Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A.A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

THE DETECTION AND THE CONTROL OF MACHINE/CHAMBER MISMATCHING IN SEMICONDUCTORMANUFACTURING

Aabir Chouichi
Jakey Blue
Claude Yugma

Department of Manufacturing Sciences and Logistics
Center of Microelectronics in Provence
Ecole des Mines de Saint-Etienne
880 Route de Mimet
Gardanne, 13541, FRANCE

Francois Pasqualini

Process Control Enginering
STMicroelectronics Crolles
850 Rue Jean Monnet
Crolles,38920,FRANCE

ABSTRACT

Achieving desired yields in advanced manufacturing industries, as is the case for semiconductor fabrication, requires perfectly matching the performance of parallel machines/chambers at all production steps. Moreover, in the high-mix/low-volume manufacturing environment of the IC makers, the need for highly precise production processes makes the task of matching identical machines/chambers more and more complicated and challenging. In this paper, a methodical approach todeal with machine/chamber mismatching using all available data such as sensor data,product measurements, and maintenance logis proposed and validated with real practices. In fact, after detecting and identifying the key variables source of variance in the process, a modeling step known as Virtual Metrology is performed to quantify accurately the recipe adjustments i.e., R2R control, which will match as much as possible both sensor data and metrology measurements between similar machines/chambers.

1 INTRODUCTION

Due to the onward evolution of microchips applications and the numerous customer inquiries, the semiconductor industry become more vital and competitive. Improve overall microchips performance,yield equivalent product quality and minimize the margins of error are from the top priorities of semiconductor manufacturers. This is not easy to achieve since semiconductor manufacturing is a highly complex process consisting of more than 200 operations with reentrant routes. If not controlled,the cumulative effect from these operations can have a detrimental effect on the final yield. One of the critical challenges in this industry is then to minimize the process variability accumulated and augmented through the consecutive production operations. In a High-Mix/Low-Volume (HM/LV) fabrication environment such as the fab under study, several hundred products are manufactured through successive steps in which identical machines/chambers are performing with objectives of maximizing the throughput and optimizing the machine utilization (see Figure 1).

Unsurprisingly, after processing several products with very different recipe set-ups, machine/chamber conditions will not be identical. To reduce the variability of the whole process, the critical and fundamental task is to obtain a perfectly matched performance between the parallel machines/chambers in one area. Usually, the electrical tests done at the end of the process provide the final evaluation indicators used to trace back the operations where similar machines/chambers are mismatched. Waiting for the results of electrical tests, despite their high relevance, is too risky given a simple fact that a large number of wafers might be defective and potentially need to be scrapped when the testing results are not validated. The exploitation of the metrology measurements collected after each operation is therefore essential.

In the studied fab, the major approach of calibrating the mismatched metrology measurements is through the Run-to-Run (R2R) control loops. For example, in order to match the post-etching metrologies from two machines, the process time is often regulated from one lot to another. However, the effect of R2R regulators is limited. Some mismatching cases have been detected and should be investigated given the required high precision of the production operations.

Figure 1:Schematic representation of parallel tools.

This article presents a methodology for tracing the mismatching issue back to its root causes to find out the key machine parameters where the problem comes from and control them correctly. The rest of this paper is structured as follows: in Section 2, the literature related to machine/chamber mismatching and which is presenting some statistical methods used to deal with it will be reviewed. Then, the proposed methodology is described in Section 3.In Section 4, a real case study illustrating the approach is presented followed by the concluding remarks in Section 5.

2 LITERATURE REVIEW

The problem of matching machines/chambers can be divided into two main axes:The identification of any variations between identical machines/chambers using all available data and the correction of the mismatching detected addressing its root causes. The detection part has been treated and handled in several works according to the problem domains.

A classical approach proposed by Davis and Lian(2005) is to use only machine sensor readings. Principal Component Analysis (PCA) was performed on the collected data to compare the component scores between studied machines/chambers and to identify the parameters source of any detected mismatching. Being the most popular dimension reduction method, PCA does not guarantee a good separation between classes since the objective is to retain as much variance as possible on the principal components.

For the same purpose, a more adapted method to the problem was adopted by Cherry and Qin (2008). In fact, Fisher Discriminant Analysis (FDA) was applied to differentiate the classes of product quality data from different machines or chambers. Preventive Maintenance (PM) was then proposed to calibrate the mismatching machines/chambers. However, FDA requires fundamental assumptions related to the data distribution which might restrict its applicability. As an alternative, non-linear approaches, such as Kernel Discriminant Analysis, were proposed by You et al. (2011) but are considered as difficult to implement especially on the data with small sample sizes.

To fix the problems related to data distribution, Zhu and Martinez (2006) have chosen to approximate the distribution of each class with a mixture of Gaussians before applying FDA. The major difficulty of this method is to determine the optimal number of Gaussians per class. This meets the initial need for class separation but may distort the equipment parameters retained as main contributors since the distance between the created subclasses is also maximized.

Apart from the problem of choosing the right discriminant methods, the limitation of the approaches used lies in considering matching the sensor data as an objective in itself and not a way to ensure a good quality production.

Pan et al. (2012) present results by including measurement data in the analysis. Key variables that best separate the classes of sensor data and that are correlated with product quality are determined using many basic statistical methods, such as F-test, IQR-test and linear regression. However, this remains insufficient given the multi-dimensional nature of the machine/chamber matching process that was clearly explained byJimmy et al.(2014). The matching process usually covers hardware, software, tool sensors, process, metrology, maintenance, end of line electrical test and yield. Due to the hetero-granularities of different data types, it was difficult to consider them all together in such a complex industrial context.

Regarding the second part which treat the correction of detected mismatching,a limited number of works approached it because of its complexity seen thata large number of criteria should be considered before modifying the values of machine parameters.

Jimmy et al.(2014) provide a chamber matching solution which consists of using R2R control to make the states of the chambers closer together.This solution is interesting but cannot be applied directly before modeling the relation between machine parameters and metrology measurements.

The methodology presented in this paper tries to complete the existing state of art in this direction. Its aim is to consolidate all the data that are regularly collected and monitored during the wafers production in order to identify and to control the root causes of any significant mismatches among the machines/chambers.

3 METHODOLOGY

In the studied wafer fab, a wide range of products is manufactured through many production steps. Each step may be performed by similar machines/chambers to optimize and speed up the production process. Quality measurements are taken after most operations to ensure the compliance with the limits set for quality targets.These metrology data are the first indicator of any worry or problem since they are directly linked with the expected features of the wafers and have an impact on the electrical tests made at the end of the process. Parallel machines or chambers should normally yield the same product quality and therefore, any significant difference between the measurement data of these machines/chambers on the same recipe/type of product has to be investigated.

Fault Detection and Classification (FDC) data are usually defined as a collection of sensor readings during a wafer process.Apart from their traditional use for tracking and detecting machine problems, FDC data can be used to compare the operating states of the identical machines/chambers and to detect the root causes of any mismatching related to equipment conditions.

To make a precise and exhaustive study, an R code is developed in order to perform the preprocessing of FDC data. This code is structured as follows: Firstly, a filtering is done to keep only the comparable FDC parameters then the setting ones are centered by subtracting the variable setpoint from each of its initial values to avoid misleading that may occur when comparing the incomparable. The quality of temporal data is verified by finding parameters/steps where the collection was interrupted due to connection break or high speed equipment. Recently, the univariate indicators are calculated based on temporal data across all the retained steps. Once the FDC data are cleaned and pretreated, the parameters are analyzed to identify the ones that best separate the mismatched machines/chambers. Since the process machine/chamber of each wafer is the prior information, the classification problem is supervised. Its objective is twofold: to verify if the classes are well separated and to find the parameters that best describe the differences between the groups.

Given the large number of FDC parameters retained for each analysis, the use of a multivariate classification method is necessary. After testing several methods on multiple datasets ,and according to the nature of the data to be analyzed, Partial Least Squares Discriminant Analysis (PLS-DA) is the classification method chosen to separate the wafers according to their machine/chamber of process. In fact,PLS-DA is a powerful method of classification widely applied in chemistry and medical sciences (Sun 2009). It is a variant of PLS where the response variable Y is categorical to indicate the class of each observation. PLS-DA is known to work well for large data sets with potentially collinear variables, and for small data sets with large

amount of variables. Usually, the loadings in PLS-DA indicates the importance of machine parameters in class separation. This can be misleading as it shows only the contribution ofeach variable to the principal components direction. Then, it doesn't take into account the variances of each variable and the correlation between them (Cadima and Jolliffe 2011).

Two variable selection methods, Variable Importance in Projection (VIP) (Farrésa et al. 2015) and Selectivity Ratio (SR) **(**Farrésa et al. 2015) were tested on two different data sets to select the one that identify the most discriminating variables. The VIP measure of the variable j is defined as :

$$VIP_j = \sqrt{\frac{p\sum_{f=1}^{F} R^2(y,t_f)(\frac{w_{fj}}{\|w_f\|})^2}{\sum_{f=1}^{F} R^2(y,t_f)}}$$

where, w_{fj} is weight of the j^{th}predictor variable in component f and $R^2(y,t_f)$ is fraction of variance in y explained by the component f. The Selectivity Ratio of the variable j is the ratio between the explained and the residual variance for this variable. It turned out that the SR criterion can keep non-significant parameters and neglect other more important ones. This result was derived using the correlation coefficients between the parameters and the discriminating criterion i.e., process chamber as well as the FDC time signals.VIP is therefore retained and calculated for each parameter on the basis of the PLS-DA results. The parameters having a VIP greater than 1 are the more relevant ones to explain the response variable variations.

Once mismatched FDC parameters between identical machines /chambers are identified, the next step is to adjust them in order to ensure a good matching level. The difficulty of this step lies in the fact that only the direct FDC parameters can be adjusted. Any modification made impact both the indirect parameters and the measurement data. The mismatching correction is therefore critical considering the need to take into consideration the link existing between all available data.

The idea is to model the relationship between the measurement data and the FDC parameters for the identical machines/chambers, then find the best recipe for both matching the FDC parameters previously selected using the VIP criterionand meeting the target for each measured parameter. In this paper, we will rather focus on the modeling part. Since we do not know the nature of the link between the FDC parameters and the measurement data, two regression methods have been tested: Elastic NET(Zou and Hastie2004)and Support Vector Regression(SVR)(Awad and Khanna 2015).

Elastic Net is a linear regression which allows both to select the variables to be introduced into the model by penalization and to circumvent the collinearity problem even with a large number of explanatory variables or predictors. If the linear equation has the following form :

$$Y = bX + e$$

The criteria to minimize is :

$$\sum_{i=1}^{n}(y_i - \sum_{j=1}^{p} x_{ij}\, b_j)^2 + \lambda\,(\,\alpha \sum_{j=1}^{p} b_j{}^2 + \lambda\,(1-\alpha)\sum_{j=1}^{p} |bj|\,)$$

where, $\alpha \in [0, 1]$ and λ a positive parameter.

Support Vector Regression (SVR) is a non-parametric technique that allows the construction of a non-linear model using kernel function. The SVR model has the following form :

$$Y_i = W\, k(x_i, x) + b$$

where, $k(x_i,x)$ is the kernel function. One of the most popularfunctions is radial basis which is expressed as:

$$k(x, x') = e^{-\frac{\|x-x'\|^2}{2\partial^2}}$$

where, ∂ is a free parameter.

Two models will be realized: Thefirst one to link the direct FDC parameters with the measurement data and the second one to link the indirect FDC parameters to the direct ones. Some indirect parameters can be affected by the wear of the parts. The duration since the last maintenance operations realized on the machine/chambers will be introduced in the model.

4 INDUSTRIAL CASE STUDY:

A real case study will be conducted to illustrate the approach introduced in the previous sections of this paper.An operation from the etch workshop is chosen due to its criticality and its impact on the whole production process. The tool responsible for carrying out this operation will be noted EQ. It's composed of three identical chambers CH1, CH2 and CH3. After the etching operation, two important parameters are measured, the thickness after etching (TH) and the dimensions of the small patterns on the wafer called Critical Dimensions (CD).The sensor data of 650 samples processed in the three chambers (EQ-CH1,EQ-CH2,EQ-CH3) were retrieved. The FDC univariate indicators across all the critical steps were calculated based on temporal data via a developed R code. In total, we get 105 summarized indicators.

4.1 Application of Discriminant Analysis on FDC data

The first part of this approach as explained previously is to detect the machine/chamber mismatching using FDC data. In fact, the discriminant method PLS-DA is applied to the FDC calculated indicators to detect if there is a significant difference between the classes, determined in this case by the process chamber of each sample. The number of components used in this method was chosen by minimizing both the overall error and the Balanced Error Rate. Figure 2 displays the results of applying PLS-DA to the three chambers. Each wafer, represented by the 105 univariate indicators is projected into the subspace based on the first two PLS components that explain 41 % of the total variability. As can be seen in Figure2, the chambers are clearly separated, especially by the first component. The key parameters retained using the VIP criterion are listed in Table 1.

4.2 The Construction of the Regression Models

The parameters given in the Table 1 are those that must be matched between these similar chambers. They can be either direct ones i.e., that are settled at the beginning of the operation respecting the associated recipe and have an impact on measurement data, or indirect ones i.e., that can be viewed as the responses from the configuration of the direct ones and may dependent on parts wear. The proposed procedure is presented in Figure 3. In fact, Virtual Metrology technique will be used to predict wafer's characteristics based on sensor data. Existing links between machine parameters i.e., direct and indirect ones, will be also modeled. The regression analysis results are used to evaluate the impact of matching the discriminating parameters, retained using VIP, on metrology measurements. The overall objectif is to adjust the recipe settings in a way thatmaintain the classical R2R control aim i.e., keeping the process output as much closer as possible to the target, while taking into account the new constraint of matching machine/chamber parameters.

The first critical step preceding the matching is then to model, for each chamber studied, the links between the FDC parameters and the measurement data, taking into account the chambers conditions.

In this section, and given the similarity of the approach, we will only focus on one of the three studied chambers (EQ-CH1) and we will try to model the impact of the direct FDC parameters on both the indirect FDC parameters and the thickness measured after etching (TH).

The most important and the most critical model to build is the one that links the TH with the direct FDC parameters. We have introduced in the model the process time of the step which is regulated with an R2R loop since the other steps all have a fixed time. As explained previously,the quality of the predictions obtained using two regression methods i.e.,Elastic Net and Support Vector Machine (SVR) with a radial basis kernel function,was compared using two different measures: Mean Absolute Deviation (MAD) and

Mean Absolute Percentage Error (MAPE).It turned out that good fitting modelisthe one build using SVR as shown in Table 2.

Apart from the classical validation of the model,another industrial validation is required. Indeed, it must be verified that the regression obtained has a meaning from the physical point of view.According to the experience of some process engineers, TH is only impacted by some steps of the operation. However, the model makes others out as main contributors. But this is possible given the link between the steps and the impact that any step can have on the following.

Regarding the indirect parameters, a simple linear regression was sufficient to build a first logical model since we already know the direct parameters impacting each indirect parameter.For example, the temperature received must depend primarily on the set temperature and other unknown parameters that were determined in the model. The calculation shows a good Root-Mean-Square Error (RMSE) of 0.1.

Once the regression models are built for all machines/chambers studied, the values of direct FDC parameters identified as source of mismatching are progressively varied, while respecting the limits of each recipe, and the impact that it generates on metrology data and indirect FDC parameters is quantified. This makes it possible to evaluate the necessary adjustments in production recipe tobring the FDC parameters of identical machines/chambers closer while keeping the measured parameters on target.

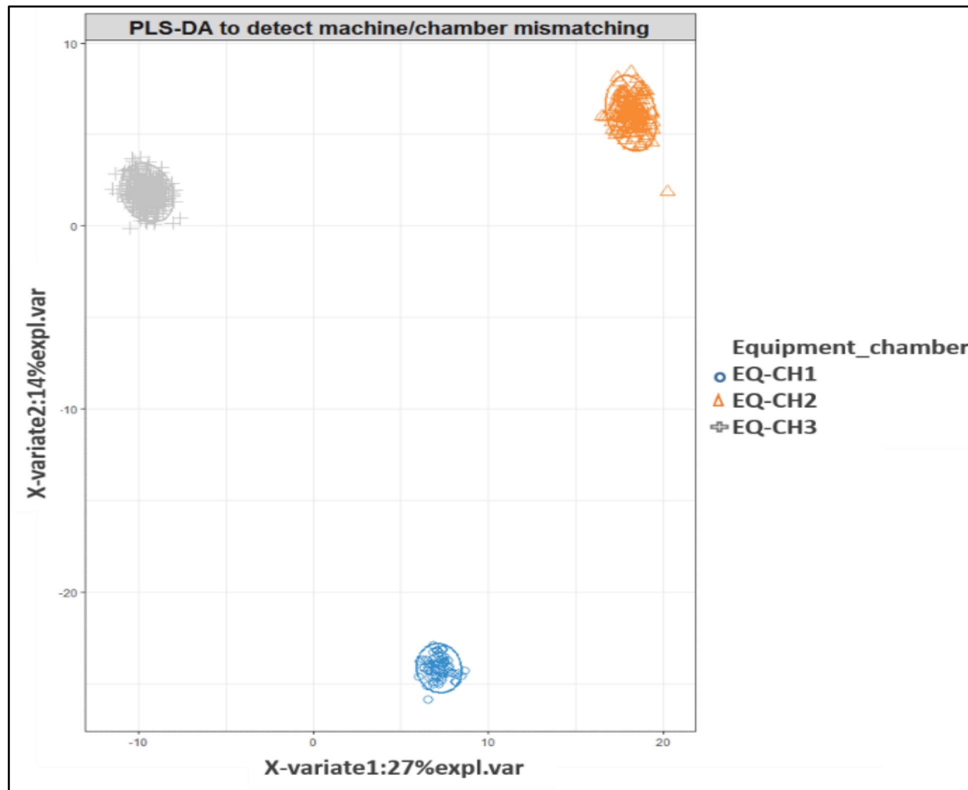

Figure 2:Scatter plot of the three chambers after projection onto the first two PLS components.

Table 1:The key parameters that separate the chambers selected using the VIP criterion.

Parameters	VIP	Parameters	VIP
mean_11_Flow_MFC_N2_500	1,897	mean_4_Pressure_ProcessManometer	1,749
mean_4_Flow_MFC_CF4_200	1,897	mean_14_Pressure_HeBackside_OuterZone	1,749
mean_11_Flow_MFC_AR_1000	1,897	mean_3_Pressure_HeBackside_OuterZone	1,749
mean_14_Flow_MFC_CO_500	1,897	mean_11_Pressure_HeBackside_OuterZone	1,749
mean_3_Flow_MFC_C4F6_50	1,897	mean_11_Pressure_HeBackside_InnerZone	1,749
mean_3_Flow_MFC_CF4_200	1,897	mean_4_Pressure_HeBackside_OuterZone	1,749
mean_11_Flow_MFC_C4F8_50	1,897	mean_7_Pressure_ProcessManometer	1,748
mean_14_Flow_MFC_C4F8_50	1,897	mean_3_Pressure_ProcessManometer	1,748
mean_14_Flow_MFC_AR_1000	1,897	mean_14_Power_RF2MHzGen_Delivered	1,716
mean_4_Flow_MFC_C4F6_50	1,897	mean_7_Power_RF27MHzGen_Delivered	1,63
mean_4_Pressure_HeBackside_InnerZone	1,791	mean_11_Power_RF27MHzGen_Delivered	1,625
mean_7_Pressure_HeBackside_InnerZone	1,791	mean_14_Power_RF27MHzGen_Delivered	1,622
mean_3_Pressure_HeBackside_InnerZone	1,791	mean_7_Flow_MFC_O2_2000	1,614
mean_3_Power_RF60MHzGen_Delivered	1,778	mean_11_Flow_MFC_TuningO2_10	1,61
mean_7_Pressure_HeBackside_OuterZone	1,749	mean_14_Flow_MFC_O2_20	1,61
mean_14_Pressure_HeBackside_InnerZone	1,749	mean_14_Flow_MFC_N2_500	1,607

Table 2: The evaluation of prediction quality of two regression methods using MAD and MAPE metrics.

Methods/Measures	MAD	MAPE
Elastic Net	0.62	1.57
Support Vector Regression	0.26	0.44

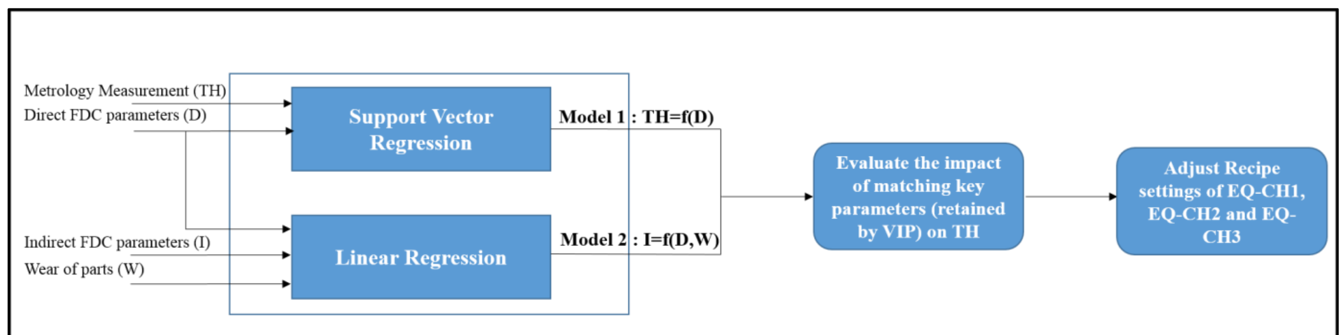

Figure 3:Proposed procedure to control machine/chamber matching.

5 CONCLUSION

In this paper, we identify the key FDC parameters that should be controlled in order to reduce the studied etching process variability. We also build up the models to quantify all the existingcausal relationship between the metrology measurements and the FDC parameters.The originality of the work comes from proposing a methodology that combine all data sources, usually used separately, in order to control process deviation in semiconductor manufacturing. Many statistical methods are applied on real fab datain a complex context as the objective is to adjust production recipes to ensure the matching of both machine parameters and process outputs between identical machine/chambers.This therefore requires a lot of precision in the results provided to be really exploited in fab. The ongoing work is then to validate these built models to ensure their compatibility with the industrial and physical side.

REFERENCES

You, D., C. Hamsici, and M. Martinez. 2011. "Kernel Optimization in Discriminant Analysis." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33(3):631–638.

Sun, D. W. 2009. *Infrared Spectrometry for Food Quality and Control*. Dublin: Academic Press.

Davis, M. F., and L. Lian. 2007. *Method for Automatic Determination of Semiconductor Plasma Chamber Matching and Source of Fault by Comprehensive Plasma Monitoring*. United States Patent: US 7169625 B2.

Cherry, G., and S. J. Qin. 2008. "Fisher Discriminant Analysis for Semiconductor Batch Tool Matching." In *Proceedings of International Federation of Automatic Control*, July 6th-11th, Seoul, Korea, 9144-9148.

Zou, H., and T. Hastie. 2005. "Regularization and Variable Selection via the Elastic Net." *Journal of the Royal Statistical Society Series B* 67(2):301–320.

Cadima, J., and T. Jolliffe. 2011. "Loading and Correlations in the Interpretation of Principle Components." *Journal of Applied Statistics* 22(2):203–214.

Moyne, J., M. Yedatore, J. Iskandar, P. Hawkins, and J. Scoville. 2014. "Chamber Matching Across Multiple Dimensions." In *Proceedings of the 25th Annual SEMI Advanced Semiconductor Manufacturing Conference*, May 19th-21st, Saratoga Springs, USA, 86–91.

Awad, M., and R. Khanna. 2015. "Efficient Learning Machines." Berkeley, California: Apress.

Farrésa, M., Y. Platikanova, L. Tsakovskib, and R.Taulera. 2015. "Comparison of the Variable Importance in Prediction (VIP) and of the Selectivity Ratio (SR) Variable Selection Methods in the Analysis of Three Different Data Sets." *Journal of Chemometrics* 29(10):528-536.

Zhu, M., and M. Martinez. 2006. "Subclass Discriminant Analysis." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(8):1274 –1286.

Pan, T., D. S.-H. Wong, and S. Jang. 2012. "Chamber Matching of Semiconductor Manufacturing Process Using Statistical Analysis." *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews* 42(4):571–576.

AUTHOR BIOGRAPHIES

AABIR CHOUICHI is a PhD student in Department of Manufacturing Sciences and Logistics at École des Mines de Saint-Étienne in partnership with STMicroelectronics Crolles. After obtaining an industrial engineering degree, she starts a thesis in January 2017 to work on the chamber/machine mismatching issue in semiconductor industry.As the corresponding author, his email address is aabir.chouichi@emse.fr.

JAKEY BLUE is Associate Professor of Department of Manufacturing Sciences and Logistics at École des Mines de Saint-Étienne, France. He received his Ph.D. in Mechanical Engineering from National Taiwan University, Taiwan in 2010. His current research interests include statistical data mining, demand planning and forecasting, object-oriented modeling and implementation, and, in particular, the applications

to semiconductor engineering data analysis and advanced process control. As the corresponding author, his email address is Ian@emse.fr.

CLAUDE YUGMA is Professor of Department of Manufacturing Sciences and Logistics at École des Mines de Saint-Étienne, France. He received his Ph.D. in Operational Research and Combinatorial from INP Grenoble, France in 2003. His current research interests include scheduling, simulation and planning inthe semiconductor production environment. Recently, he began to work on Inventory Routing Problems (IRP). His email address is yugma@emse.fr.

FRANCOIS PASQUALINI is Senior Member of Technical Staff in charge of technical coordination ofAdvanced Process Control activities and Statistics in STMicroelectronics Crolles that he joined in 1999.From 1983 to 1988 he was a research engineer at Philips Research Laboratories in Limeil-Brevannes, France. From 1989 to 1999 he worked at IBM Microelectronics Division in Corbeil-Essonnes, France.His interests were in the field of line monitoring, defect density, yield modeling, design of experiments, and statistical process control. His email address is francois.pasqualini@st.com.