Modelling and decision support system for intelligent manufacturing: An empirical study for feedforward-feedback learning-based run-to-run controller for semiconductor dry-etching...

Article *in* The International Journal of Industrial Engineering: Theory, Applications and Practice · January 2018

Marzieh Khakifirooz
Tecnológico de Monterrey

Michel Fathi
University of North Texas

International Journal of Industrial Engineering, 25(6), 828-842, 2018

MODELLING AND DECISION SUPPORT SYSTEM FOR INTELLIGENT MANUFACTURING: AN EMPIRICAL STUDY FOR FEEDFORWARD-FEEDBACK LEARNING-BASED RUN-TO-RUN CONTROLLER FOR SEMICONDUCTOR DRY-ETCHING PROCESS

Marzieh Khakifirooz[1], Mahdi Fathi[2], and Chen-Fu Chien[1, *]

[1]Department of Industrial Engineering and Engineering Management, National Tsing Hua University
Hsinchu, Taiwan
[*]Corresponding author's e-mail: cfchien@mx.nthu.edu.tw

[2]Department of Industrial and Systems Engineering, University of Florida
Gainesville, FL, United States

Shrinkage in semiconductor devices affects the process window of all wafer fabrication steps including plasma etching. Drifts or shifts are most significant effects on the etching process due to shrinkage in semiconductor devices. Any drift or shift affects on critical dimensions (CD) of the wafer and changes the thickness and the width over time. Therefore, there would be an essential need for estimation and minimization of CD variation on a wafer-to-wafer basis by optimization techniques. This study aims to design a learning-based control system for monitoring the CD in Dry-Etching process. Feedforward-feedback control technique is used to reduce CD variation. Among all learning-based control systems, the Iterative Learning Control (ILC) integrated with Virtual Meteorology (VM) data, as a well-known system which can involve both feedforward signal from the past events, and feedback signals from the output of the current event is used to learn the behavior of the system and enhance the performance of the controller run-by-run. The proposed control model is optimized by gradient learning approach. The result is validated through the simulated study manipulated from empirical data and shows the advantage of the proposed feedforward-feedback learning controller than the common run-to-run exponentially weighted moving average (EWMA) control design.

Keywords: critical dimension; disturbance rejection controller; feedforward-feedback control system; iterative learning controller; decision support system; virtual metrology

(Received on September 15; Accepted November 05, 2018)

1. INTRODUCTION

Dry etching (DE) is one of the critical wafer fabrication processes for semiconductor devices to remove selected layers of photoresist materials and maintain the quality of critical dimension (CD). DE technology as a core part of semiconductor device fabrication, cannot avoid many difficulties to cope with the challenges associated with new circumstances regarding the future development directions of semiconductor manufacturing. In general, DE suffers from the following challenges:

- Rapid development of lithography technology from double patterning technology to quadruple patterning technology, which is required tighter control law for global and local distribution control of CD (Turkot *et al.*, 2017).
- Tremendous increase of difficulties for High Aspect Ratio Contacts (HARC) etching such as DRAM capacitor, VNAND channel hole etching, and Metal Contact (MC) etching (Tandou *et al.*, 2016).
- More demand on lot-to-lot, wafer-to-wafer, and tool-to-tool or even die-to-die process variation management in recent years (Chien *et al.,* 2016).
- Request for innovative approaches to achieve goals of the etching process while utilizing new material/new structure (Ghodssi and Lin, 2011).

Specifications for DE are typically a tradeoff among rate, directionality, uniformity, selectivity, pattern dependencies, damage, and cleaning of etching process (Sawin, 1994). These requirements are contradictory, and typically resulting in the loss of CD control process because of photoresist erosion or lifetime of etching tools. Therefore, designing an excellent CD control system is one of the essentials to achieve high etching rates.

This study aims to propose an optimal control system for monitoring CD after the etching process. The CD is measured by metrology tool and is compensated by modifying the setup parameters of the controller or replacing the etching tools before their life cycle. A feedback message is sent to the next run for pre-adjustment, and a feedforward message is transmitted from the previous chemical-mechanical polishing (CMP) for post-adjustment. Then, the input recipe is updated for the next run based on recently measured process data through the metrology tools.

In this paper, we drive a learning control design technique based on the frequency of measurement data that integrates feedforward and feedback control on the etch process. To the best of our knowledge, there is only limited studies investigated on the Advanced Process Control (APC) system for CD of DE process. Some significant contributions in controlling CD is summarized in Table 1. Most of the developed model from literature are based on the run-to-run (R2R) controller which has a fixed profile structure and is not suitable for CD of DE process. We will broadly discuss this phenomenon in Section 2.

Based on both feedback and feedforward control signal and the mixture characteristics of control variables (i.e., lifetime, width, depth) plus the high-mixed production plan and short lifetime of etching tools, we need an advanced learning control system to be adapted with dynamic structure of DE process. Hereupon these needs, we design a novel feedforward-feedback learning R2R control system, to support all obligations in controlling CD during DE process.
The proposed approach is developed in a way that can be utilize the feedback information from metrology step combined with feedforward information of the previous CMP step. In addition, to overcome the challenges of metrology delay and shortage of data, the proposed control system is designed based on the learning-based control process and frequent measurement system (iteration). The applications of learning-based controller can cover a wide range of operations in the semiconductor industry such as chemical vapor deposition (CVD) process (Xu *et al.*, 1999; Chen *et al.*, 2011), batch processing (Kim *et al.*, 2013a; Kim *et al.*, 2013b, and temperature uniformity control (Won *et al.*, 2017). The learning-based control design in this study is inspired by Iterative Learning Control (ILC) system for R2R control design where iteration data is generated by Virtual Metrology (VM) tool.

The remainder of this paper is organized as follows: Section 2 introduces the core challenges, construction of problems, and intellectual foundations of investigating DE fabrication processes in this study. Section 3 presents the core structure of learning-based controller and the augmented feedforward-feedback R2R for controlling the CD. Section 4 demonstrates the case study of manufacturing data. Finally, the paper will be concluded in Section 5.

Table 1. Related studies on advanced process control of CD of DE process

Reference	Control System	Control Objective	Main Contribution
El Chemali *et al.* (2003)	feedforward and feedback R2R	Kalman filtering	minimizing and modeling etch-rate disturbance using a model of relationship between toll-etching life time and CD
El Chemali *et al.* (2004)	feedback R2R	Kalman filtering	manipulating the inputs and estimate disturbances; control sidewall angle
Mao *et al.* (2007)	multiple-dimensional closed-loop feedback R2R	EWMA; Kalman filtering	analyzing the effect of model mismatch and the controller's sensitivity to unknown noise
Wu *et al.* (2008)	R2R	Nonlinear Multiple Exponential-Weight Moving-Average (NMEWMA); Dynamic Model-Tuning Minimum-Variance (DMTMV)	modeling the relationship between exposure dose and focus and CD
Yang *et al.* (2010)	R2R	-	monitoring photoresist parameters such as photoresist thickness, photoactive compound and CD in-situ and in real-time
Ngo *et al.* (2013)	Linear Model Predictive (LMP)	EWMA; Kalman filtering	manipulating etch time to estimate state variables
Chien *et al.* (2015)	feedforward R2R	Analysis of Variance (ANOVA)	determining tool affinity
Hsu and Wu (2016)	R2R	error-smoothing EWMA	compensating the process variation

2. PROBLEM DEFINITION AND IDENTIFICATION

2.1. Terminologies and Notations

The notation and terminologies used in this study are listed as follows:

j	The iteration index.
t	The process run index, $t \geq 1$.
T	The total number of instance at each iteration.
$\boldsymbol{u}_j(t)$	Vector of input variables for iteration j at run t.
$\boldsymbol{y}_j(t)$	Vector of process outputs for iteration j at run t.
$\boldsymbol{y}_d(t)$	Vector of desired process outputs at run t.
$d(t)$	Process disturbance at run t.
$\boldsymbol{e}_j(t)$	Vector of deviation from the desired output for iteration j at run t.
J_t	Cost function at run t.
FF	Difference between the output and target of pre-layer.
RF	Etching tool lifetime.
C_{FF}	Coefficient of linear regression between intercept and output of pre-layer.
C_{RF}	Coefficient of linear regression between intercept and tool lifetime.
$P(q)$	Rational function of learning-based controller.
$Q(q)$	Q-filter of learning-based controller.
$L(q)$	Learning function of learning-based controller.
q	Forward time-shift operator.

2.2. Semiconductor DE Process

The etching process removes material from areas identified by the lithography process, to create structures for functional use (see Figure 1). Etching is a critically important process which every wafer needs to undergo this step many times before its completion. There are two main types of etching, wet-etching (liquid-based etchants) and DE (plasma-based etchants).

In wet-etching, the wafers are immersed in a tank of chemical solution or etchant. In wet etching, the photoresist material is removed through chemical reaction between wafer surface and etchants. DE is the removal process of material in the absence of the solvent. In DE process, the etching materials such as gases or plasma remove the substrate layer by the physical reaction. The wet-etching has some limitations in its applicability including the large size of patterning, isotropic process (Ivanov, 2017), hazardous of chemical materials, long completion time, and combination with a subsequent rinse (Jörg *et al.*, 2018). Therefore, DE technique is more applicable in the wafer fabrication process than wet-etching.

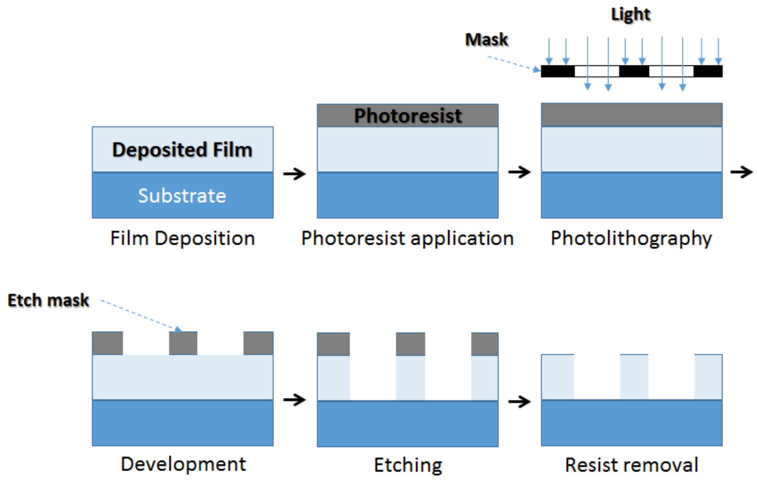

Figure 1. The etching process for wafer fabrication

Any etch process is characterized by certain properties and quality measurements as follows:

- Etching rate: The amount of material removed from the wafer over a defined period of time.
- Selectivity: The ability of the etch process to distinguish between the photoresist layer and substrate to be carved and the material not to be carved.
- Feature profile: Isotropic, etching proceeds at equal rates in both horizontal and vertical directions; Anisotropic, etching flows faster in one plane than in another.
- CD: dimensions of the delicate patterns formed on a wafer, i.e., width and depth.
- Residue: Remaining polymer after a post-etching cleaning process.
- Thickness: The photoresist thickness after the post-etching process.

In DE process, usually due to the lifetime of etching tools and mixed-product production process, it is challenging to obtain sufficient historical data as the reference information for controlling the production process. Therefore, the control system of DE process should design in a way that can be learned during short life-cycle of etching tools. In this study, among all properties, CD as the key characteristics of the etching process is selected for further investigations in designing the learning-based control system for DE.

2.3. Control System of CD

The CD is one of the important quality characteristics for wafer fabrication that its toleration should be continuously controlled, and keep it tight for yield enhancement. In semiconductor fabrication device, there is two different source of CD, the CD of scanned pattern via photolithography processes or photo-CD (PCD), and the CD measured by metrology tools after the etching process or etch-CD (ECD). The ECD represents the final line-width of fabricated patterns of each layer on the wafer. Thus, the ECD is the target for process control. The fundamental objective of the control process for ECD is minimizing the ECD's variation. Nevertheless, the ECD's variation is affected by the variability of both etching and photolithography processes. Therefore, the process control in etching should reduce the cumulative process variation from photolithography to the end of etching time. Controlling ECD is an inter-process R2R control (Chien *et al.*, 2015; Qin *et al.*, 2006) (see Figure 2), which requires the information from the past production step as well as measurement information from metrology tools to build a sufficient control system.

In the rest of this study for simplicity, the general CD refers to the ECD.

Showing in Figure 2, the inter-process R2R control deals with the process control of two or more inter-related process modules. In practice, although CVD and CMP are affecting the thin film thickness, the CMP processes can affect the control strategy of the CD. In particular, after the CMP process removes the unwanted photoresist materials, process engineers can report the photoresist thickness. The CMP measurement will be applied as a feedforward controller to the DE process. CD will be measured during the metrology step and will send feedback to the etching process (see Figure 3). As each wafer undergoes many times under DE for completing the fabrication process, therefore, the control purpose of CD is to minimize the effect of unmeasurable cumulative disturbance on the CD from the first layer of DE.

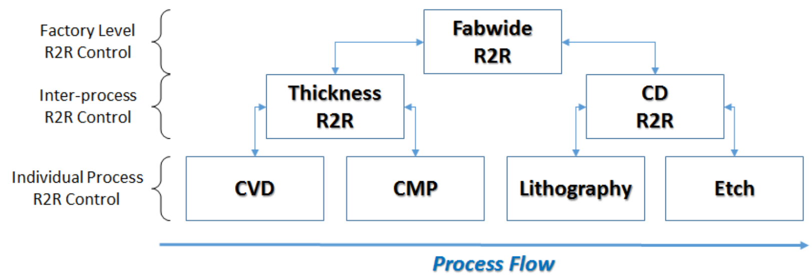

Figure 2. Classification the level of control process during the process flow of wafer fabrication (Chien *et al.*, 2015; Qin *et al.*, 2006)

In Figure 3, consider the system design of a controller which only contains the feedback signal from the metrology tools. Therefore, the linear dependency simply holds for input and output, and the relationship is usually known and formulated as:

$$\boldsymbol{y}(t) = slope * \boldsymbol{u}(t) + intercept \tag{1}$$

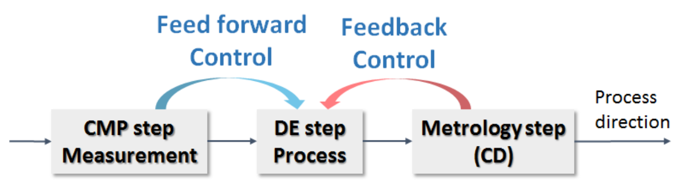

Figure 3. The process flow in system design of controller for CD

where the input $\boldsymbol{u}(t)$ is the etching time measuring the time of chemical or physical reaction to remove the photoresist material, the output $\boldsymbol{y}(t)$ is CD such as depth and width, and $intercept$ is cumulative disturbance and uncertainties. However, the process variation usually includes the disturbance ($d(t)$) from the effects of DE of previous photoresist layers (in short called pre-layer effect) or the etching tool lifetime in radio frequency (RF) hours. In this situation, both effects of pre-layer disturbance and tool lifetime can contribute into the "intercept". The effect of pre-layer disturbance could be reported from CMP measurement called the feedforward signal (see Figure 3). Therefore, the intercept in (1) is divided into two portions

$$intercept = C_{FF} * FF_{d(t)} + C_{RF} * RF_{time} + intercept' \tag{2}$$

and (1) is updated as follow:

$$\boldsymbol{y}(t) = slope * \boldsymbol{u}(t) + \left(C_{FF} * FF_{d(t)} + C_{RF} * RF_{time} + intercept'\right) \tag{3}$$

where C_{FF} is the coefficient of linear regression between intercept and output of pre-layer, $FF_{d(t)}$ is the difference between the output of pre-layer and target of pre-layer (or the disturbance from pre-layer), and C_{RF} is the coefficient of linear regression between intercept and tool lifetime. In practice, C_{FF} and C_{RF} are fixed, and if process engineers detect any change in CD, the coefficient parameters in the model will update to their optimal setting. This study aims to design a control system for controlling both feedback and feedforward signals to avoid any significant changes in the system and optimize the coefficient parameters without relying on the expert knowledge.

The sparse behavior of semiconductor manufacturing data (one observation per run), makes the R2R controller as the most applicable and suitable control design for this industry. However, for some process like DE, which is engaged with a high level of dynamicity the regular R2R controller is not efficient. On the other hand, in general design of R2R, first, input profiles for t-th run, named $\boldsymbol{u}(t)$, then the conventional R2R is used to update output $\boldsymbol{y}(t)$. Therefore, $\boldsymbol{y}(t)$ has no profile, while $\boldsymbol{u}(t)$ has a fixed profile structure. The fixed profile structure of R2R controller does not allow a different structure for input and output signal. In another word, the R2R controller is designed for the system when both input and output are describing the same characteristics. Therefore, regarding to the linear relationship between input and output of controller for

CD of DE as described in (3), when output is representing the CD and input is a function of etching time, etching tool's lifetime, and disturbance from pre-layer, obviously R2R controller is not the best choice for DE process.

In this study, we propose a learning-based control design for CD of DE to deal with the sparsity of semiconductor manufacturing data with a varying profile structure.

3. OPTIMIZATION BASED LEARNING IN CONTROL SYSTEM

The performance of a system which repeats a task multiple times can be improved through learning procedure from previous iterations. As the production repeats cyclically, at each loop/cycle/run the optimal decision is made and becomes the initial setting for the next loop/cycle/run. The control system can learn from this cyclic repetition and iteratively improve the performance accuracy. Learning control can deal with the problem of synthesizing an appropriate control input to make the system produce the desired action by repeated trails even with incomplete knowledge. In this study, we used the advantages of learning by repeating to support control system of the CD during the DE process.

There are many extensions of applying learning terminology in control theory which some of its advantages can be briefly classified as follows (Antsaklis, 2001):

1. To learn about the plant; how to derive new plant models and to learn how to incorporate changes.
2. To learn about the environment; this can be done using methods ranging from passive observation to active experimentation.
3. To learn about the controller; learn how to adjust specific control parameters to enhance performance.
4. To learn new design goals and constraints.

There is a variety of control strategy can be learned by historical information to design a new control system. In particular, the Iterative Learning Control (ILC) (Chen *et al.*, 2012), R2R control (Moyne *et al.*, 2000; Chien *et al.*, 2014), Adaptive Control (AC) (Åström and Wittenmark, 2013), Neural Networks (NN) (Hunt *et al.*, 1992), and Repetitive Control (RC) (Steinbuch, 2002) are commonly used methodology. Table 2, summarizes the applications and characteristics of the aforementioned learning-based control model. In this study, we design a hybrid learning-based feedforward-feedback control system to optimize C_{FF} in (3) based on the problem definition in Section 2. The proposed control system can compensate limitations of R2R controller with the following properties:

1. Automatic coefficient optimization of disturbance.
2. Robustness improvement through the use of causal feedback of metrology tools and feedforward of data from previous CMP step.
3. Iterative learning procedure to deal with lack of historical information for learning.

Table 2. Comparison key characteristics and objectives of ILC, R2R control, Adaptive Control, Neural Networks, and Repetitive Control system

Control System	Key Characteristics	Advantage	Limitation
ILC (Bristow *et al.*, 2006)	• time-based function • can be applied for dynamic batch processing • is built upon feedback and feedforward controller • can be built based on state space model • modifies the control input/signals • intended for discontinues operation • the initial setting is fixed for entire of process	• input has varying profile • output has frequent measurement • feedforward control can eliminate the lag in the transient tracking of feedback control • does not need the distribution of repeating disturbances • highly robust to system uncertainties	• has closed loop structure • friction, unmodeled nonlinear behavior, and disturbances can limit the effectiveness of feedforward control • cannot provide perfect tracking in every situation

R2R (Tan *et al.*, 2015)	• time-based function • can be applied for static model • can be built upon on state space model	• has close loop and open loop structure	• only a single product is manufactured on a single tool • input has fixed profile structure • output has sparse measurement
RC (Wang *et al.*, 2009)	• time-based function • can be applied for dynamic continuous process with periodic input • can be built upon transfer function • the initial setting is based on last trail	• input has varying profile • output has frequent measurement	• has single close loop structure • intended for continues operation
NN	• can be applied for nonlinear network • modifies the control parameters	• solve problems that do not have an algorithmic solution	• requires extensive training data • convergence rate is slow
AC	• modifies the control system	-	• does not use the historical data

3.1. Framework of an Intelligent Control System

To design a robust control system based on feedforward-feedback learning-based structure an intelligent control system is demanded. The framework of an intelligent control system is partitioned into three main parts: plant, data management center, and optimal controller where all three components continuously are connected to decision support system. The schematic of an intelligent control system is illustrated in Figure 4.

In the first part or production plant, the information is produced, and collected, then process engineers apply the decision rules. The entire of demanded information including metrology data, control parameters, scheduling and recipe information, and environmental factors are collected in this part and then stored in the data management center for further investigation.

Data management center is a vital part for statistical process control (i.e., fault detection, recipe management, and yield enhancement). In fact, the data management center is a feeding part of decision support system. The whole information, from raw data, decision rules or control law could be restored in the data management center.

The control system plays the analyzer role. Regardless the structure of the control system, all controllers are using the information from data management center, and again sending the control law and predicted information to there, where the decision support system can make the decision rules for plant performance enhancement.

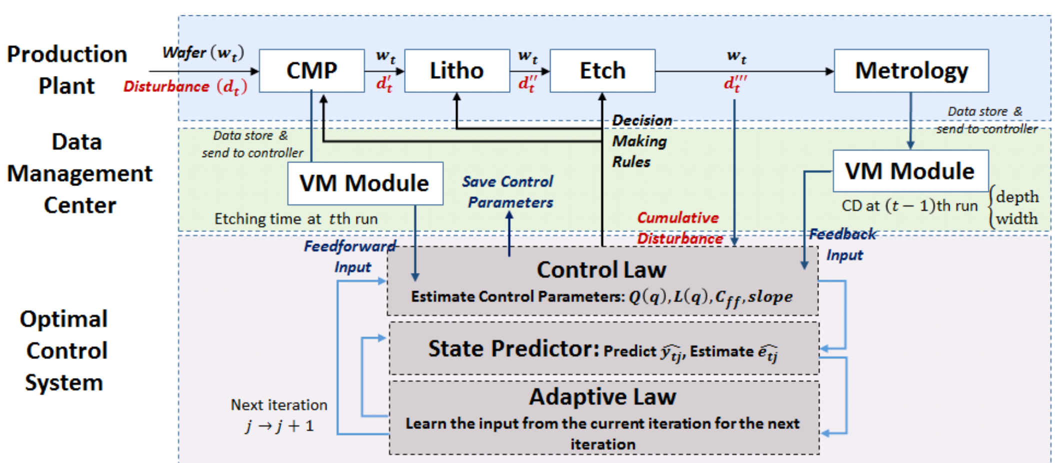

Figure 4. The framework of feedforward-feedback control system equipped with VM.

3.2. Feedforward-Feedback Learning-Based Control System

Consider the regular R2R controller which is equipped with only a feedback signal, as the most widely applied control system in semiconductor manufacturing. In controlling the CD of DE process, the regular feedback R2R controller cannot attain precisely near zero error due to time dependency of tool lifetime, tool heath, and cumulative disturbance that may transfer from pre-layers (Chien and Hsu, 2011; Chien *et al.*, 2014; Yu *et al.*, 2014).

To adapt the R2R controller with dynamic change in the system, there should be multiple measurements, however, the sparsely sampled output measurement from metrology tools can't support the dynamic change. One solution to overcome this weakness is to use the advantage of VM (Kang *et al.*, 2011; Tsai *et al.*, 2013; Baseman *et al.*, 2016; Jebri *et al.*, 2017) in the control system. The role of VM is to produce data for iterative control, therefore, the R2R controller can learn how to compensate the cumulative disturbance through iteration.

To adopt the VM into R2R controller we conduct the "just in time learning" approach (Cheng and Chiu, 2004) as the following steps:

1. At t-th run, virtual data is built upon the historical measured data through k-mean clustering.
2. Iteration is run using the data in the same cluster with the t-th measured data.
3. The process is separately conducted for both feedforward and feedback signals.

The proposed controller in Figure 4 is a controller that due to learning procedure can produce zero tracking error during repetitions of a command or eliminate the effects of a repeating disturbance on a control system output. Therefore, VM technology can be emerged into the R2R controller as a powerful tool to obtain models of imperfection and noisy data with a high degree of interpretability.

The next question to design a powerful control system for CD of DE process is how to eliminate the uncertainty using past performance information on the current trial (or how to bring the feedforward signal from CMP step into R2R control design)? The answer to this question can be given if the R2R controller will be formulated in a time-domain format.

Assume the discrete-time, linear time-invariant (LTI), Single-Input Single-Output (SISO) system as the principal structure for our proposed controller as follow:

$$\boldsymbol{y}_j(t) = P(q)\boldsymbol{u}_j(t) + d(t) \tag{4}$$

where q is the forward time-shift operator which means for any input signal at $(t+1)$-th run it can defined by the input signal at t-th run by $q.u(t) \equiv u(t + 1)$ and the plant $P(q)$ is a proper rational function of q and has a delay, or equivalently relative degree of 1. We assume that $P(q)$ is asymptotically stable. Repeating disturbances (Boeren *et al.*, 2016), repeated nonzero initial conditions (Gal and bars, 2013), and systems augmented with feedforward-feedback control (Kuo, 2002) can be captured in $d(t)$. Therefore, with regards the desired system output $\boldsymbol{y}_d(t)$ at t-th run, the system performance or error signal can be defined as follows:

$$\boldsymbol{e}_j(t) = \boldsymbol{y}_d(t) - \boldsymbol{y}_j(t) \tag{5}$$

$$\boldsymbol{u}_j(t) = Q(q)\boldsymbol{u}_{j-1}(t) + L(q)\boldsymbol{e}_j(t) \tag{6}$$

where $Q(q) \in (0,1)$ is Q-filter (transforming the feedforward information) and $L(q)$ is the learning function (updating law). The dynamic control system with plant dynamics in (4) and learning dynamics in (6) are shown in Figure 5. By the definition of (4) and (6), the R2R controller is changed to the form of ILC.

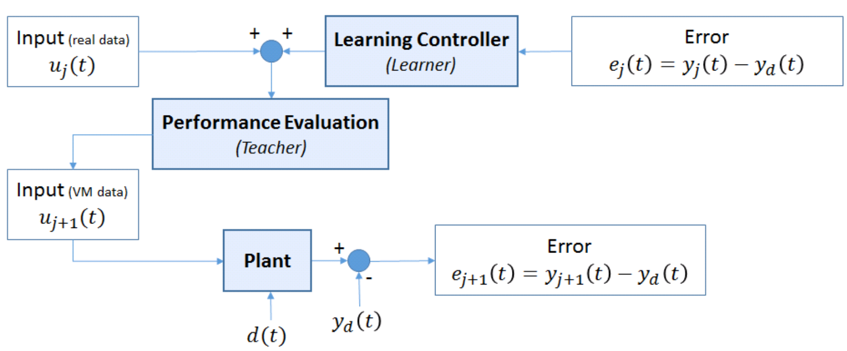

Figure 5. The block diagram of learning procedure of feedforward-feedback controller for one iteration.

3.3. Optimization the Control System for CD

The augmented feedforward-feedback learning-based controller similar to the regression model in (3) is formed into two parts:

1. A linear regression between CD and etching time.
2. The model of disturbances.

Therefore, the objective function of control system with regards to (3), (4), and (6) and the definition of error signal in (5) is:

$$\begin{aligned} &\min_{u(t)} J_t = \left|\boldsymbol{e}_j(t)\right|^2 \\ &\boldsymbol{e}_j(t) = \boldsymbol{y}_d(t) - \boldsymbol{y}_j(t) \\ &\boldsymbol{y}_j(t) = slope * \boldsymbol{u}_j(t) + \ d(t) \\ &\boldsymbol{u}_j(t) = Q(q)\boldsymbol{u}_{j-1}(t) + L(q)\boldsymbol{e}_j(t-1) \end{aligned} \tag{7}$$

where $d(t) = C_{FF} * FF_{d(t)} + C_{RF} * RF_{time}$.

Deriving the decision variables $slope$, C_{FF}, $Q(q)$ and $L(q)$ is necessary to solve the optimization problem in (7). There is infinite possible iterative procedures to solve the optimization problem in (7). The gradient approach (Owens *et al.,* 2009) has the most straightforward form and has been widely investigated in literature for optimizing the learning-based error (Amann *et al.,* 1996).

4. EMPIRICAL STUDY

Follow modeling a predictive control design for minimizing the variation of CD, the next step is to implement the proposed controller with process data and under unmeasurable disturbance. The primary objective of the controller is to regulate the CDs in the face of all sources of uncertainties, in which the difference between actual output and desired output will be minimized. Accordingly, if the control model can reject the effect of uncertainties, then the actual output and the input should be very close to each other.

In addition, for any control system design, it is essential to understand the system configuration, calibration, and initialization before system assembles in the real plant. In the case of learning-based feedforward-feedback control system, we should understand how the learning process can reject the disturbance and how the learning rate can transfer the disturbance-free output at the current run to the input at the next run. Furthermore, for the DE process, we would like to understand the effect of various process parameters on CD from the quality of photolithography or CMP process.

To estimate the validity of the proposed control system and the effect of learning in feedforward-feedback control design, we implement a simulation study for 200 lots of manufacturing data. To design the simulation scenario, we firstly collect empirical data, and the density plot of real data is illustrated in Figure 6. Manufacturing data in this study are accumulated the effect of etching tools' lifetime and etching time together. Therefore, equation (7) updates as follow:

$$\begin{aligned} &\min_{u(t)} J_t = \left|\boldsymbol{e}_j(t)\right|^2 \\ &\boldsymbol{e}_j(t) = \boldsymbol{y}_d(t) - \boldsymbol{y}_j(t) \\ &\boldsymbol{y}_j(t) = slope * \boldsymbol{u}_j(t) + \ C_{FF} * FF_{d(t)} \\ &\boldsymbol{u}_j(t) = Q(q)\boldsymbol{u}_{j-1}(t) + L(q)\boldsymbol{e}_j(t-1) \end{aligned} \tag{8}$$

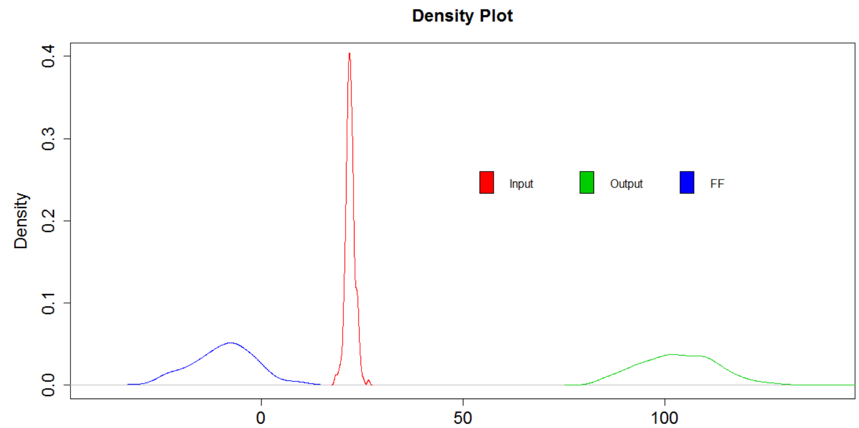

Figure 6. The density plot of empirical $\boldsymbol{u}(t)$ (red line), $\boldsymbol{y}(t)$ (green line), and $FF_{d(t)}$ (blue line)

The following steps design the simulation process for performance evaluation of the proposed feedforward-feedback learning-based R2R for controlling the CD in DE process for a SISO system.

Step 1: Consider $FF_{d(t)}$ in simulation design equivalent to the empirical data as illustrated in Figure 6. In addition, initiate $\boldsymbol{y}_1(1)$ and $\boldsymbol{u}_1(1)$ equal to the empirical result.

Step 2: Consider 200 data set $(\boldsymbol{y}(t), \boldsymbol{u}(t), FF_{d(t)})$ as the information for 8 lots.

Step 3: Set number of iteration $j = 20$.

Step 4: Initiate the parameter setting for $Q(q)$, $L(q)$, $slope$, and C_{FF} as $(0.5,1,1,1)$, respectively.

Step 5: Set $\boldsymbol{y}_d(t)$ equal to the $\boldsymbol{u}_1(t-1)$.

Step 6: Generate virtual data by "just in time learning" approach as mentioned in Section 3.2, where parameters of k-mean clustering method are selected by tuning algorithm.

Step 7: Optimize the model in (8) by gradient decent method in Owens *et al.* (2009).

Step 8: Adopt Residual Mean Square Error (RMSE) in (9), and Range in (10) for performance comparison between the feedforward-feedback learning-based R2R controller and empirical data.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{200} |\boldsymbol{e}(t)|^2}{t}} \tag{9}$$

$$Range = \max_t \boldsymbol{y}(t) - \min_t \boldsymbol{y}(t) \tag{10}$$

Figure 7 illustrates the effect of iteration on the performance of the learning control system for the first 11 iterations of the simulation scenario. It is clear that always a high number of repetition cannot guarantee the better performance. Among all situations, $j = 7$ and $j = 8$ perform the best result in case of Range and RMSE, respectively. In addition, error reaches to the steady-state condition after 8-th iterations. Table 3 summarizes the effect of RMSE and Range for each iteration after 200 runs.

Table 3. RMSE, and Range for each iteration

Index	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j=7$	$j=8$	$j=9$	$j=10$
$RMSE$	-	960	472	221	107	44	26	19	24	31
$Range$	-	1710	797	370	157	82	52	75	103	148
Index	$j=11$	$j=12$	$j=13$	$j=14$	$j=15$	$j=16$	$j=17$	$j=18$	$j=19$	$j=20$
$RMSE$	41	53	68	85	107	133	164	202	249	306
$Range$	198	255	331	425	543	682	853	1062	1318	1640

Figure 7 shows that how iteration can be helpful for controller to learn and eliminate the effect of unmeasurable disturbance. As we can see for $j = 2$, error is calculated by $\boldsymbol{e}_1(t) = \boldsymbol{u}_1(t) - \boldsymbol{y}_2(t)$ and due to the initial setting of parameters defined in **Step 4**, the cumulative error is positive. Since the estimated error $\boldsymbol{e}(t)$ in comparison with the estimated input $\boldsymbol{u}(t)$ is very small (see Figure 8), therefore, the effect of input is more stronger than the effect of error on estimating the output $\boldsymbol{y}(t)$. In this example, as we can see on Figure 7 the value of $\boldsymbol{u}(t)$ is always positive, therefore, it is expected that error has positive value at the first iteration is expected. For the second iteration $\boldsymbol{y}_d(t) = \boldsymbol{u}_1(t)$ remains fix, however, $\boldsymbol{y}_j(t)$ is indirectly affected by $\boldsymbol{u}_{j-1}(t)$ and this causes that $\boldsymbol{e}_2(t)$ has opposite sign of $\boldsymbol{e}_1(t)$. This pattern is repeated until learning algorithm (iterations) can eliminate the effect of unmeasurable disturbance.

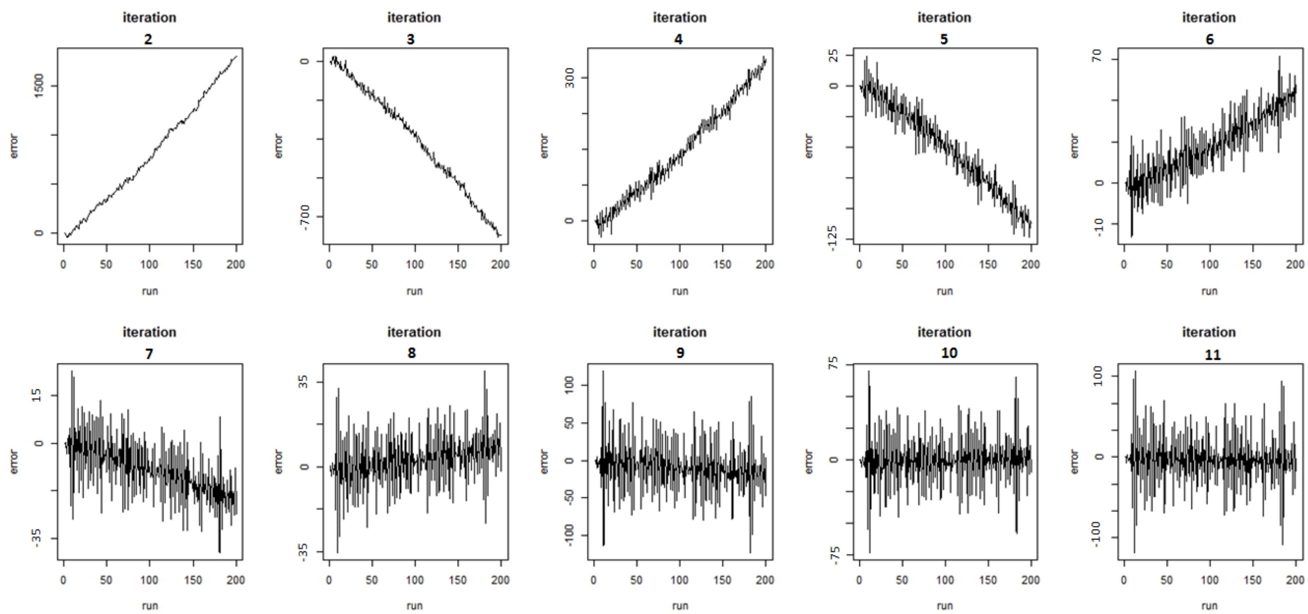

Figure 7. The effect of iteration on optimization of (8).

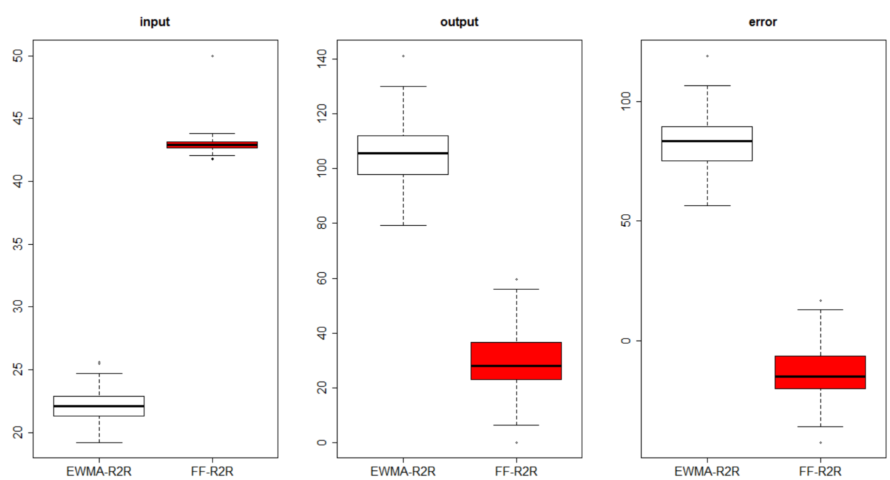

Figure 8. Performance comparison between proposed R2R controller and empirical data, where the y-axis indicates $\boldsymbol{u}(t)$, $\boldsymbol{y}(t)$, and $\boldsymbol{e}(t)$ for each plot from left to right, respectively.

Figure 8, shows the comparison results between empirical data (EWMR-R2R with fixed discount factor 0.3) and feedforward-feedback learning-based R2R. The results indicate that the proposed R2R controller tightens up the excellent performance bound, and eventually achieves a lower cost, together with an extensive disturbance, in comparison with the empirical control design (EWMR-R2R with fixed discount factor 0.3). In total, although in contrast, the proposed R2R controller has weakness in the performance improvement for Range (see Table 4), it can be taught to the system to estimate the output very close to the input which resulted in improvement of the error.

As discussed in Amann *et al.* (1996), the feedforward-feedback learning-based R2R algorithm employed in this study can be implemented in practice if and only if the feedback loop is available. For implementation, the free parameters Q, and L in (8) must be chosen appropriately. The parameter L is related to the size of the error, and the parameter Q to the size of the change of the input. Therefore the sensitivity analysis is essential for study the speed of convergence.

Table 4. RMSE, and Range for each iteration

Index	**RMSE**		**Range**	
	EWMA-R2R	Learning-R2R	EWMA-R2R	Learning-R2R
$\boldsymbol{e}(t)$	82.147	19.051	42.75	75.68
$\boldsymbol{u}(t)$	22.2	4.37	5.579	4.88
$\boldsymbol{y}(t)$	104.21	22.43	40.74	77.97

To illustrate the effect of Q and L, we could consider Q to be fixed and set $L = aQ$, $a > 0$. Therefore, for small a the algorithm is expected to change the incremental input substantially to achieve a small error and causing a fast rate of decrease of J_t. Contrariwise, for large a the convergence rate of J_t will hold with a slow rate in decreasing pattern.

5. CONCLUSION

Smart decision support system is critical for intelligent manufacturing, integrating operations research modeling, optimization, big data analytics, and AI to empower flexible decisions with complicated objectives involved in strategic, operational, and tactical decisions. For a process that is repetitive or cyclic, the learning type control method should be the first choice for control. The specific type of learning type control should then be selected according to the characteristics of the process. In this study regards to the cycling nature of the semiconductor manufacturing and the different profiles of input variables, the hybrid feedforward-feedback learning-based R2R control system is selected for process monitoring of DE process.

The feedforward-feedback learning-based R2R system has proven to be accurate and flexible for monitoring the CD during the DE process. This approach has modeled wafer fabrication processes with the low error for empirical data compare to EWMA-R2R control design. The presented methodology is able to learn from the input variable and ignore the effect of unmeasurable uncertainties through the iterative learning process and the help of virtual data, and compensate the variation of CD.

The constraint-free optimization algorithm in (8) is evolved by gradient learning approach and has the capability to work online and offline for the supervisory control plant. However, to facilitate the optimization algorithm, the system can be modeled firstly by historical data in the offline mode to initiate the parameter setting for online mode. Regards to the data-warehouse management strategy, the offline model can also store the feedforward/feedback signal in the data warehouse till receiving the feedback/feedforward signal then can optimize the system.

The result presented in this paper was regarding the capabilities of the feedforward-feedback learning-based R2R controller for LTI system. Furthermore, there are a number of extensions that could be considered as the future research direction. Future research can be done to employ big data analytics (e.g. Chien and Chuang, 2014; Khakifirooz *et al.*, 2018) to enhance CD control and the yield. Also, more studies can be done to address the issues of convergence rate and robustness of the proposed control system. As it is an often case in semiconductor manufacturing, there would be a potential of the technique for investigating the non-linear optimization model for CD instead of the linear model in (3). Also, similar control system can be considered with time-varying metrology delay of LTI system. There is a great deal of research needed for comparison in the area of learning-based control system, such as NN or kernel-based optimization control systems. The learning algorithm could be designed for learning the different source of disturbance (i.e., non-stationary disturbance or stationary disturbance) and the model in (3) could expand by the effect of the other environmental variables such as gas flows, and temperature. The optimization algorithm could involve the control variable with constraint and enhance the performance of the CD controller.

ACKNOWLEDGEMENTS

This research is supported by the Artificial Intelligence for Intelligent Manufacturing Systems (AIMS) Research Center of Ministry of Science and Technology, Taiwan (MOST 107-2634-F-007-002; MOST 107-2634-F-007-009).

REFERENCES

Amann N, Owens DH, Rogers E. Iterative learning control for discrete-time systems with exponential rate of convergence. IEE Proceedings-Control Theory and Applications. 1996 Mar;143(2):217-24.

Antsaklis PJ. Intelligent control. Wiley Encyclopedia of Electrical and Electronics Engineering. 2001 Aug 21.

Åström KJ, Wittenmark B. Adaptive control. Courier Corporation; 2013 Apr 26.

Baseman RJ, He J, Yashchin E, Zhu Y, inventors; International Business Machines Corp, assignee. Run-to-run control utilizing virtual metrology in semiconductor manufacturing. United States patent US 9,299,623. 2016 Mar 29.

Boeren F, Bareja A, Kok T, Oomen T. Frequency-domain ILC approach for repeating and varying tasks: With application to semiconductor bonding equipment. IEEE/ASME Transactions on Mechatronics. 2016 Dec 1;21(6):2716-27.

Bristow DA, Tharayil M, Alleyne AG. A survey of iterative learning control. IEEE Control Systems. 2006 Jun;26(3):96-114.

Chen J, Cheng N, Cheng YC. Hybrid model based iterative learning control for semiconductor processes with uncetain metrology delay. In Industrial Electronics and Applications (ICIEA), 2011 6th IEEE Conference on 2011 Jun 21 (pp. 1519-1524). IEEE.

Chen Y, Moore KL, Ahn HS. Iterative learning control. InEncyclopedia of the Sciences of Learning 2012 (pp. 1648-1652). Springer, Boston, MA.

Cheng C, Chiu MS. A new data-based methodology for nonlinear process modeling. Chemical Engineering Science. 2004 Jul 1;59(13):2801-10.

Chien CF, Chen YJ, Hsu CY. A novel approach to hedge and compensate the critical dimension variation of the developed-and-etched circuit patterns for yield enhancement in semiconductor manufacturing. Computers & Operations Research. 2015 Jan 1;53:309-18.

Chien CF, Chen YJ, Hsu CY, Wang HK. Overlay Error Compensation Using Advanced Process Control with Dynamically Adjusted Proportional-Integral R2R Controller. IEEE Transactions on Automation Science and Engineering. 2014;11(2), 473-484.

Chien CF, Chuang SC. A Framework for Root Cause Detection of Sub-batch Processing System for Semiconductor Manufacturing Big Data Analytics. IEEE Transactions on Semiconductor Manufacturing. 2014;27(4): 475-488.

Chien CF, Hsu CY. UNISON Analysis to Model and Reduce Step-and-Scan Overlay Errors for Semiconductor Manufacturing. Journal of Intelligent Manufacturing. 2011;22(3), 399-412.

El Chemali C, Freudenberg J, Hankinson M, Collison W, Ni T. Critical dimension control of a plasma etch process by integrating feedforward and feedback run-to-run control. Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena. 2003 Nov;21(6):2304-12.

El Chemali C, Freudenberg J, Hankinson M, Bendik JJ. Run-to-run critical dimension and sidewall angle lithography control using the PROLITH simulator. IEEE Transactions on Semiconductor Manufacturing. 2004 Aug;17(3):388-401.

Gal PL, Bars ML. Rotating thermal flows in natural and industrial processes. Geophysical & Astrophysical Fluid Dynamics. 2013 Sep; 107(6): 720-723

Ghodssi R, Lin P, editors. MEMS materials and processes handbook. Springer Science & Business Media; 2011 Mar 18.

Hsu CY, Wu JZ. Error-smoothing exponentially weighted moving average for improving critical dimension performance in photolithography process. International Journal of Industrial Engineering. 2016 Sep 1;23(5).

Hunt KJ, Sbarbaro D, Żbikowski R, Gawthrop PJ. Neural networks for control systems—a survey. Automatica. 1992 Nov 1;28(6):1083-112.

Ivanov A. Silicon Anodization as a Structuring Technique: Literature Review, Modeling and Experiments. Springer; 2017 Sep 10.

Jebri MA, El Adel EM, Graton G, Ouladsine M, Pinaton J. Virtual Metrology applied in Run-to-Run Control for a Chemical Mechanical Planarization process. InJournal of Physics: Conference Series 2017 Jan (Vol. 783, No. 1, p. 012042). IOP Publishing.

Jörg T, Hofer AM, Köstenbauer H, Winkler J, Mitterer C. Oxidation and wet etching behavior of sputtered Mo-Ti-Al films. Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films. 2018 Mar;36(2):021513.

Kang P, Kim D, Lee HJ, Doh S, Cho S. Virtual metrology for run-to-run control in semiconductor manufacturing. Expert Systems with Applications. 2011 Mar 1;38(3):2508-22.

Kim H, Park JH, Lee KS. Methods and properties of quadratic iterative learning control for semi-conductor processes under different perturbations. In Control, Automation and Systems (ICCAS), 2013 13th International Conference on 2013a Oct 20 (pp. 570-573). IEEE.

Kim HT, Lee KW, Yang HJ, Kim SC. A self-learning method for automatic alignment in wafer processing. International Journal of Precision Engineering and Manufacturing. 2013b Feb 1;14(2):215-21.

Khakifirooz M, Chien CF, Chen YJ. Bayesian inference for mining semiconductor manufacturing big data for yield enhancement and smart production to empower Industry 4.0. Applied Soft Computing. 2018;68: 990-999.

Kuo AD. The relative roles of feedforward and feedback in the control of rhythmic movements. Motor control. 2002 Apr;6(2):129-45.

Mao ZJ, Kang W. Benchmark study of run-to-run controllers for the lithographic control of the critical dimension. Journal of Micro/Nanolithography, MEMS, and MOEMS. 2007 Apr;6(2):023001.

Moyne J, Del Castillo E, Hurwitz AM. Run-to-run control in semiconductor manufacturing. CRC press; 2000 Nov 30.

Ngo YS, Ang KT, Tay A. Method for real-time critical dimensions signature monitoring and control: Sensor, actuator, and experimental results. Review of Scientific Instruments. 2013 Jan;84(1):015116.

Owens DH, Hatonen JJ, Daley S. Robust monotone gradient‐based discrete‐time iterative learning control. International Journal of Robust and Nonlinear Control: IFAC‐Affiliated Journal. 2009 Apr;19(6):634-61.

Qin SJ, Cherry G, Good R, Wang J, Harrison CA. Semiconductor manufacturing process control and monitoring: A fab-wide framework. Journal of Process Control. 2006 Mar 1;16(3):179-91.

Sawin HH. Challenges in dry etching: uniformity, selectivity, pattern dependencies, damage, and cleaning. Microelectronic Engineering. 1994 Jan 1;23(1-4):15-21.

Steinbuch M. Repetitive control for systems with uncertain period-time. Automatica. 2002 Dec 1;38(12):2103-9.

Tan F, Pan T, Li Z, Chen S. Survey on run-to-run control algorithms in high-mix semiconductor manufacturing processes. IEEE Transactions on Industrial Informatics. 2015 Dec;11(6):1435-44.

Tandou T, Kubo S, Negishi N, Izawa M. Improving the etching performance of high-aspect-ratio contacts by wafer temperature control: Uniform temperature design and etching rate enhancement. Precision Engineering. 2016 Apr 1;44:87-92.

Tsai PF, Tsen A, Sung JN, inventors; Taiwan Semiconductor Manufacturing Co (TSMC) Ltd, assignee. System and method for implementing a virtual metrology advanced process control platform. United States patent US 8,437,870. 2013 May 7.

Turkot B, Carson S, Lio A. Continuing Moore's law with EUV lithography. In Electron Devices Meeting (IEDM), 2017 IEEE International 2017 Dec 2 (pp. 14-4). IEEE.

Wang Y, Gao F, Doyle III FJ. Survey on iterative learning control, repetitive control, and run-to-run control. Journal of Process Control. 2009 Dec 1;19(10):1589-600.

Won W, Park K, Kim J. Combined iterative learning and delta-operator adaptive linear quadratic Gaussian control of a commercial rapid thermal processing system. Chemical Engineering Science. 2017 Dec 31;174:146-56.

Wu CF, Hung CM, Chen JH, Lee AC. Advanced process control of the critical dimension in photolithography. International Journal of Precision Engineering and Manufacturing. 2008 Jan;9(1):12-8.

Xu JX, Chen Y, Lee TH, Yamamoto S. Terminal iterative learning control with an application to RTPCVD thickness control. Automatica. 1999 Sep 1;35(9):1535-42.

Yang G, Ngo YS, Putra AS, Ang KT, Tay A, Fang ZP. Monitoring and control of photoresist properties and CD during photoresist processing. In Metrology, Inspection, and Process Control for Microlithography XXIV 2010 Apr 2 (Vol. 7638, p. 763828). International Society for Optics and Photonics.

Yu HC, Lin KY, Chien CF. Hierarchical Indices to Detect Equipment Condition Changes with High Dimensional Data for Semiconductor Manufacturing. Journal of Intelligent Manufacturing, 2014;25(5), 933-943.