

Virtual Metrology applied in Run-to-Run Control for a Chemical Mechanical Planarization process*

M.A. Jebri¹, E.M. El Adel¹, G. Graton^{1,2}, M. Ouladsine¹ and J. Pinaton³

¹Aix-Marseille Université, CNRS, ENSAM, Université de Toulon, LSIS UMR 7296, Av. Normandie-Niemen, 13397 Marseille, France

²Ecole Centrale Marseille, Technopôle de Château-Gombert, 38 rue Frédéric Joliot-Curie, F-13451 Marseille Cedex 13, France

³STMicroelectronics, ZI Rousset, 13106 Rousset Cedex, France

*This work is part of the European project "INTEGRATE", and carried by STMicroelectronics Fab of Rousset

E-mail: mohamed-ali.jebri@lsis.org

Abstract. This paper deals with missing data in semiconductor manufacturing derived from a measurement sampling strategies. The idea is to construct a virtual metrology module to estimate non measured variables using a new modified Just-In-Time Learning approach (JITL). The aim of this paper is to integrate estimated data into product control loop.

In collaboration with our industrial partner STMicroelectronics Rousset, the accuracy of the proposed method is illustrated by using industrial data-sets derived from Chemical Mechanical Planarization (CMP) process that enables us to compare results obtained with the classical and the modified version of JITL approach. Then, the contribution of the estimated data is shown in product quality improvement.

1. Introduction

In semiconductor manufacturing, a lot which contains usually 25 wafers, can undergo plenty of different treatments (over than 350 according to [1]) to obtain the final product. Among the known and used treatments in semiconductor manufacturing is the Chemical Mechanical Planarization (CMP) process. This latter is a polishing process, which uses a chemical slurry formulation and mechanical polishing process to remove unwanted conductive or dielectric materials on the silicon wafer. In a same CMP equipment, many products characterized by different targets (desired properties) and different recipes, can be processed.

At the end of this step, and to control product quality, the layer thickness of each wafer of the lot can be measured using one measurement strategy (according to [2] and [3]). Depending on whether the measurement equipment is located within or outside the CMP process, two measurement strategies can be considered. The first refers to the Integrated Metrology Module (IMM) and the second is called Stand-Alone Metrology (SAM), where the metrology tool is located outside the process equipment. In the Integrated Metrology Module (IMM) case, all products (lots and wafers) are measured. These physical measurements can then be used for the R2R control, in order that the process operates in the best conditions and improves the product quality.

While this configuration of measuring (IMM) is very expensive, and requires a high cycle time, it provides a good process control and it is applied when the CMP action seems to be critical in the production chain. However, in the Stand Alone Metrology (SAM) case, two strategies are opposed. The first one consists in a control at the end of all treatments. This makes the situation hazardous because in the case of defective lots, the fault is detected too late and lots of resources may be lost. The second strategy consists of measuring the products after each CMP treatment. This is not beneficial because it



requires additional processing time and may slow down the production and reduce the Fab efficiency. It implies also a high cost and introduces an added time delay to the process [4].

To overcome these two main issues, a compromise is achieved by sampling the physical metrology. According to [5], at lot level, the sampling rate is fixed using the criticality of step, machine and the product requirement. However, the main drawback of these measurements sampling is the lack of information. When layer thicknesses $y_j(k)$ are not provided, the recipe parameters $u_j(k)$ are not updated at each run. Then R2R controller is not working properly. Consequently, product quality cannot be well controlled. So, it is not possible to verify whether the variable respects the control limits which are defined as the upper control limit (UCL) and the lower control limit (LCL). Thus, it is necessary to find out an approach that fills in the lack of information. This is the main purpose of virtual metrology (VM). VM is introduced in semiconductor manufacturing in order to cover the lack of layer thickness measurements by providing estimation $\hat{y}_j(k)$ for non measured products (when $y_j(k)$ is not available). Then, R2R takes into account estimations of missing data and the recipe parameters $u_j(k)$ can be updated at each run. Therefore, VM module with the SAM strategy have the same function of IMM strategy and reacts with the same way with R2R controller with a lower cost.

The interaction between VM module, SAM strategy and R2R controller is described in Figure 1.

The main goal of this paper is to create a VM module for n types of product, manifested by different targets (T_j , where $j \in \{1, \dots, n\}$), using the old sampled measurement $y_j(k)$ with their corresponding old process information $x_j(k)$. Then, to improve a reliable control of product quality, the VM module should provide accurate estimations to R2R control by improving the recipe parameters $u_j(k)$.

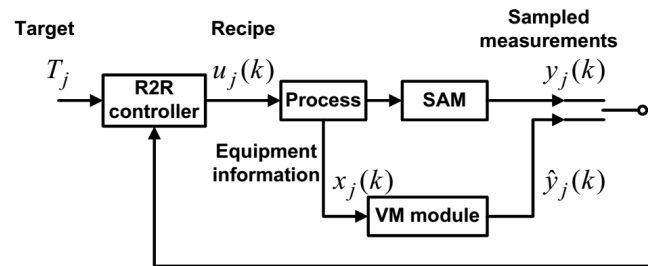


Figure 1. VM and R2R interaction

The paper is organized as follows: The R2R control and the virtual metrology are introduced in section 2. In section 3, a description of the proposed approach to deal with the lack of information and its interaction with R2R control is presented. By considering some data-sets extracted from CMP process and delivered by our industrial partner, the efficiency of the proposed method is given in section 4 followed by concluding remarks. The conclusion fenced by perspectives is given at the end of the paper.

2. Run-to-Run Control and Virtual Metrology

II-A. Virtual Metrology

In the literature, several definitions are given for the virtual metrology (VM). In [6], VM is defined as the use of mathematical models with available data to estimate the variables of interest. According to [7]; VM is presented as the prediction of metrology variables (either measurable or non-measurable) using the process state and the product's information. VM aims also to estimate non-measured variables called difficult-to-measure (DTM) variables [8], from a directly collected equipment data called also easy-to-measure (ETM) variables [8].

Due to the diversity of processes and products in semiconductor manufacturing, many approaches are proposed in the literature to build the VM module. However, there are no standard or generalized methods that can be applied.

Otherwise, the lack of knowledge of physio-chemical relationships in semiconductor manufacturing processes makes the determination of mathematical models describing the behavior of equipment

practically difficult. For this reason the majority of approaches are data-based methods.

There exist two strategies for VM model creation. The first one is based on global modeling strategies which exploits whole databases to provide a single estimation model for all areas of operation. The second one is based on local modeling strategies, also called "modeling on demand", where the model is dynamically built from the similar samples stored in the database to provide a single estimation, when it is needed.

In semiconductor manufacturing, equipment is characterized by many operating domains. Here, local models can provide more precise estimations than the global models which provide more general estimations [1].

The Just-In-Time Learning (JITL) approach is considered as a strategy based on local modeling. JITL approach provides an estimation when it is required [9].

For the two already listed strategies, there are two categories of modeling approaches.

The first strategy is based on linear approaches. The multiple linear regression (MLR) [11], the partial least squares regression (PLSR) defined in [10] which combines a principal components analysis (PCA) and MLR techniques, the recursive partial least squares regression (R-PLSR) [7], the principal component regression (PCR) [4] and Kalman filter [4] are considered as linear based approaches to build VM module.

The second strategy is based on non-linear based approaches. For instance, in [13], the authors adopt a neural network of the radial basis function to build VM model for a CVD process.

In another context, Su *et al.* develop, in [14], a VM module based on neural networks by back-propagation and recurrent neural networks on CVD equipment of 5th generation. Then, they make a comparative study of VM-based models.

In [15], Besnard *et al.* adopt tree ensembles method as a non-linear approach to estimate the oxide thickness.

In another context, in [16], techniques based on data mining are developed on a photo-lithography equipment. Output estimations are used then in R2R control.

To summarize, we notice that in semiconductor manufacturing, CMP equipment features are characterized by their non-linear aspect. We notice also that many product types are processed into CMP machines. So the idea is to use the combination of local models to cover different operation ranges (multiples product types) and linear modeling methods to keep the simplicity of the calculation and make their implantation easier for industrial context. This strategy allows to avoid the high cost of calculation time of techniques based on non-linear modeling.

In this context, JITL approach is proposed as a local modeling technique based on local linear models. This approach is already used for the development of VM modules in chemical processes which are characterized by non-linear aspects. In the next section we aim to introduce JITL algorithm as described by [9].

For our case, JITL, with some contributions, is opted as a VM module in CMP equipment of our industrial partner. Estimation outcomes from VM module are used then in R2R controller.

II-B. Run-to-Run control

The R2R Control module improves essentially product quality, increases yields and reduces defective products by optimizing recipe parameters from lot-to-lot using metrology data or estimations if necessary.

To implement R2R controllers, there exist different types of control laws (recipe parameters): the Exponentially Weighted Moving Average (EWMA) and the double EWMA (dEWMA) are the most widely explored in semiconductor manufacturing process.

The CMP process, for many processed type of product, can be modeled as:

$$y_j(k) = \alpha_j + \beta_j u_j(k) + \Delta(k) + \varepsilon(k) \quad (1)$$

where, $j \in \{1, \dots, n\}$ is the product type, $y_j(k)$ is the output at the end of run k , α_j is the linear process model intercept term, $u_j(k)$ is the recipe parameters at the start of run k and β_j is the system gain. $\varepsilon(k)$

and $\Delta(k)$ are respectively a white noise and systematic disturbance. Note that, these two terms depend only on the considered machine.

First, an offline estimation of the process gain $\hat{\beta}_j$ is required and remains unchanged. Then, at each run k , an estimation of the process model intercept term ($\hat{\alpha}_j(k)$) must be done.

When output is measured (SAM is applied), $\hat{\alpha}_j(k)$ is computed as follows:

$$\hat{\alpha}_j(k) = \lambda_j[y_j(k-1) - \hat{\beta}_j u_j(k-1)] + [1 - \lambda_j]\hat{\alpha}_j(k-1) \quad (2)$$

When output is estimated (VM is applied), $\hat{\alpha}_j(k)$ is computed by replacing the measured output $y_j(k-1)$ by its estimation $\hat{y}_j(k-1)$ in (2). $\lambda_j \in [0, 1]$ is a discounting factor used to minimize the influence of old data in favor of the new data. The control law for the next run is calculated as follows:

$$\hat{u}_j(k+1) = \frac{T_j - \hat{\alpha}_j(k)}{\hat{\beta}_j} \quad (3)$$

where T_j is the desired target (which reflects product type) value for the process output.

The simple EWMA control method as described above is adequate for controlling equipment with small disturbances and slow changes in output variation. However, the process output results a continuous error from the target if there is a consistent process drift. Therefore, a dEWMA control scheme must be employed. This controller involves two expressions, the first for estimating step-change $\hat{\alpha}_j(k)$ and the second for drift speed $\hat{p}_j(k)$. When the output is measured $\hat{\alpha}_j(k)$ and $\hat{p}_j(k)$ are computed as follows:

$$\begin{cases} \hat{\alpha}_j(k) &= \lambda_{1j}[y_j(k-1) - \hat{\beta}_j u_j(k-1)] + (1 - \lambda_{1j})[\hat{\alpha}_j(k-1) + \hat{p}_j(k-1)] \\ \hat{p}_j(k) &= \lambda_{2j}[y_j(k-1) - \hat{\beta}_j u_j(k-1) - \hat{\alpha}_j(k-1)] + (1 - \lambda_{2j})\hat{p}_j(k-1) \end{cases} \quad (4)$$

where λ_{1j} and $\lambda_{2j} \in [0, 1]$, are discounting factors used to minimize the influence of old data in favor of the new data.

Otherwise, when the system output is estimated, $\hat{\alpha}_j(k)$ and $\hat{p}_j(k)$ are computed by replacing the measured output $y_j(k-1)$ by its estimation $\hat{y}_j(k-1)$ in (4).

Then the control law for the next process run is calculated as follows:

$$\hat{u}_j(k+1) = \frac{T_j - \hat{\alpha}_j(k) - \hat{p}_j(k)}{\hat{\beta}_j} \quad (5)$$

3. Proposed Approach

As mentioned before, to deal with the lack of information, and for our industrial case, the JITL approach as described in [9], is used to estimate non measured DTM data ($y_j(k)$). Then, for the R2R controller, and due to the amount of disturbances in CMP process which we have dealt with, the dEWMA is used. The classical JITL approach, as described in [9], is based on a three-step VM module to the query data. First of all, the relevant data samples in the database are searched to match the query data by some nearest neighborhood criterion. Then, a local model is built based on the most similar data. Finally, the output estimation is calculated based on the local model and the current query data. Right after that, the local model is discarded. When the next estimation is required, a new local model will be built based on the three aforementioned steps.

By observing carefully the classical approach, some contributions can be integrated in the classical JITL version to cope with some drawbacks. In fact, the JITL approach described in [9] has at least three main drawbacks.

For the first step, when some nearest neighborhood criterion is evaluated, there is no fixed similarity threshold. In other words, "most relevant data samples" can be selected while these data are not truly similar to the query data at the run k . Therefore, these selected data are not considered as real neighbors to the query point observation. In this case, local VM model construction is based on dissimilar

observations. Consequently, this model cannot reflect the query point domain. Therefore, current DTM estimation will be biased. In order to face the first drawback, a threshold of similarity must be fixed. Then, data is considered as relevant only if it reaches this threshold.

For the second step of the classical JITL, and just after the threshold definition, it may be that a sufficient number of real neighbors can not be found to build the modeling database. In this case, it is impossible to build a local model. So, the idea is to define minimum number of real similar observations which allows to construct a local model. In this case, the estimation becomes possible.

A third weakness of the classical approach can be listed when using a static sampling measurement. Indeed, a physical measurement must be provided at each time the run of measurement is coming. However, a sufficient number of similar observations could be found before proceeding to physical measurement. Therefore, an estimation of the measured variable could be easily performed instead of physical measurement which increases the manufacturing cost.

The detailed description of the modified JITL approach with its interaction with R2R controller is described in the sequel.

III-A. JITL approach description

As mentioned before, the ETM variables (collected data from equipment), denoted by $x_j(k)$, are measured or collected during the entire processing task (or run). They have also a direct impact on wafer properties which are defined throughout DTM variables denoted by $y_j(k)$. The sequence of the approach is described as follows:

1. To initialize measurement and estimation counter

Measurement and estimation procedures need to be initialized. To do this, let's define two counters k_{est} and k_{mes} which define respectively the number of estimates and measures. At the beginning, both counters are set to zero.

2. To find similar observations

In this step, all anterior ETM variables of the complete database are extracted when the DTMs were measured, which gives a collection $[X_c, Y_c]$ of N_c data. Then, the current ETM $x_j(k)$ is compared to all x_c variables. Such a comparison is achieved by using the similarity criterion (6), defined by [9] as follows:

$$\begin{cases} s_i(x_j(k), x_c(i)) = \gamma \sqrt{e^{-d(x_j(k), x_c(i))^2}} + (1 - \gamma) \cos \theta_i \\ \text{with } \cos \theta_i = \frac{x_j(k)^T x_c(i)}{\|x_j(k)\| \|x_c(i)\|} \end{cases} \quad (6)$$

where $d(x_j(k), x_c(i))$ denotes the Euclidean distance between the current ETM data $x_j(k)$ and the i^{th} extracted variable $x_c(i)$ from $[X_c, Y_c]$. $\gamma \in [0, 1]$ is a weight parameter, i an index with $1 \leq i \leq N_c$ and θ_i is the angle between $x_j(k)$ and $x_c(i)$. Note that from (6), $s_i \in [0, 1]$ and s_i tends to unity when $x_j(k)$ is close to $x_c(i)$. Referring to [9], it is also important to note that $s_i(x_j(k), x_c(i))$ will not be computed if $\cos \theta_i$ is negative ($x_j(k)$ and $x_c(i)$ are not collinear and consequently are not similar).

Then, a threshold of similarity s_{min} is defined. This means that the observations $x_j(k)$ and $x_c(i)$ are considered as similar if the following inequality holds:

$$s_i(x_j(k), x_c(i)) \geq s_{min} \quad (7)$$

Then, M data of the database which satisfy the last condition (Eq.7) are selected and noted $[X^M, Y^M]$.

3. To test if the number of similar observations is sufficient

In this step, a minimum number (threshold) of real similar observations noted N_{min} is fixed. This allows

the construction of a local model if the number of nearest ETM neighbors M exceeds the predefined threshold N_{min} ($M \geq N_{min}$). In this case, the estimation becomes possible and k_{est} is incremented ($k_{est} = k_{est} + 1$). Otherwise, if $M \leq N_{min}$, the measurement procedure must be started instead of estimation and k_{mes} counter is incremented ($k_{mes} = k_{mes} + 1$). Physical measurement is then explored in R2R controller to provide control laws $u_j(k)$.

4. To build VM local model

Using $[X^M, Y^M]$ and MLR, the linear local model for each needed estimation can be modeled as follows [17]:

$$y_j^M(l) = \sum_{p=1}^{N_p} B_p x_j^M(l) \quad (8)$$

where $l \in \{1, \dots, M\}$, $x_j(l) = [x_{1,j}(l), \dots, x_{N_p,j}(l)]^T$ denotes the vector of ETM variables of the l^{th} run, N_j the number of ETM variables and β_j the regression coefficients.

Let's define $B_j = [B_{1,j}, \dots, B_{N_p,j}]^T$ the matrix of regression coefficients, $Y_j^M = [y_j^M(1), y_j^M(2) \dots y_j^M(M)]^T$ the matrix of M nearest DTM neighbors and $X_j^M = \begin{pmatrix} x_{1,j}^M(1) & \dots & x_{N_p,j}^M(1) \\ \vdots & & \vdots \\ x_{1,j}^M(M) & \dots & x_{N_p,j}^M(M) \end{pmatrix}$ the matrix of M nearest

ETM neighbors.

Let's define the estimation error $\varepsilon(k) = y_j(k) - \hat{y}_j(k) = y_j(k) - \hat{B}_j^T x_j(k)$

The main objective is to find the best parameter vector B_j such that the estimated output $\hat{y}_j(k)$ is close to the $y_j(k)$. It is well know that this objective is achieved by using the least square minimization [17], and the solution is given as follows:

$$\hat{B}_j = (X_j^{MT} X_j^M)^{-1} X_j^{MT} Y_j^M \quad (9)$$

5. To estimate DTM and to discard local model

Once the model is built, $\hat{y}_j(k)$ can be estimated as follows:

$$\hat{y}_j(k) = \hat{B}_j^T x_j(k) \quad (10)$$

where $x_j(k) = [x_{1,j}(k), \dots, x_{N_p,j}(k)]^T$ denotes vector of ETM variables of current run (when estimation is needed).

Once $\hat{y}_j(k)$ is estimated, the local model will be discard.

6. To integrate estimation in R2R control

Using dEWMA strategy and when the estimation $\hat{y}_j(k)$ is provided, the control law $\hat{u}_j(k+1)$ of the next run is calculated using (4) and (5). The whole description of the approach is illustrated by Figure 2.

III-B. Modified JITL approach highlights

Using the new version of the JITL, a new dynamic sampling measurement strategy is newly integrated. This strategy does not take into account the old static sampling measurement strategy already used in semiconductor manufacturing.

For the new proposed version of JITL approach, the sampling frequency of the measurement is not predefined in advance by the user. The number of physical measurements is conditioned by the number of products which cannot be estimated. The number of physical measurements will be reduced using the modified version of the approach. Therefore, more estimations will be provided. Consequently, R2R controller which will find less difficulty to update control laws. Then product quality will be well

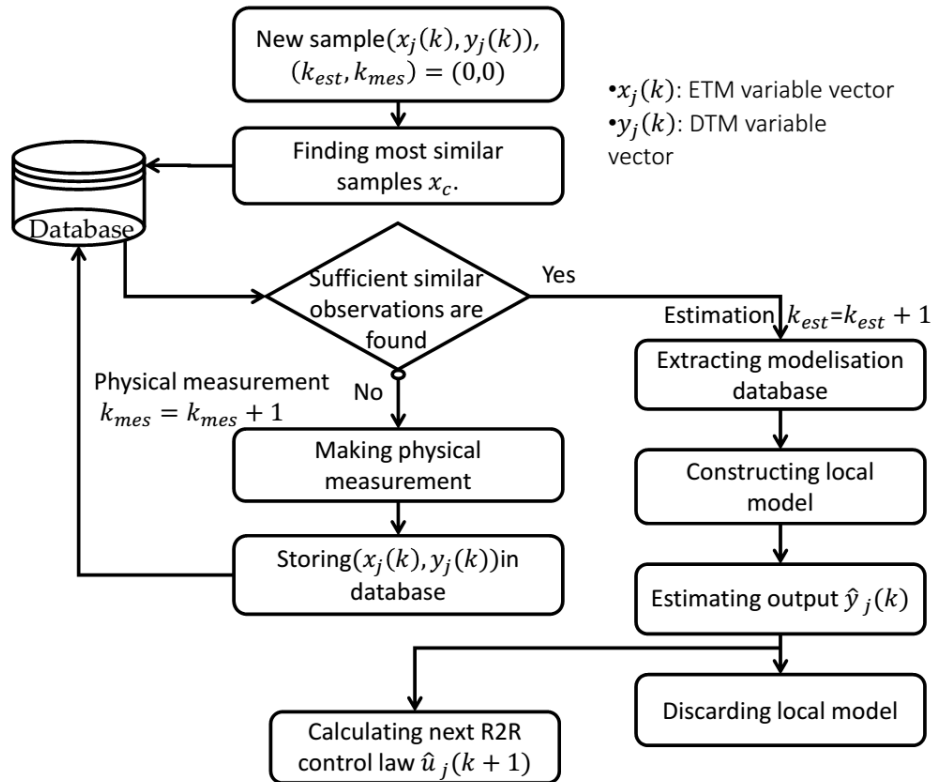


Figure 2. Modified JITL algorithm description

controlled. For example if the sampling measurement rate is prefixed and if accidentally for many runs of measurement (where physical measurement is required), the selected measured product is the same and does not cover all products. The estimation of measurements in other runs will not be possible or the estimation will be very biased.

Another point can be discussed. If physical measurements are forced at these runs, a redundancy of measurements could be made. These redundancy of measurements could be avoided using the modified JITL by improving estimations instead of physical measurements. Therefore, only needed products will be measured. This reduces measurement and production costs.

We note that it is important to include a threshold of non-measured products (threshold of successive estimation) k_{est}^{max} . When the number of successive estimation, k_{est}^{succ} reaches the threshold ($k_{est}^{succ} = k_{est}^{max}$), a physical measurement must be made to guarantee the estimation and the product quality. The next paragraph will be dedicated to a case study using industrial examples which will be used to compare results of the two approaches. The industrial examples will be dedicated also to show how R2R manifests with estimations provided by the second approach.

4. Case study

To test our approach and in collaboration with our industrial partner STMicroelectronics Rousset, 5 historical databases were extracted from 5 different CMP machines. In each equipment, many different products are processed (3 for most machines).

As a first step, a comparison was made between estimation results provided by the classical approach and the modified one.

The first approach is based on a static measurement sampling plan. For this reason, the sampling measurement rate must be defined. For CMP machines, we consider that the sampling rate is equal to 10, i.e. only one product is measured over 10 processed products.

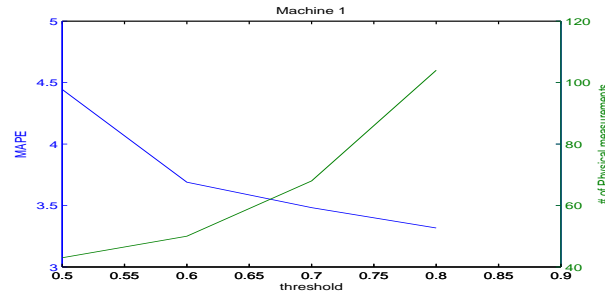


Figure 3. Threshold selection

As mentioned before, the number of real similar observations M must reach the threshold N_{min} . For our example, input vector $x_j(k)$ is made from 3 parameters. For this reason, $N_{min} = N_j = 3$, during second approach processing.

The true similar observations are obtained with the threshold that realize a compromise between the number of physical measurement and the Mean Absolute Percentage Error (MAPE). According to Figure 3, the optimum choice of the threshold can be obtained from the interval $[0.65, 0.7]$. In the present example, the value 0.7 is considered as optimum choice.

According to Figure 4, we note that the estimation of the layer thickness $\hat{y}_j(k)$, using the second approach has been improved qualitatively for the two first represented machines. It is noted that the obtained estimation of $\hat{y}_j(k)$ is filtered i.e. the peaks vanish and variability is reduced. In other words, the estimation error decreases.

Table 1 shows the quality estimation of DTM variables using both approaches, including the proposed new method and the classical approach of the algorithm JITL. One can remark clearly that the number of physical measurements (k_{mes}) is reduced to nearly third when using the second approach for all machines. For example, for the first machine, when the first approach is applied, the total number of estimations is 1331 among 1502. This means that the operation required 171 physical measures to estimate 1331 outputs. However, using the modified version of the approach shows that the number of estimations increases and reaches 1434 among 1502. This means that the operation needed only 68 physical measurements to estimate 1482 outputs. Here, the second approach requires less physical measurements than the first one and contributes in the decrease of the production cost. This intake is also valid for the other machines. This also allows to show the interest of integrating dynamic sampling using the proposed new version of JITL.

Table 1. Results comparison between the classical approach and the modified approach

Machine	Approach	k_{mes}	MAPE% ⁽¹⁾	σ_r ⁽²⁾
1	Classical JITL	171	9.49	0.2
	Modified JITL	68	3.5	0.045
2	Classical JITL	157	12.4	1.43
	Modified JITL	55	4.38	0.046
3	Classical JITL	167	7.3	0.3373
	Modified JITL	69	3.8	0.045
4	Classical JITL	164	6.5	0.18
	Modified JITL	62	3.2	0.04
5	Classical JITL	155	7.4	0.13
	Modified JITL	59	3.5	0.044

$$^{(1)} MAPE\% = 100 \frac{1}{n} \sum_{k=1}^n \frac{|\hat{y}(k) - y(k)|}{|y(k)|}, \quad ^{(2)} \sigma_r = \sqrt{\frac{1}{n-1} \sum_{k=1}^n \frac{|\hat{y}(k) - y(k)|^2}{|y(k)|^2}}$$

Referring also to Table 1, we note that MAPE (the Mean Absolute Percentage Error) undergoes a slight decrease for all machines. The MAPE is between two times and three times smaller using the modified JITL approach.

The relative standard deviation σ_r of error of estimation also decreases to nearly a third for all machines. So, it can be concluded that the proposed approach contributes to reducing the number of physical measurements by introducing a dynamic sampling measurement plan which contributes in the diminution of the cost of production, and ensures better estimation quality compared to the ones given by the classical JITL.

Table 2. Comparison the MAPE of static error using R2R based on VM and using industrial R2R

Product	R2R type	MAPE %
1	Industrial R2R controller	2.35
	R2R controller based on VM	0.37
2	Industrial R2R controller	1.13
	R2R controller based on VM	0.3
3	Industrial R2R controller	1.6
	R2R controller based on VM	1.4

The second step of the test, consists of showing how the VM based on modified JITL interacts with R2R controller. To highlight the contribution of the estimated outputs, static errors of the R2R controller based on VM and that based on industrial data are compared. Without loss of generality, the comparison is made only for the first machine. In this one, 3 product types are processed.

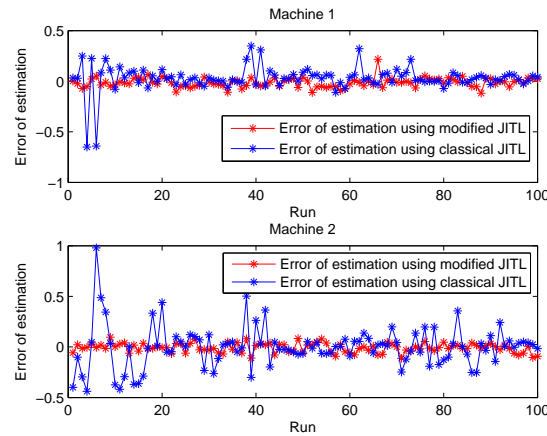


Figure 4. Comparison of error of estimation performance using the two approaches

Referring to Figure 5 and Table 2, we note that static error is improved using R2R controller based on VM for the three products and most noteworthy for the two first ones. In fact, by investigating the MAPE, this latter is 6 and 4 times less for the first and the second products using R2R controller based on VM than using the industrial R2R controller.

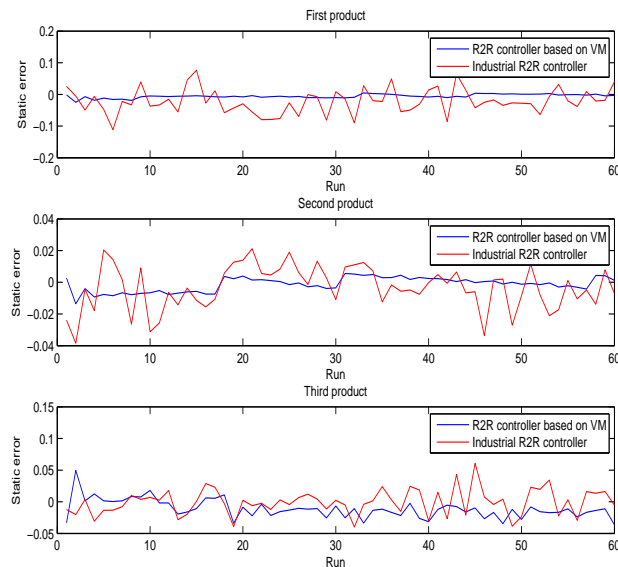


Figure 5. Comparison of static error using R2R based on VM and using industrial R2R

5. Conclusion

In this paper, VM module has been used to estimate non-measured DTM variables (layer thickness) based on current ETM variables. This module is calculated using anterior measured DTM with their corresponding ETM data. We have opted for that a new proposed version of JITL approach which allows to overcome the problem related to static sampling measurement, and the drawbacks of classical JITL version. We have also introduced the interaction and the interest of VM in R2R controller. Finally, a

comparison between the classical JITL approach and the modified one is made on real industrial data-sets derived from CMP machines and delivered by our industrial partner. It has been shown that the contribution of the proposed approach is manifested by its improved results, especially by defining a new dynamic sampling measurement plan reduces the number of physical measurements.

Without loss of generality, we focused the work only on the first machine to see how VM estimations can react with R2R controller. We noted that quality product is improved using R2R controller based on VM.

In partnership with our industrial leader, future works will concern the application of the proposed approach to deal with missing data for the case of interconnected control loops.

References

- [1] Lynn, S., MacGearailt, N., and Ringwood, J.V. (2012). Real-time virtual metrology and control for plasma etch. *Journal of Process Control*, 22(4), 666-676.
- [2] Moyne, J., Del Castillo, E., and Hurwitz, A.M. (2000). *Run-to-run control in semiconductor manufacturing*. CRC press.
- [3] Maas, J., Ebert, M., Bhattacharyya, K., Cramer, H., Becht, A., Keij, S., Plug, R., Fuchs, A., Kubis, M., Hoogenboom, T., (2011). Yieldstar: a new metrology platform for advanced lithography control. In 27th European Mask and Lithography Conference, 79850H-79850H. International Society for Optics and Photonics.
- [4] Gill, B.S. (2011). Development of virtual metrology in semiconductor manufacturing. Ph.D. thesis, The University of Texas at Austin.
- [5] Elliott, R.C., Nurani, R.K., Gudmundsson, D., Preil, M., Nasongkhla, R., and Shanthikumar, J.G. (1999). Critical dimension sample planning for sub-0.25 micron processes. In *Advanced Semiconductor Manufacturing Conference and Workshop*, 1999 IEEE/SEMI, 139-142. IEEE
- [6] Lynn, S. (2011). Virtual metrology for plasma etch processes. Ph.D. thesis, National University of Ireland, Maynooth.
- [7] Khan, A.A., Moyne, J.R., and Tilbury, D.M. (2008). Virtual metrology and feedback control for semiconductor manufacturing processes using recursive partial least squares. *Journal of Process Control*, 18(10), 961-974.
- [8] Slišković, D., Grbić, R. and Hocenski, Ž (2012). Methods for plant data-based process modeling in soft-sensor development. *AUTOMATIKA: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije* 52(4), 306-318.
- [9] Cheng, C. and Chiu, M.S. (2004). A new data-based methodology for nonlinear process modeling. *Chemical Engineering Science*, 59(13), 2801-2810.
- [10] Abdi, H. (2010). Partial least squares regression and projection on latent structure regression (PLS regression). *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(1), 97-106.
- [11] Pan, J.C.H. and Tai, D.H. (2009). Implementing virtual metrology for in-line quality control in semiconductor manufacturing. *International Journal of Systems Science*, 40(5), 461-470.
- [12] Purwins, H., Barak, B., Nagi, A., Engel, R., Hockele, U., Kyek, A., Cherla, S., Lenz, B., Pfeifer, G., and Weinzierl, K. (2014). Regression methods for virtual metrology of layer thickness in chemical vapor deposition. *IEEE/ASME Transactions on Mechatronics*, 19(1), 1-8.
- [13] Hung, M.H., Lin, T.H., Cheng, F.T., and Lin, R.C. (2007). A novel virtual metrology scheme for predicting CVD thickness in semiconductor manufacturing. *IEEE/ASME Transactions on Mechatronics*, 12(3), 308-316.
- [14] Su, Y.C., Lin, T.H., Cheng, F.T., and Wu, W.M. (2008). Accuracy and real-time considerations for implementing various virtual metrology algorithms. *IEEE Transactions on Semiconductor Manufacturing*, 21(3), 426-434.
- [15] Besnard, J., Gleispach, D., Gris, H., Ferreira, A., Roussy, A., Kernafl, C., and Hayderer, G. (2012). Virtual metrology modeling for cvd film thickness. *International Journal of Control Science and Engineering*, 2(3), 26-33.
- [16] Kang, P., Kim, D., Lee, H. J., Doh, S., and Cho, S. (2011). Virtual metrology for run-to-run control in semiconductor manufacturing. *Expert Systems with Applications*, 38(3), 2508-2522.
- [17] Ghasemi, J., Saaidpour, S., and Brown, S.D. (2007). Qspr study for estimation of acidity constants of some aromatic acids derivatives using multiple linear regression MLR analysis. *Journal of Molecular Structure: THEOCHEM*, 805(1), 27-32.