PLS Based Run-to-Run Control Design for MIMO Non-Squared Semiconductor Processes

Fan Wang (王凡) and Junghui Chen*(陳榮輝)
Department of Chemical Engineering
Chung-Yuan Christian University,
Chung-Li, Taiwan 320, Republic of China

Abstract

The PLS (partial least squares) based dEWMA (double exponentially weighted moving average) control algorithm is proposed to adjust a MIMO non-squared process which has an unequal number of inputs and outputs from run to run in semiconductor manufacturing. Recently, the multivariable dEWMA controller was developed, but it could not avoid the ill-conditioned problem when the control rule was solved. Thus, an arbitrary nonnegative value is often included. In this paper, the enhancement of the multivariable dEWMA controller is done by incorporating a popular linear technique, PLS. The PLS based enhancement is an improvement over the above method that has the occurrence of the collinearity of the input and the output variables. It is particularly useful for inherent noise suppression. The major advantage of the proposed method is that the non-squared MIMO process is decomposed into several independent SISO systems. The dEWMA controller can be separately applied to each SISO system. The performance of the proposed method is illustrated through a Chemical-Mechanical Polishing (CMP) process, which is a critical semiconductor manufacturing processing step.

1. Introduction

In the semiconductor manufacturing industries, they are challenged by reduced circuit critical dimension and increased wafer size in order to improve and achieve maximum performance, like processor speed and analogy frequency response, out of the existing equipments. Nowadays, the industries are getting closer to the sub-0.1 μm technology. In order to continuously achieve high yield and reduce the process variability, the tight process control solution plays a significant role in increasing throughput (Smith and Boning, 1999).

Run-to-Run (RtR) control has been widely applied in the semiconductor manufacturing industry (Ingolfsson and Sachs, 1993; Hu *et al*., 1992; Sachs *et al*., 1991). The control algorithms were pioneered by researcher at MIT and at the University of Michigan as well as by workers at various semiconductor manufactures. The RtR process control consists of two major steps. First, a linear regression model is updated based on the available in-situ and ex-situ measurements of the past runs. The model relates the input variable to the output variable. The static model was often used in the literature (Boning *et al*., 1996; Sachs et al., 1995). The second step in RtR control is to determine a control action which is known as the recipe in the microelectronics literature. It is used to reach the desired values or improve the performance for the next run. The exponentially weighted moving average (EWMA) gives different weights to measurement data from past runs used to computer the RtR controller. It has been applied to chemical mechanical polishing (Boning *et al*., 1996) and plasma etching (Moyne *et al*., 1995). However, due to aging of the process, like depletion of etch solution or degradation of thermocouples in high temperature furnace in etch reactors, the existing EWMA RtR controllers cannot compensate for such process drifts or ramp disturbances and, as a result, offset in the process outputs occurs [Smith and Boning, 1997]. This will lead to poor performance, such as deposition nonuniformities. Butlter and Stefani (1994) proposed the double EWMA (dEWMA) controller which eliminated this offset for a polysilicon gate etch process. Further extensive studies were done to improve or tune the RtR controller, but the development was limited to single-input single-output processes. Controlled processes in nearly all-semiconductor manufacturing frequently encounters with inherently more than one variable to be controlled (Sachs *et al*., 1995; Roy *et al*., 1994). The system may either have more inputs than outputs or more outputs than inputs. For example, two important output variables of the CMP process are the remaining thickness and its uniformity within a wafer. The input recipes are consisted of polishing time, table speed, downforce etc. They are known as multivariable or multi-input multi-output (MIMO) processes. The control of multivariable systems is not always an easy task due to its complex and interactive nature. Few of them are intended for MIMO processes. Tseng *et al*. (2002) proposed a multivariate extension to a single EWMA controller. Del Castillo and Rajagopal (2002, 2003) presented a multivariable extension of dEWMA controller and analyzed the robustness and stability conditions. However, an appropriate value of the Lagrange multiplier constant should be selected in MIMO controller in order to have an invertible matrix and reduce the larger variation of inputs.

The development of chemometric techniques has spurred a torrent of research in multivariable processes. Those techniques can be used to extract the state of the system via applications of mathematical and statistical methods from the stored data. Several chemometric techniques were proposed, like principal component analysis and partial least squares (PLS). They have received considerable attention in the field of chemical process problem and have been applied to system monitoring and diagnosis (Kourti and MacGregor, 1996). Still, it was rarely on the control problem. Recently, a PLS projection-based model was proposed (Lakshminaraynan et al., 1997). A PLS outer model was first constructed. The relationship between the input and the output scores were built on the inner model. However, the control objective was still lumped when model predictive control was used. Another PLS decomposition strategy for PID control system design was addressed (Chen and Chang, 2004). PLS first decompose a MIMO process into a multi-loop control system in a reduced subspace. Then the optimal tuning multiloop PID controller can be directly and separately applied to each control loop. The control performances of the above methods were satisfactory. They could successfully tackle operational data analysis. It showed that only a few principal components could capture most of the characteristics of the system pattern in a multivariable process behavior. Based on the PLS modeling method, it is worth extending it to RtR controller design for improving the currently met problems in MIMO RtR controller.

The remainder of this paper is structured as follows: The second section defines the type of MIMO RtR control problem. PLS technique for the decomposition of MIMO is discussed in Section 3. In Section 4, the PLS based MIMO model is first derived. Then the proposed decomposition control design for directly applying the dEWMA RtR control algorithm onto each loop is developed. The effectiveness of the proposed method is demonstrated through a simulation benchmark of the CMP process in Section 5. The example investigates the performance of the proposed method and makes a comparison with the conventional algorithms. Finally, concluding remarks are made.

2. MIMO RtR Control Problem

The model of the MIMO process is similar to that of the SISO process, but there are M inputs or controllable factors and N outputs or responses. The number of inputs and outputs may be unequal. It is given in Fig. 1. The model can be described as

$$\mathbf{y}_k = \boldsymbol{\alpha} + \mathbf{B}\mathbf{x}_{k-1} + \mathbf{n}_k \qquad (1)$$

where

$$\mathbf{n}_k = \boldsymbol{\delta}k + \boldsymbol{\varepsilon}_k \qquad (2)$$

where $\mathbf{x}_k = \begin{bmatrix} x_{1,k} & \cdots & x_{M,k} \end{bmatrix}^T$ denotes the vector of input recipes, $\mathbf{y}_k = \begin{bmatrix} y_{1,k} & \cdots & y_{N,k} \end{bmatrix}^T$ denotes the vector of output variables, $\boldsymbol{\varepsilon}_k = \begin{bmatrix} \varepsilon_{1,k} & \cdots & \varepsilon_{N,k} \end{bmatrix}^T$ denotes the process disturbance with the white noise sequence, k is a run number, $\boldsymbol{\delta}$ is a vector equal to average drift rate per run, $\boldsymbol{\alpha}$ is a vector of the offset parameters, and $\mathbf{B}$ is a process gain matrix. $\boldsymbol{\delta}$, $\boldsymbol{\alpha}$ and $\mathbf{B}$ are unknown vectors and a matrix to be estimated.

At the end of each run k, the goal of the control design for the MIMO system is to seek a new recipe $\mathbf{x}_k$ for use in the next run in order to correct the deviation from the desired target vector ($\boldsymbol{\tau}$). That is, the objective function of the MIMO system is expressed as

$$\min \left\| \hat{\mathbf{y}}_{k+1} - \boldsymbol{\tau} \right\|^2 \qquad (3)$$

subject to $\hat{\mathbf{y}}_{k+1} = \hat{\boldsymbol{\alpha}} + \hat{\mathbf{B}}\mathbf{x}_k + \hat{\mathbf{n}}_{k+1}$.

In the above objective equation, a one-step-ahead output is predicted for the estimated deviation from the target in the future of the next run. A MIMO dEWMA controller can be derived (Del Castillo and Rajagopal, 2002),

$$\mathbf{x}_k = \left(\hat{\mathbf{B}}^T \hat{\mathbf{B}}\right)^{-1} \hat{\mathbf{B}}^T \left(\boldsymbol{\tau} - \hat{\boldsymbol{\alpha}} - \hat{\mathbf{n}}_{k+1}\right) \quad (4)$$

Note that the matrix $\left(\hat{\mathbf{B}}^T \hat{\mathbf{B}}\right)$ is often not a positive definite. This comes from the high degree of coupling and correlation in this MIMO process. In such case, the ordinary least-squares technique will result in the estimated recipe with large variances owing to the ill-conditioned nature of the problem. This may cause oscillation and even instability. One way to circumvent this problem is to add a positive definite matrix $\lambda \mathbf{I}$ to $\left(\hat{\mathbf{B}}^T \hat{\mathbf{B}}\right)$, where $\mathbf{I}$ is the identity matrix and λ is some nonnegative value. This way is just to suppress all inputs, not just some inputs which have the weak relationships with the outputs. This may degrade the performance.

Past research shows multivariable statistical techniques such as PCA and PLS can offer the data-compression facility to condense the variance of the process into a very low-dimensional latent subspace. This data-compression feature provides a technique that break up a multivariate problem into a series of univariate problems. The model is constructed based on the signals with sufficient low-frequency content, which is the essential behavior of the system. Before going through the proposed PLS based RtR controller, we have to digress a little to get some background information about the properties of PLS.

3. **Partial Least Squares: Overview**

PLS regression derived from the classical linear regression is often used to predict properties of processes based on variables only indirectly related to the properties. The given process data are subdivided into two blocks, a dependent block ($\mathbf{Y}$) and an independent block ($\mathbf{X}$). $\mathbf{Y}$ block with a two-way array ($I \times M$) summarizes the I runs and the M final properties (or responses). $\mathbf{X}$ block with a two-way array ($I \times N$) organizes the controllable N factors. PLS is used to extract latent variables. The latent variables explain the best correlation between the product response block ($\mathbf{Y}$) and the controllable factor block ($\mathbf{X}$).

The standard PLS regression (Höskuldsson, 1988) relies on decomposing the dependent block ($\mathbf{Y}$) and the independent block ($\mathbf{X}$) into a sum of rank one component matrices. Before applying PLS, each measurement variable that centers and scales the variance to unit one is typically applied; This will put all variables on an equal basis. Initially, let $\mathbf{X}_0 = \mathbf{X}$, $\mathbf{Y}_0 = \mathbf{Y}$ and $r = 0$. Find out a vector (or component) ($\mathbf{w}_r$) which is correlated with $\mathbf{Y}_r$ while describing a large amount of the variation in $\mathbf{X}_r$. It can be formulated as

$$\mathbf{w}_r = \arg\max_{\mathbf{w}_r} \left(\mathbf{X}_r^T \mathbf{Y}_r \middle| \|\mathbf{w}_r\| = 1\right) \quad (5)$$

Then the component is subtracted from $\mathbf{X}_r$ and $\mathbf{Y}_r$

$$\mathbf{X}_r = \mathbf{X}_{r-1} - \mathbf{t}_r \mathbf{w}_r^T \quad (6)$$

$$\mathbf{Y}_r = \mathbf{Y}_{r-1} - b_r \mathbf{t}_r \mathbf{c}_r^T \quad (7)$$

And

$$\mathbf{t}_r = \mathbf{X}_r \mathbf{w}_r \quad (8)$$

$$\mathbf{c}_r = \frac{\mathbf{Y}_r^T \mathbf{t}_r}{\mathbf{t}_r^T \mathbf{t}_r}, \quad \mathbf{u}_r = \mathbf{Y}_r^T \mathbf{c}_r \quad (9)$$

where $\mathbf{w}_r$ and $\mathbf{c}_r$ are the loadings of $\mathbf{X}_r$ and $\mathbf{Y}_r$, respectively. The score ($\mathbf{t}_r$) is the projection of $\mathbf{X}_r$ into the direction $\mathbf{w}_r$. The score ($\mathbf{u}_r$) is the projection of $\mathbf{Y}_r^T$ into the direction $\mathbf{c}_r$. b_r is the regression coefficient related to $\mathbf{t}_r$ and $\mathbf{u}_r$,

$$b_r = \frac{\mathbf{u}_r^T \mathbf{t}_r}{\mathbf{t}_r^T \mathbf{t}_r} \quad (10)$$

Eqs. (6) and (7) are used to remove the variance associated with the already calculated r-th directions of $\mathbf{w}_r$ and $\mathbf{c}_r$ in the variance of process variables and quality variables respectively. Then set $r = r + 1$ and repeat the above procedures (Eqs.(5)-(10)) until the description of $\mathbf{Y}$ convergence is properly gotten. Finally, the matrices $\mathbf{Y}$ and $\mathbf{X}$ are separately decomposed into the summation of the product of score vectors $\mathbf{t}$ and loading vectors $\mathbf{w}$ and $\mathbf{c}$ plus some residual matrix $\mathbf{E}$ and $\mathbf{F}$, respectively:

$$\mathbf{X} = \sum_{r=1}^{R} \mathbf{t}_r \mathbf{w}_r^T + \mathbf{E} = \mathbf{T}\mathbf{W}^T + \mathbf{E}$$
$$\mathbf{Y} = \sum_{r=1}^{R} \mathbf{t}_r \mathbf{c}_r^T + \mathbf{F} = \mathbf{T}\mathbf{C}^T + \mathbf{F} \qquad (11)$$

where R is the number of principal components retained in PLS. Due to its simplicity and easy interpretation, the applications of this approach can be found in an abundant literature. In this paper, the PLS technique is used to eliminate the interaction of the MIMO control problems. The controllable factors and responses are transformed into a smaller informative set via a set of linear functions which model the combinational relationship between the response variables and latent controlled variables, and between the controllable factors and latent manipulated variables, respectively. Then, the dEWMA controller can be directly applied to each independent SISO system.

4. Multiloop SISO Run-to-Run Controller design

The block diagram of the multi-SISO control system to be considered is shown in Fig. 2. The MIMO system model is decomposed into several pairs of the input-output score. The multi-control loop is then applied onto each pair to form a SISO RtR control design problem.

4.1 Conventional dEWMA RtR Controller

Generally, the SISO process model is taken to be linear regression of the form

$$\hat{y}_k = a_k + bx_{k-1} + n_k \qquad (12)$$

where $\hat{y}_k$ is the predicted output at run k, b is the process gain, x_{k-1} is the process input calculated based on previous observation through run $k-1$, a_k is the estimation for intercept and n_k is the deterministic trend disturbance, given by the terms $\delta k + \varepsilon_k$. Typically, the system gain and the initial value of the intercept are modeled a priori from design experiments.

The control law for the dEWMA-based RtR controller is a process inversion of the form

$$x_k = \frac{\tau - a_k - n_k}{b} \qquad (13)$$

where τ is the desired target. The intercept and the deterministic trend disturbance are updated recursively by the observer of the form

$$a_k = \lambda_1(y_k - bx_{k-1}) + (1-\lambda_1)a_{k-1}, \quad 0 < \lambda_1 \le 1 \qquad (14)$$

and

$$n_k = \lambda_2(y_k - bx_{k-1} - a_{k-1}) + (1-\lambda_1)n_{k-1}, \quad 0 < \lambda_2 \le 1 \qquad (15)$$

where λ_1 and λ_2 are the exponential weighting factor of the observer. The observer equations are geometric averages that weight the past observations in an exponentially decreasing manner. Small value of λ_1 and λ_2 are appropriate for systems with small deterministic drifts and relatively large natural variance. Conversely, highly correlated output error is best compensated through use of higher values of the weighting factors.

4.2 PLS Based Decoupling MIMO

The goal of the RtR controller design for the MIMO system is to seek control actions $\mathbf{x}_k$ that can minimize the difference between the responses $\mathbf{y}_{k+1}$ and the desired targets $\boldsymbol{\tau}$ at the next run; it can is expressed as

$$\min_{\mathbf{x}_k} J = \frac{1}{2}\min_{\mathbf{x}_k}\left\|\mathbf{e}_{k+1}\right\|^2 = \frac{1}{2}\min_{\mathbf{x}_k}\left\|\boldsymbol{\tau} - \mathbf{y}_{k+1}\right\|^2 \qquad (16)$$

Let $\boldsymbol{\tau}$ and $\mathbf{y}_{k+1}$ can be decomposed into the lower dimensional space $\boldsymbol{\tau} = \sum_{r=1}^{R} u_r^{set}\mathbf{c}_r$ and $\mathbf{y}_{k+1} = \sum_{r=1}^{R} u_{r,k+1}\mathbf{c}_r$. The above equation can be represented as

$$J = \frac{1}{2}\min_{\mathbf{x}_k}\left\|\sum_{r=1}^{R}\left(u_r^{set} - u_{r,,k+1}\right)\mathbf{c}_r\right\|^2 \qquad (17)$$

Since the objective function involves a term in the future of the next run; namely $u_{r,k+1}$, which is not available at time k. When an one-step ahead output can be predicted; that is, $u_{r,k+1} \cong \hat{u}_{r,k+1}$,

$$\begin{aligned} J &= \frac{1}{2}\min_{\mathbf{x}_k}\left\| \sum_{r=1}^{R}\left(u_r^{set} - \hat{u}_{r,k+1}(k)\right)\mathbf{c}_r \right\|^2 \\ &\le \frac{1}{2}\min_{\mathbf{x}_k}\left[\sum_{r=1}^{R}\left(u_r^{set} - \hat{u}_{r,k+1}\right)^2 \left\|\mathbf{c}_r\right\|^2\right] \\ &= \min_{\mathbf{x}_k}\left[J_1 + J_2 + \cdots\cdots + J_R\right] \\ &= \left[\min_{x_{1,k}} J_1 + \min_{x_{2,k}} J_2 + \cdots\cdots + \min_{x_{R,k}} J_R\right] \end{aligned} \quad (18)$$

This is the consequence of the Schwarz inequality. Let $J_r \equiv \frac{1}{2}\left(u_r^{set} - \hat{u}_{r,k+1}\right)^2 \left\|\mathbf{c}_r\right\|^2$, the objective function is decomposed into R subobjective functions in the lower dimensional subspace, $J = \sum_{r=1}^{R} J_r$. Only R score variables ($t_r, r = 1,2,\cdots R$) require separate design compared with M responses to be lumped together without the decomposition. This multi-SISO design method, like a decentralized control strategy, have a simpler structure and, accordingly, less tuning parameters are needed than the fully cross-coupled one.

4.3 dEWMA RtR Controller Design of Each SISO System

After the objective function is decoupled into R objective functions, the conventional SISO dEWMA controller design technique can be directly applied to each score variable respectively in the decomposed space, because the MIMO system is decomposed using PLS and the interactions which exist between control loops are also eliminated. The only difference is that the process variables are converted into the score variables in the subspace. Each subobjective (J_r) is rearranged into

$$\min_{x_{r,k}} J_r = \frac{1}{2}\min_{x_{r,k}}\left(u_r^{set} - \hat{u}_{r,k+1}\right)^2 \left\|\mathbf{c}_r\right\|^2 \quad (19)$$

Although the input-output model is $\hat{u}_{r,k+1} = b_r t_{r,k}$, it is not always accurate due to unmodeled process dynamics and disturbances that enter the process. Therefore, the bias term is recursively estimated as each run according to the traditional dEWMA filter

$$a_{r,k} = \lambda_{r,1}(u_{r,k} - b_r t_{r,k-1}) + (1 - \lambda_{r,1}) a_{r,k-1} \quad (20)$$

and the trend estimation filter is also estimated as

$$\begin{aligned} n_{r,k} &= \lambda_{r,2}(u_{r,k} - b_r t_{r,k-1} - a_{r,k-1}) \\ &\quad + (1 - \lambda_{r,2}) n_{r,k-1} \end{aligned} \quad (21)$$

Thus, the input for the k+1 run is

$$\mathbf{x}_k = \sum_{r=1}^{R} t_{r,k}\mathbf{w}_r \quad (22)$$

where

$$t_{r,k} = \frac{u_r^{set} - a_{r,k} - n_{r,k}}{b_r} \quad (23)$$

5. Illustration Examples

To study the performance of the proposed PLS-based MIMO RtR control, simulations of the equipment model that characterize a real CMP processes is utilized (Rajagopal and Castillo, 2003). The mode was developed at SEMATECH. In sequent polishes of CMP operation, the polish pad undergoes plastic deformation which cause the pad surface to become smoother and its pores become filled with pad material. This will causes the polish rate to decay over the course of subsequent runs. In this model, the controllable factors are: plate speed (x_1), back pressure (x_2), polishing downforce (x_3), and the profile of the conditioning system (x_4). Two responses are: removing rate (y_1) and within-wafer non-uniformaity (y_2). The objective of the CMP RtR controller is not only to reach the target values at 2000 for y_1 and 100 for y_2 but also to reduce the lot-to-lot variations of the response variables. Two sets of simulation equations described as follows are conducted. The simulation condition and relevant parameters are the same as that of Rajagopal and Castillo (2002).

(1) The approximate linear models:

$$\begin{aligned} y_1 &= 1563.5 + 159.3x_1 - 38.2x_2 + 178.9x_3 + \\ &\quad 24.9x_4 - 0.9k + \varepsilon_{1,k} \\ y_2 &= 254 + 32.6x_1 + 113.2x_2 + 32.6x_3 + \\ &\quad 37.1x_4 + 0.05k + \varepsilon_{2,k} \end{aligned}$$

where $\varepsilon_{1,k} \sim N(0,60^2)$ and $\varepsilon_{2,k} \sim N(0,30^2)$. It is a typical MIMO process with interaction. First the aim is to build up the PLS model based on the collected data. The identification data set contains 100 runs. Another 50 runs which does not come from the training sets are produced in a similar way for validation. The percentage of variance captured by each PLS component is listed in Table 1. It is observed that two principal components capture over 80% of the variance in the relationships of the MIMO process, which suggests that the process variables are fairly well correlated between inputs and outputs. With the proposed PLS based RtR control design strategy, the setpoint targets are traced by the on-line updated algorithm that is in control of the process. Fig. 3 indicates that the proposed multi-SISO updated algorithm is able to trace the setpoint signal in the MIMO process. Meanwhile, the MIMO control algorithm (Eq. (4)) is also applied (Fig. 4). The means and standard deviation of y_1 and y_2 of PLS based RtR are [1999.9, 63.4] and [100.0, 31.0], respectively. The means and standard deviation of y_1 and y_2 of MIMO RtR are [2000.0, 70.0] and [100.0, 35.2], respectively. The controlled variables deviated from the set points are much reduced based on the proposed control algorithm. Also the variations of the control actions of the proposed control are smaller. Furthermore, the proposed decomposition structure can significantly reduce the computation load and more feasibly avoid the non-invertible problem.

(2) The true equipment models:

$$y_1 = 1563.5 + 159.3x_1 - 38.2x_2 + 178.9x_3 + 24.9x_4 - 67.2x_1x_2 - 46.2x_1^2 - 19.2x_2^2 - 28.9x_3^2 - 12x_1k' + 116x_4k' - 50.4k' + 20.4(k')^2 + \varepsilon_{1,k}$$

$$y_2 = 254 + 32.6x_1 + 113.2x_2 + 32.6x_3 + 37.1x_4 - 36.8x_1x_2 + 57.3x_4k' - 2.42k' + \varepsilon_{2,k}$$

where $k' = (k-53)/52$. In this case, the true non-linear equipment equations are carried out. The objective of this study is to test the robustness of the proposed method for a realistic non-linear process, because the RtR controllers are based on the assumption of the linear process and the linear model has a limited range of validity for the nonlinear process. Based on the previously built PLS model, two responses are still around the desired targets in 2000 runs (Fig. 5). The means and standard deviation of y_1 and y_2 are [2000.5, 176.9] and [99.9 74.4], respectively. As for the MIMO control design (Eq. (4)), the only results over the first 1000 runs are plotted here (Fig. 6.). Because of the linear model error and difficulty in control, the controlled variables become unstable. The means and standard deviation of y_1 and y_2 for the 1000 runs are [2000.4, 95.5] and [99.9, 38.9], respectively. The variations of the proposed control are much smaller. Based on these comparisons, it indicates the proposed model can retain the most essential process information and filter out the large, high-frequency nonlinear variation to keep the kernel behavior of the nonlinear system over a long period of runs even if the nonlinear process with the inherent noise.

6. Conclusion

In this paper a SISO dEWMA RtR controller design strategy is developed for the design of the MIMO dEWMA RtR controller system. Rather than using the traditional MIMO control design with the lump structure, the proposed method explores the benefits of using the decomposition PLS framework in the reduced subspace and the optimal control design of the MIMO system. The PLS model structure can decompose the MIMO control problem into the several independent SISO control problems. Thus, the conventional SISO RtR design strategy can be directly used to determine each SISO system. Of additionally important to multivariable process is PLS applicability to non-square systems. To sum up, the proposed algorithm has the following advantages: (i) It is simple to design the controllers based on the SISO control algorithms individually since it is not necessary to design the MIMO system based on the whole system variables. (ii) The coupling effect in the MIMO system can be overcome effectively. The PLS structure can be decomposed into several pairs of inputs and outputs, so the number of SISO control loops can be selected based on the variation captured by each pair. The potential of the proposed technique for prediction and process control are demonstrated by means of a CMP simulation study. Modeling and control performed on the large-scale problems and the real lab-scale experiments will be included in our next research.

References

(1) Boning, D. S.; Moyne, W. P.; Smith, T. H.; Moyne, J.; Telfeyan, R.; Hurwitz, A.; Shellamn, S. and Taylor, J., "Run by run control of chemical-mechanical polishing," *IEEE Trans. on Components, Packages, Packing, and Manufacturing Technology-Part C*, 19, 307, 1996.

(2) Butler, S. W. and Stefani, J. A., "Supervisory run-to-run control of a polysilicon gate etch using in situ ellipsometry," *IEEE Trans. on Semiconductor Manufacturing*, 7, 193, 1994.

(3) Chen, J. and Cheng, Y.-C., "Applying PLS-based decomposition structure to multi-loop adaptive PID controllers in nonlinear processes, *Ind. Eng. Chem. Res.*, 43, 5888, 2004.

(4) Del Castillo, E. and Rajagopal, R., "A multivariable double EWMA process adjustment scheme for drifting processes," *IIE Transactions*, 34, 1055, 2002.

(5) Höskuldsson, A., “PLS regression methods,” *J. of Chemometrics*, 2, 211, 1988.
(6) Hu, A.; Sachs, E. and Ingolfsson, A., “Run by run process control: performance benchmarks,” *IEEE/SEMI int’l Semiconductor Manufacturing Science Symposium*, 73, 1992
(7) Ingolfsson, A. and Sachs, E., “Stability and sensitivity of an EWMA controller,” *Journal of Quality and Technology*, 25, 271, 1993.
(8) Kourti, T. and MacGregor, J. F., “Multivariate SPC methods for process and product monitoring,” *J. Quality Technology*, 28, 409, 1996.
(9) Lakshminaraynan, S.; Shan, S. L. and Nandakumar, K., “Modeling and control of multivariable processes: the dynamic projection to latent structure approach,” *AIChE J.,* 43, 2307, 1997.
(10) Moyne, J. R.; Chaudry, N. and Tefeyan, R., “Adaptive extensions to a multi-branch run-to-run controller for plasma etching,” *Journal of Vacuum Science and Technology*, 13, 1787, 1995
(11) Rajagopal, R. and Del Castillo, E., “An analysis and MIMO extension of a double EWMA run-to-run controller for non-squared systems,” *International Journal of Reliability, Quality and Safety Engineering*, 10, 417, 2003.
(12) Roy, S. R.; Glynn, P.; Hogan, R. and Reynolds, J., “A novel reactor design configuration for contamination control and improved performance in the polysilicion doping process using pocl3,” *Journal of Electrochemical Society*, 41, 2257, 1994.
(13) Sachs, E.; Hu, A. and Ingolfsson, A., “Run by run process control: combining SPC and feedback control,” *IEEE Trans. on Semiconductor Manufacturing*, 8, 26, 1995.
(14) Sachs, E.; Guo, R. and Hu., A., “Process control system for VLSI fabrication,” *IEEE Trans. on Semiconuctor Manufacturing*, 4, 134, 1991
(15) Smith, T. H. and Boning, S. D., “Process control in the semiconductor industry,” *Industrial Engineering Research Conference*, May 1999.
(16) Smith, T. H. and Boning, S. D., “Artificial neural network exponentially weighted moving average controller for semiconductor process,” *Journal of Vacuum Science and Technology* A, 15, 1377, 1997
(17) Tseng, S. T.; Chou, R. J. and Lee, S. P., “A study of a multivariable EWMA controller,” *IIE Transactions*, 34, 541, 2002.

Table 1: Percentage of variance captured by each PLS component

	Percent Variance Captured by Each PLS Component			
Component	Xblock	Total	Yblock	Total
1	28.80	28.80	62.46	62.46
2	24.73	53.53	20.93	83.38
3	25.04	78.57	1.30	84.68
4	21.43	100.00	0.38	85.07

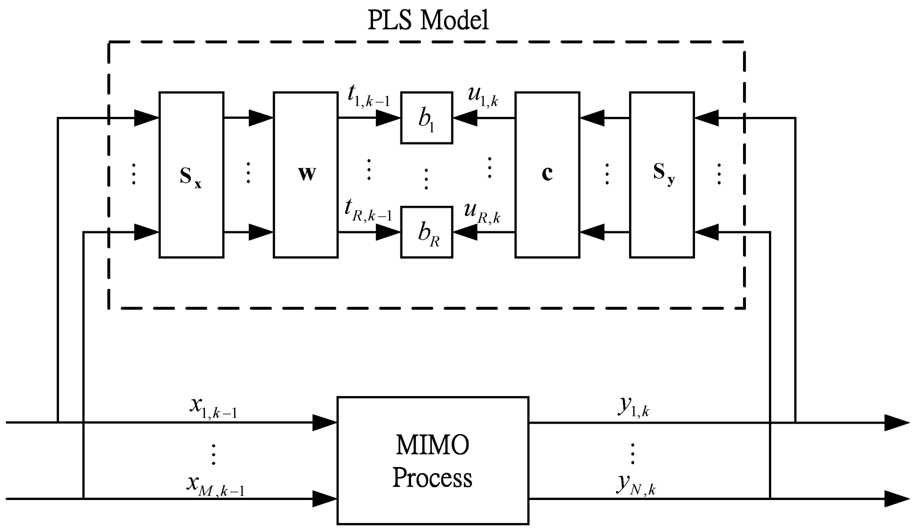

Fig. 1. Structure of PLS-based MIMO model. $\mathbf{S}_x$ and $\mathbf{S}_y$ are the factors that scale the controllable factors and responses respectively.

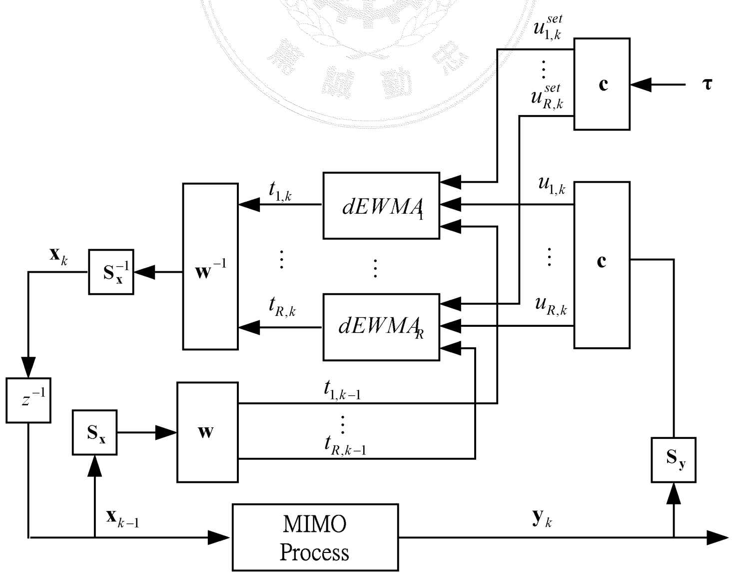

Fig. 2. Implementation of the PLS model-based RtR controller design. $\mathbf{S}_x$ and $\mathbf{S}_y$ are the factors that scale the controllable factors and responses respectively. $\mathbf{S}_x^{-1}$ is the factor that rescales the controllable factors.

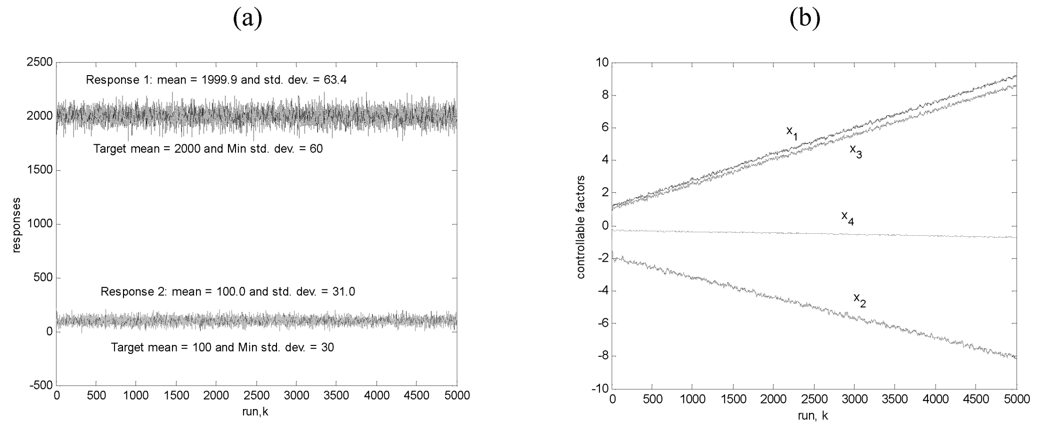

Fig. 3. Control performance of a linear CMP process controlled by PLS-based MIMO controller: (a) responses and (b) controllable factors.

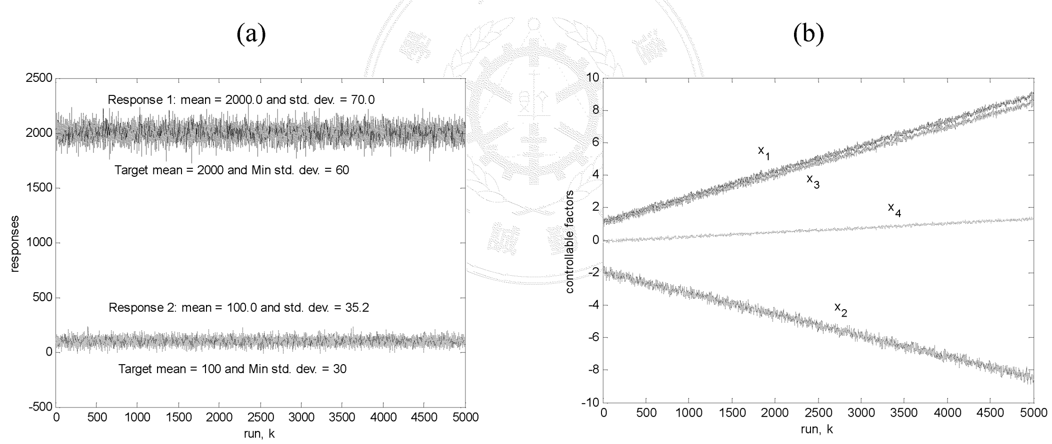

Fig. 4. Control performance of a linear CMP process controlled by MIMO controller: (a) responses and (b) controllable factors.

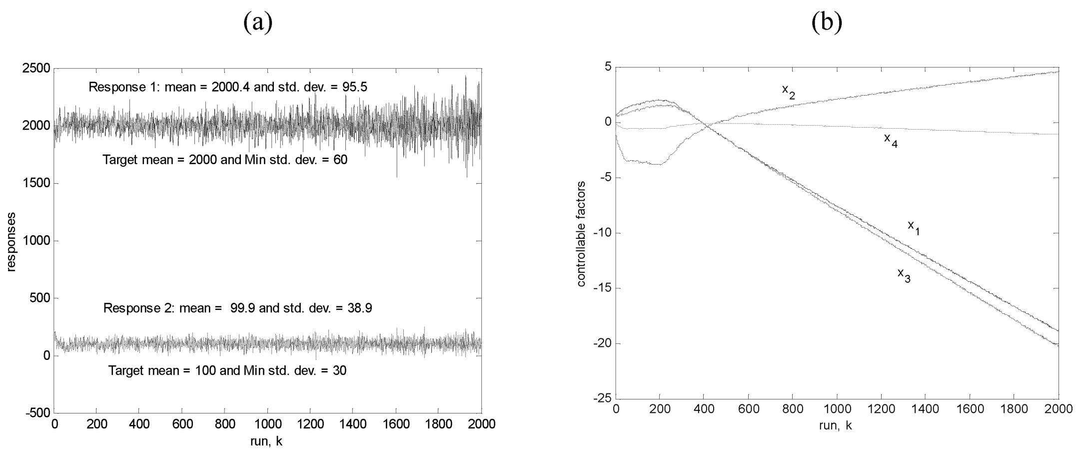

Fig. 5. Control performance of a non-linear CMP process controlled by PLS-based MIMO controller: (a) responses and (b) controllable factors.

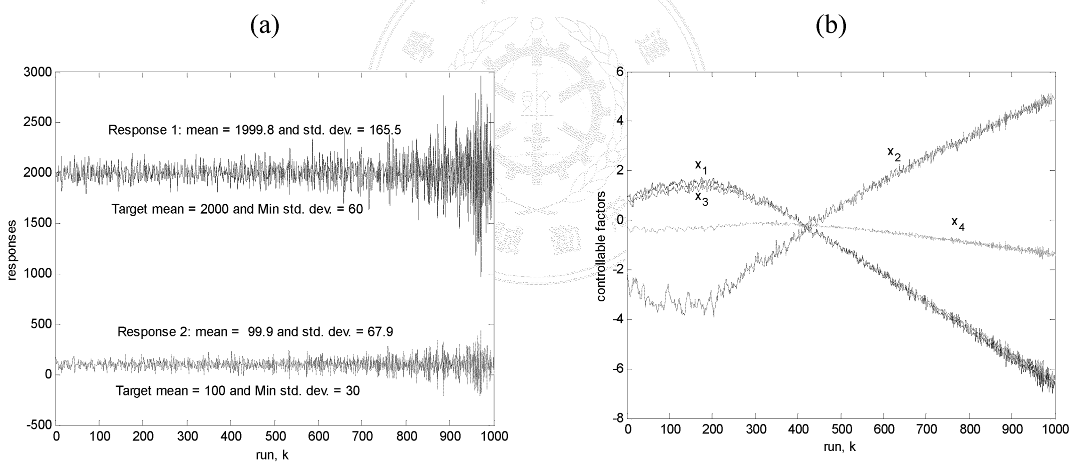

Fig. 6. Control performance of a non-linear CMP process controlled by MIMO controller: (a) responses and (b) controllable factors.