the new toolset qualification configuration (Figure 3) illustrates how the workload variability is reduced after performing one new qualification. Note that both diagrams have the same scale.

By continuing to perform new qualifications, the toolset capacity allocation improves. However, instead of showing the workload diagram, variability and performance indicators are presented in Table 1. It shows the workload variability, overload and unused capacity reduction and used capacity increase percentages after new qualifications, i.e. creating new links between recipe set and machine set. It can be observed that performing more and more new qualifications reduces less and less the production variability. A trade-off must be made between the cost of performing new qualifications and the benefit of reducing workload variability.

Figure 4 depicts the results of Table 1. It is worth to note that workload variations are not linear as the number of new qualifications increases linearly. Some qualifications decrease variability more than others. However, too many new qualifications do not decrease variability very much.

Table 2 presents how one new machine-to-recipe qualification affects production variability for ten industrial instances. In general, one new qualification reduces variability, overload and unused capacity drastically while only slightly increasing the total workload. Note that, in the first instance, the overload is

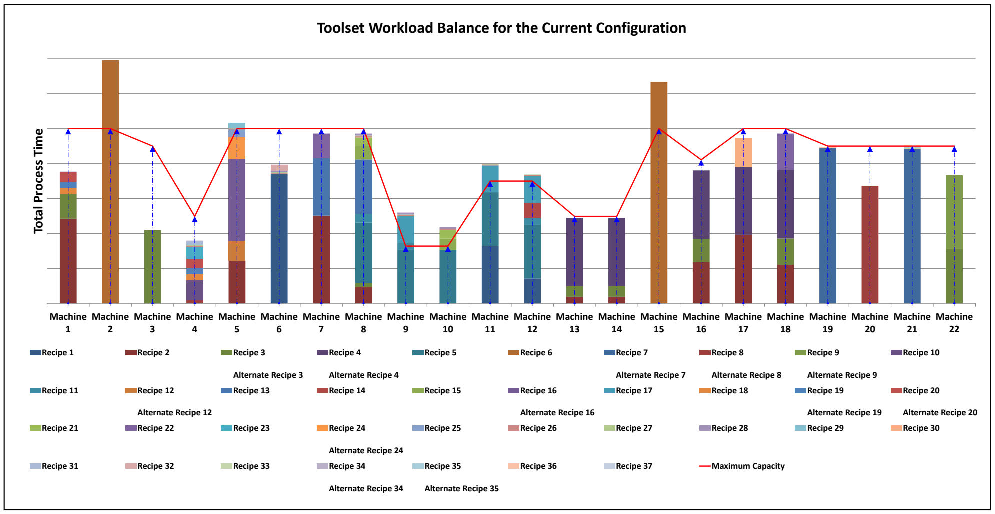

Figure 2: Toolset Workload Balance for the Current Recipe-to-Machine Qualification Configuration.

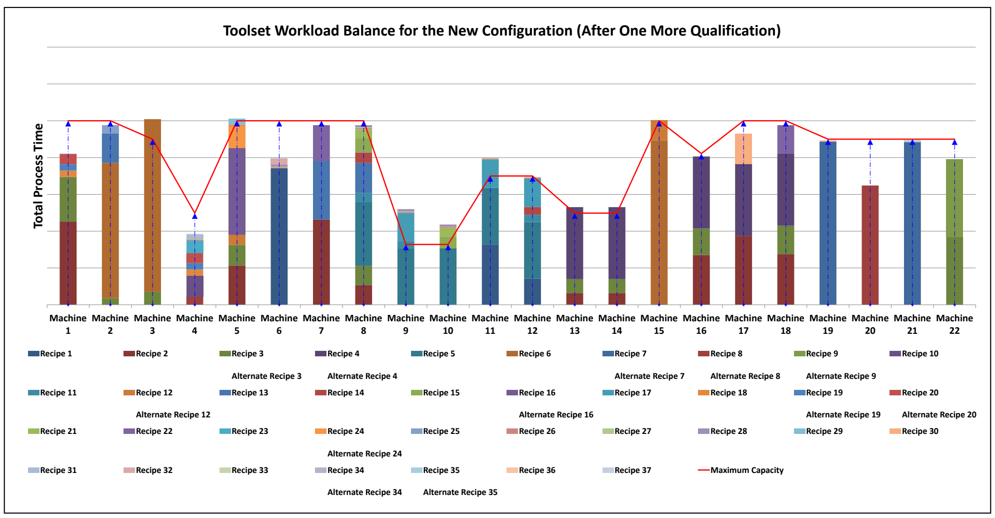

Figure 3: Toolset Workload Balance for the Configuration After Performing One New Qualification.

completely eliminated (-100%), capacity utilization is highly increased (20.20%) while the total workload only increases by 3.40%. The experiments illustrate again that the variations do not follow a linear pattern. Depending upon the data set and toolset configuration, nearly the same amount of variability reduction can lead to more or less impact on the performance measures. This is illustrated in Instances 1 and 10. While both instances record a variability decrease of about four percent 4% (-4.13% and -4.71%), the variations of workload (3.40% and 1.20%), overload (-100% and -1.41%) and unused capacity (-20.20% and -5.49%) are very different.

Tables 1 and 2 show that the toolset workload increase percentage is not equal to the variability, overload and unused capacity decrease. This implies that, by performing one new qualification, only a small increase of workload leads to a high decrease of variability. The reason is that the process times of recipes are different from machine to machine and, by creating a new link between a recipe-machine couple via qualification, a better capacity allocation becomes possible.

Table 1: Number of New Qualifications versus Variability Reduction and Performance Indicators Variations.

	Variation			
New Qualification(s)	Variability	Workload	Overload	Unused Capacity
1	-77.44%	3.33%	-37.33%	-32.72%
2	-86.32%	4.08%	-47.11%	-40.02%
3	-89.94%	4.55%	-57.58%	-44.66%
4	-94.41%	4.94%	-59.48%	-48.48%
5	-95.58%	5.22%	-62.22%	-51.20%
6	-96.08%	5.28%	-61.99%	-51.83%
7	-96.12%	5.23%	-65.82%	-51.31%
8	-96.60%	5.48%	-66.75%	-53.78%
9	-97.55%	5.84%	-72.56%	-57.33%
10	-97.55%	5.84%	-72.56%	-57.33%

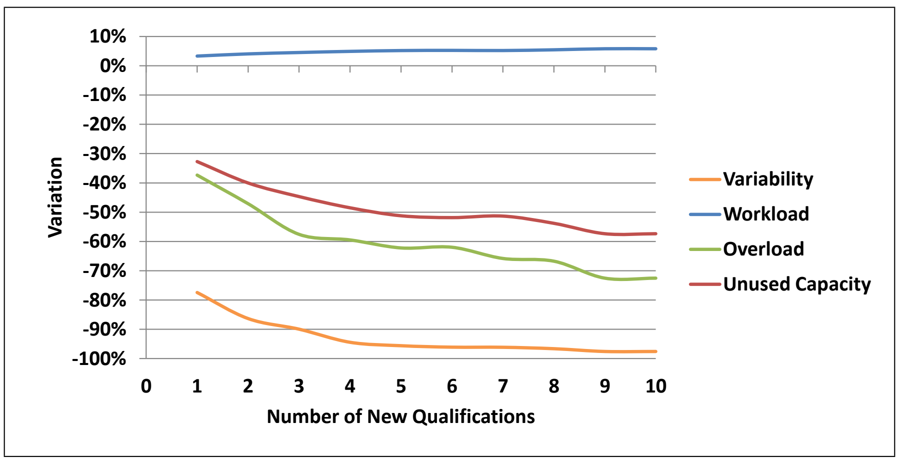

Figure 4: Toolset Variability, Workload, Overload and Unused Capacity Variation versus Number of New Qualifications.

6 CONCLUSIONS AND PERSPECTIVES

In this paper, we have studied the relationship between toolset qualification configuration (toolset manufacturing flexibility) and workload variability. A product can only be processed in a toolset when its associated recipe is qualified on at least one machine. Increasing the number of qualifications adds flexibility to the toolset and allows better capacity allocation.

Variability measures are presented to evaluate the workload variability of a toolset qualification configuration. Based on industrial data, we showed that, by performing the best qualification according to the proposed variability measures, toolset overload, unused capacity and also variability are reduced.

In conclusion, more manufacturing flexibility is required where the workload variability is the largest and not simply where the workload is the largest. In other words, more manufacturing flexibility absorbs workload variability. If the workload variability is low, meaning that (almost) all machines of a toolset are loaded equally according to their capacity, more manufacturing flexibility does not reduce variability. In this case, acquiring new machines might be necessary.

Table 2: Variability Reduction by Performing One New Qualification.

	Variation			
Instance Number	Variability	Workload	Overload	Unused Capacity
1	-4.13%	3.40%	-100.00%	-20.20%
2	-34.52%	2.16%	-46.76%	-7.36%
3	-76.11%	4.43%	-89.42%	-22.41%
4	-43.39%	3.99%	-28.97%	-16.76%
5	-3.29%	2.20%	-21.47%	-18.06%
6	-21.90%	2.29%	-20.05%	-15.80%
7	-62.82%	1.06%	-5.80%	-5.03%
8	-99.76%	4.57%	-30.10%	-87.41%
9	-49.65%	1.75%	-18.86%	-18.23%
10	-4.71%	1.20%	-1.41%	-5.49%

Several perspectives are possible for this study. An important source of variability is batching. The same variability measures can be used to evaluate how batches affect (increase) workload variability.

The proposed approach leads to local variability reduction. Using stochastic modeling, it would be interesting to integrate the present toolset variability measurement to decrease the global production variability. In this case, the impact of manufacturing flexibility on buffer stock requirements between workcenters can be studied.

Machine failure is often a major element of toolset variability. Although they reflect machine breakdowns, the machine capacities used in our measures are deterministic. It could be interesting to explicitly consider machine breakdown probabilities.

Finally, it could also be relevant to formalize the trade-off between the costs associated to performing and maintaining new qualifications and the gains related to the improved flexibility and variability quantified with our measures. This could lead to an interesting bi-criteria optimization problem. It remains to be seen if this will bring enough added value to the decision makers that are currently using our decision support system, since they will have to provide more information.

ACKNOWLEDGMENTS

This study has been done within the framework of a joint collaboration between the Center of Microelectronics in Provence of the École des Mines de Saint-Étienne (Gardanne, France) and SOITEC (Bernin, France). The authors would like to thank the ANRT (*Association Nationale de la Recherche et de la Technologie*) which has partially financed this study.

REFERENCES

Boyer, K. K., and G. Leong. 1996, October. "Manufacturing Flexibility at the Plant Level". *Omega-International Journal of Management Science* 24 (5): 495–510.

Ciprut, P., M.-O. Hongler, and Y. Salama. 1999. "On theVariance of the Production Output of Transfer Lines". *IEEE Transactions on Robotics and Automation* 15 (1): 33–43.

Delp, D., J. Si, and J. W. Fowler. 2006. "The Development of the Complete X-Factor Contribution Measurement for Improving Cycle Time and Cycle Time Variability". *IEEE Transactions on Semiconductor Manufacturing* 19 (3): 352–362.

Graves, S. C., and B. T. Tomlin. 2003. "Process Flexibility in Supply Chains". *Management Science* 49 (7): 907–919.

Griva, I., S. G. Nash, and A. Sofer. 2008. *Linear and Nonlinear Optimization*. 2nd ed.
He, X.-F., S. Wu, and Q.-L. Li. 2007, May. "Production Variability of Production Lines". *International Journal of Production Economics* 107 (1): 78–87.
Johnzén, C. 2009. *Modeling and Optimizing Flexible Capacity Allocation in Semiconductor Manufacturing*. Ph. D. thesis, Department of Manufacturing Sciences and Logistics - Center of Microelectronics in Provence, École des Mines de St-Étienne.
Johnzén, C., S. Dauzère-Pérès, and P. Vialletelle. 2011, January. "Flexibility Measures for Qualification Management in Wafer Fabs". *Production Planning & Control* 22 (1): 81–90.
Jordan, W. C., and S. C. Graves. 1995. "Principles on the Benefits of Manufacturing Process Flexibility". *Management Science* 41 (4): 577–594.
Kim, D. S., and J. Alden. 1997, December. "Estimating the Distribution and Variance of Time to Produce a Fixed Lot Size Given Deterministic Processing Times and Random Downtimes". *International Journal of Production Research* 35 (12): 3405–3414.
Muriel, A., A. Somasundaram, and Y. Zhang. 2006, April. "Impact of Partial Manufacturing Flexibility on Production Variability". *Manufacturing & Service Operations Management* 8 (2): 192–205.
Rowshannahad, M., and S. Dauzère-Pérès. 2013. "Qualification Management with Batch Size Constraint". In *Proceedings of the 2013 Winter Simulation Conference*, edited by E. R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, Number 2007, 3707–3718: Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
Rowshannahad, M., S. Dauzère-Pérès, and B. Cassini. 2013, May. "Qualification Management and its Impact on Capacity Optimization". *Advanced Semiconductor Manufacturing Conference (ASMC)*:174–179.

AUTHOR BIOGRAPHIES

MEHDI ROWSHANNAHAD holds an MSc in "Science and Technology of Management" from the University of Montpellier II in France, and an MSc in "Industrial Engineering" from the Grenoble National Polytechnics Institute. He is currently working at Soitec (Bernin, France) and is a PhD student at the Center of Microelectronics in Provence of the École des Mines de St-Étienne. His research interests are optimization, capacity and production planning, and scheduling. His email address is rowshannahad@emse.fr.

STÉPHANE DAUZÈRE-PÉRÈS is Professor at the Center of Microelectronics in Provence (CMP) of the EMSE. He received the Ph.D. degree from the Paul Sabatier University in Toulouse, France, in 1992; and the H.D.R. from the Pierre and Marie Curie University, Paris, France, in 1998. He was a Postdoctoral Fellow at the Massachusetts Institute of Technology, U.S.A., in 1992 and 1993, and Research Scientist at Erasmus University Rotterdam, The Netherlands, in 1994. He has been Associate Professor and Professor from 1994 to 2004 at the Ecole des Mines de Nantes in France where he headed the team Production and Logistic Systems (about 20 members) between 1999 and 2004. He was invited Professor at the Norwegian School of Economics and Business Administration, Bergen, Norway, in 1999. Since March 2004, he is Professor at the Ecole des Mines de Saint-Etienne, where he headed the research department Manufacturing Sciences and Logistics (SFL, about 20 members) from 2004 to 2013 and the CMP from 2013 to 2014. His research interests broadly include modeling and optimization of operations at various decision levels (from real-time to strategic) in manufacturing and logistics, with a special emphasis on semiconductor manufacturing. He has published more than 50 papers in international journals and contributed to more than 120 communications in conferences. Stéphane Dauzère-Pérès has coordinated multiple academic and industrial research projects, and also five conferences. His email address is dauzere-peres@emse.fr.

BERNARD CASSINI is currently Industrial Engineering manager of Electronics Division at SOITEC, in Bernin, France. He has 20 years experience in Semiconductor Industry. He has joined Soitec in 2001 as Bernin Fab operations manager. Then he moved to Business operations manager for the Digital Business Units. Prior to joining Soitec, Bernard Cassini has held different management positions within Industrial

Operations at Motorola Semiconductor, in Toulouse, France. He holds a Master's degree in Material Engineering from École Centrale de Lyon, France. His email address is bernard.cassini@soitec.com.