

Run-To-Run Control Schemes For CMP Process Subject to Deterministic Drifts

Ruey-Shan Guo

Department of Industrial Management and Business Administration
National Taiwan University, Taipei 106, Taiwan

Argon Chen

Graduate Institute of Industrial Engineering,
National Taiwan University, Taipei 106, Taiwan

Jin-Jung Chen

Department of Mechanical Engineering
National Taiwan University, Taipei 106, Taiwan

ABSTRACT

During IC fabrication, the chemical-mechanical polishing (CMP) process usually suffers from a drift in the run-to-run removal rate due to the wear of the polishing pad. In this paper two run-to-run control schemes, PCC and d-EWMA schemes, for processes under deterministic drifts are presented. Control performance of both schemes is analytically derived and the determination of the optimal control parameters is provided. Validation results using the field CMP data demonstrate the control effectiveness of both control schemes.

Keywords: run-to-run control, EWMA, drift

INTRODUCTION

Manufacturing processes are often subject to special disturbances such as process drifts. For example, during IC fabrication, the chemical-mechanical polishing (CMP) process suffers from a drift in the run-to-run removal rate due to the wear of the polishing pad. Another example is the metal sputtering process in which the deposition rate decreases from run to run due to the consumption of the metal target. In such cases and if the following conditions apply, a run-to-run feedback control scheme will be useful:

- the causes of variation are known in advance,
- the causes of variation are difficult to remove,
- the adjustment of the process recipe is easy and fast,
- the adjustment cost is relatively inexpensive, and
- the quality loss due to process drifts is significant.

Since the feedback control action is exercised after observing the process output on a run-to-run basis, there is no input-output transient (dynamics) effect involved, as seen in the real-time control problems.

In recent years, the use of Exponentially Weighted Moving Average process controller (EWMA controller) for processes subject to small disturbances has been widely studied and adopted in practice. Box and Jenkin [1] proved

that the EWMA controller is a minimum mean square error (MMSE) controller when the underlying process disturbance follows an IMA(1,1) (first-order integrated moving average) process. In IC fabrication industry, several applications of the EWMA controller (also named as run-to-run or run-by-run process controller) are shown to be effective even when the underlying disturbances don't follow the IMA pattern but follow an approximately linear drift pattern [2]-[5].

Although effective in many applications, the EWMA controller is insufficient for processes subject to systematic drifts. Several authors [6][7] have addressed this problem and propose using two EWMA formulas: one for estimating "step-change" deviation and the other for estimating the process "drift" speed. Such a control scheme is referred to as Predictor Corrector Control (PCC) scheme, termed by Butler and Stefani [6].

The PCC approach provides a better solution than the EWMA approach for processes subject to deterministic drifts. However in the current literature, there is limited rigorous study on the PCC's performance and in many practices, the control parameters are decided empirically. To enhance the current understanding and performance on the run-to-run control schemes for CMP process subject to deterministic drifts, the goals of this papers are:

- to analyze the PCC control scheme,
- to propose and analyze an adjusted PCC control scheme which we refer as d-EWMA control scheme,
- to provide a methodology for determining the optimal control parameters of both PCC and d-EWMA schemes, and
- to characterize and compare their control performance using the field CMP data.

In the following sections, we will first review the EWMA control scheme, followed by the analysis of PCC scheme and d-EWMA scheme. The determination of the optimal control parameters for both schemes is then discussed. Data validation and conclusions are presented in the end.

REVIEW OF EWMA CONTROLLER

Suppose that the process output can be controlled linearly by an input variable and is subject to a natural process disturbance and a systematic disturbance δ_t . The process model can be written as:

$$\begin{aligned} Y_t &= \alpha + \beta X_t + \varepsilon_t + \delta_t \\ &= (\alpha + \delta_t) + \beta X_t + \varepsilon_t \end{aligned} \quad (1)$$

- Y_t : output measurement (t is the current run no.)
 α : process model constant term
 β : process model slope term
 X_t : equipment settings
 ε_t : white noise $\varepsilon_t \sim N(0, \sigma^2)$

To compensate for the systematic disturbances, an EWMA controller is used. Its main concepts are briefly explained below:

- a) Assume that the underlying process model in (1) is estimated by a control model with the following form:

$$\hat{Y}_t = a_{t-1} + bX_t \quad (2)$$

\hat{Y}_t : predicted output value

a_{t-1} : control model's constant term

b : control model's slope term (equal to β)

If we compare (1) with (2), we have $a \equiv \alpha + \delta_t$ providing that β is correctly estimated by b . This means that the systematic disturbances result in the changes of the constant term of the control model.

- b) Collect the output measurement Y_t after one run and calculate the residual:

$$e_t = Y_t - \hat{Y}_t \quad (3)$$

- c) Update the constant term using the EWMA algorithm. Notice that W is the control parameter to decide the weighting schemes on the historical data and $a_0 = \alpha$.

$$\begin{aligned} a_t &= W e_t + a_{t-1} \\ &= W(Y_t - bX_t) + (1-W)a_{t-1} \end{aligned} \quad (4)$$

- d) Obtain the equipment settings (recipe) for the next run.

$$X_{t+1} = \frac{T - a_t}{b} \quad (T \text{ is the target value}) \quad (5)$$

- e) Repeat steps b) to d) on a run-to-run basis.

Suppose the systematic disturbance follows a linear drift (output drifts away by a size of $c\sigma$ per unit of time), Ingolfson and Sachs [8] have shown that the long-run mean square error (MSE) of the EWMA controller is given by:

$$\begin{aligned} \lim_{t \rightarrow \infty} E[(Y_t - T)^2] &= \lim_{t \rightarrow \infty} Var(Y_t) + \lim_{t \rightarrow \infty} [E(Y_t) - T]^2 \\ &= \sigma^2 + \frac{W\sigma^2}{2-W} + \frac{(c\sigma)^2}{W^2} \end{aligned} \quad (6)$$

where

$$\lim_{t \rightarrow \infty} Var(Y_t) = \sigma^2 + \frac{W}{2-W} \sigma^2 \quad (7)$$

$$\lim_{t \rightarrow \infty} [E(Y_t) - T]^2 = \frac{(c\sigma)^2}{W^2} \quad (8)$$

The above equations show that for the EWMA controller, there is a tradeoff between the variance and bias: bias is minimized when $W=1$ while variance is minimized when $W=0$. Equation (8) also shows that EWMA controller is insufficient (biased) for processes subject to deterministic drifts.

PCC AND d-EWMA SCHEMES

PCC scheme

In the PCC scheme (Figure 1), there are two EWMA formula: one for estimating "step-change" deviation (a_t) and the other for estimating the process "drift" speed (p_t).

$$\begin{aligned} a_t &= W_1(e_t + p_{t-1}) + a_{t-1} \\ &= W_1(Y_t - bX_t) + (1-W_1)a_{t-1} \\ p_t &= W_2(Y_t - bX_t - a_{t-1}) + (1-W_2)p_{t-1} \end{aligned} \quad (9)$$

where W_1 and W_2 are the control parameters for the first and second EWMA equations respectively. The control model and the process recipe becomes:

$$\begin{aligned} \hat{Y}_t &= a_{t-1} + p_{t-1} + bX_t \\ X_{t+1} &= \frac{T - (a_t + p_t)}{b} \end{aligned} \quad (10)$$

To see how a_t and p_t work for a process subject to a linear drift, we can examine the process's steady state as time approaches infinity (Appendix A):

$$\begin{aligned} \lim_{t \rightarrow \infty} E(a_t) &= \alpha - \frac{c\sigma}{W_1} + c\sigma(t+1) \\ \lim_{t \rightarrow \infty} E(p_t) &= \frac{c\sigma}{W_1} \end{aligned} \quad (11)$$

This result somehow surprises us since in this PCC scheme a_t is not really an estimate for $\alpha + \delta_t$ and p_t is not really the estimate for the drifting speed $c\sigma$. The steady-state recipe becomes:

$$\begin{aligned} \lim_{t \rightarrow \infty} E(X_{t+1}) &= \frac{T - \left[\alpha - \frac{c\sigma}{W_1} + c\sigma(t+1) + \frac{c\sigma}{W_1} \right]}{b} \\ &= \frac{T - [\alpha + c\sigma(t+1)]}{b} \end{aligned} \quad (12)$$

The expected process output at time $t+1$ is then:

$$\lim_{t \rightarrow \infty} E(Y_{t+1}) = \lim_{t \rightarrow \infty} E(Y_t) = T \quad (13)$$

Here, we show that the system with PCC controller is a stable system and the expected process output converges to the desired target.

The estimate for $a_t + p_t$ in the PCC scheme can be further rewritten as (Appendix B):

$$\begin{aligned} a_t + p_t &= (W_1 + W_2 - W_1 W_2) \sum_{i=1}^t e_i + W_1 W_2 \sum_{i=1}^t \sum_{j=1}^i e_j + a_0 \\ &= w_I \sum_{i=1}^t e_i + w_{II} \sum_{i=1}^t \sum_{j=1}^i e_j + a_0 \end{aligned} \quad (14)$$

This is in effect an Integral-double-Integral (I-II) controller, in contrast to the Proportional-Integral-Derivative (PID) controller. In this I-II controller, the control action is proportional to the summation of the output errors and to the summation of summations of the output errors. The I-II controller can shown to be an MMSE controller for processes subject to IMA(2, 2) disturbances [1]. In (14), we can observe that the controller's integral constant is $w_I = W_1 + W_2 - W_1 W_2$ and the double-integral constant is $w_{II} = W_1 W_2$. We define a *control space* of the I-II controller as the space filled up by all possible settings of control constants (w_I, w_{II}) , as shown in Figure 2. In Figure 2, the controller allows both w_I and w_{II} to be set between 0 and 1. For PCC controller, we usually set W_1 and W_2 between 0 and 1 as well. This however leads to a limited control region in the (w_I, w_{II}) control space, as the shaded area shown in Figure 2.

d-EWMA scheme

Now, we would like to propose an adjustment of the PCC scheme such that the control region can be extended to the entire space and the steady-state estimates are cleaner. We refer to this adjusted PCC scheme as d-EWMA (double EWMA) scheme. Figure 1 illustrates the functional block diagram of the d-EWMA controller. Here, we have:

$$\begin{aligned} a_t &= W_1 e_t + a_{t-1} + p_{t-1} \\ &= W_1 (Y_t - bX_t) + (1 - W_1)(a_{t-1} + p_{t-1}) \end{aligned} \quad (15)$$

$$p_t = W_2 (Y_t - bX_t - a_{t-1}) + (1 - W_2) p_{t-1}$$

As can be observed in (15), the only difference from the original PCC formula is that in the first EWMA formula for a_t , we add p_{t-1} into the formula. As a result, the I-II controller becomes (Appendix C):

$$a_t + p_t = W_1 \sum_{i=1}^t e_i + W_2 \sum_{i=1}^t \sum_{j=1}^i e_j + (t+1)p_0 + a_0 \quad (16)$$

That is, $w_I = W_1$ and $w_{II} = W_2$ and we will have an I-II controller that fills up the entire control space.

With this d-EWMA formula, we can also obtain a cleaner steady state:

$$\begin{aligned} \lim_{t \rightarrow \infty} E(a_t) &= \alpha + c\sigma \\ \lim_{t \rightarrow \infty} E(p_t) &= c\sigma \end{aligned} \quad (17)$$

This result is much cleaner than the original PCC formula. Moreover, the two estimates a_t and p_t here have very clearly defined meanings. a_t in this d-EWMA scheme represents an estimate for $\alpha + \delta_t$ and p_t is an estimate for the size of the process drift from t to $t+1$. Together $a_t + p_t$ is the estimate for $\alpha + \delta_{t+1}$. The steady-state recipe at time $t+1$ becomes:

$$\begin{aligned} \lim_{t \rightarrow \infty} E(X_{t+1}) &= \lim_{t \rightarrow \infty} E \left[\frac{T - (a_t + p_t)}{b} \right] \\ &= \frac{T - [\alpha + c\sigma(t+1)]}{b} \end{aligned} \quad (18)$$

This is exactly the same as (12) and thus the process output with d-EWMA scheme will also converge to the desired target T . Though both PCC and d-EWMA schemes are "unbiased" schemes, the advantages of d-EWMA scheme over the PCC scheme are two folds:

- d-EWMA scheme is a direct form of I-II controller and its control space fills up the entire I-II control space.
- Unlike the PCC scheme, a_t and p_t in the d-EWMA scheme have clear definitions.

OPTIMAL CONTROL PARAMETERS

The determination of the optimal control parameters depends on a trade-off between steady state process variance and transient bias performance. Castillo [9] used this approach to determine the optimal control parameters for PCC scheme (although he used the term "double EWMA controller"). Here we extend the analysis to both PCC and d-EWMA schemes and define a different transient bias performance index.

Steady state variance

In Appendix A, we derive the equation for Y_t under the PCC scheme. It is then straightforward to derive the steady state variance of Y_t under the PCC scheme:

$$\begin{aligned} \lim_{t \rightarrow \infty} \text{Var}(Y_t) &= \\ & \left(1 + \frac{W_1^2 W_2^2 - 3W_1 W_2 (W_1 + W_2) + 2(W_1^2 + W_2^2) + 6W_1 W_2}{(2 - W_1)(2 - W_2)(W_1 + W_2 - W_1 W_2)} \right) \sigma^2 \end{aligned} \quad (19)$$

Figure 3 shows the relationship as defined in (19). As can be seen, steady state variance increases as W_1 and W_2 increase.

With similar approach, we can also derive the steady state variance of Y_t under the d-EWMA scheme:

$$\lim_{t \rightarrow \infty} \text{Var}(Y_t) = \left(1 + \frac{W_1(2W_1 + W_2) + 2W_2}{W_1(4 - 2W_1 - W_2)} \right) \sigma^2 \quad (20)$$

Figure 4 shows the relationship as defined in (20).

Transient bias performance

Castillo [9] has defined average mean square deviation statistic as the transient bias performance index:

$$\overline{MSD} = \frac{1}{m} \sum_{t=1}^m (E[Y_t] - T)^2 \quad (21)$$

The decision of m , however, will cause some difficulty in real applications. First, the decision of m might be empirical. Second, the mean square deviation statistic must be calculated numerically and there is no closed-form solution. To overcome this difficulty, we propose a better performance index:

$$SSD = \sum_{t=1}^{\infty} (E[Y_t] - T)^2 \quad (22)$$

Notice that under the PCC and d-EWMA schemes, $E[Y_t]$ will be equal to T as t approaches infinity. As a result, components of SSD mostly come from the transient bias. With the new definition, we can derive SSD for both PCC (23) and d-EWMA (24) schemes.

$$SSD = \frac{(2 - W_1 - W_2 + W_1 W_2) c^2 \sigma^2}{W_1 W_2 (2 - W_1)(2 - W_2)(W_1 + W_2 - W_1 W_2)} \quad (23)$$

$$SSD = \frac{(2 - W_1) c^2 \sigma^2}{W_1 W_2 (4 - 2W_1 - W_2)} \quad (24)$$

Determination of optimal control parameters

As mentioned earlier, the determination of the optimal control parameters depends on a trade-off between steady state process variance and transient bias performance.

One approach is to minimize the transient bias under an upper bound of steady state process variance constraint. The formulation is shown below:

$$\min_{W_1, W_2} SSD \quad \text{st.} \quad \lim_{t \rightarrow \infty} Var(Y_t) \leq (1 + V) \sigma^2 \quad (25)$$

Result for the PCC scheme is shown in (26) and illustrated in Figure 3 (solid line).

$$W_1 = W_2 \quad (26)$$

Result for the d-EWMA scheme is shown in (27) and illustrated in Figure 4 (solid line).

$$4(-W_1^2 W_2 - 4W_2 + 4W_1 W_2 + W_2^2) + (2 - W_1)(2 - W_1 - W_2)(4W_1^2 - 4W_2 + 4W_1 W_2 + W_2^2) = 0 \quad (27)$$

It is interesting to notice that the determination of the optimal control parameters is independent of the degree of the drift speed $c\sigma$ in both schemes. Figures 5 and 6 also show the values of control parameters under different constraint of steady state variance. By substituting these control parameters into (23)(24), we can obtain the optimal transient bias performance versus steady state variance (Figure 7).

The second approach to determine the optimal control parameters is to minimize an objective function based on a weighted sum of steady state variance and transient bias performance:

$$\min_{W_1, W_2} \left[k \lim_{t \rightarrow \infty} Var(Y_t) + (1 - k)SSD \right] \quad (28)$$

where k is the weighting factor.

By substituting (19)(20)(23)(24) into (28), we can obtain the objective functions for PCC and d-EWMA schemes respectively. Results are shown in Figures 8 and 9. Again, the optimal W_1 and W_2 will be equal to each other in the PCC scheme. In the d-EWMA scheme, optimal W_1 is usually highly than optimal W_2 .

DATA VALIDATION AND CONCLUSIONS

To demonstrate the effectiveness of both PCC and d-EWMA schemes, field CMP data are used. A typical CMP process consists of a wafer carrier, a rotating platen with a replaceable abrasive pad mounted on the surface, and a pad conditioner (Figure 10). Because the process includes mechanical abrasion of the surface, the polishing pad wears rapidly, which results in a decreasing and irregular run-to-run process removal rate.

A simple strategy for controlling the CMP process is to predict the run-to-run process removal rate and then adjust the polishing time based on the prediction [3]. Based on the PCC scheme derived in (9), the PCC prediction of removal rate can be expressed as:

$$\begin{aligned} a_t &= W_1 R_t + (1 - W_1) a_{t-1} \\ p_t &= W_2 (R_t - a_{t-1}) + (1 - W_2) p_{t-1} \\ \hat{R}_{t+1} &= a_t + p_t \end{aligned} \quad (29)$$

Based on the d-EWMA scheme derived in (15), the d-EWMA prediction of removal rate can be expressed as:

$$\begin{aligned} a_t &= W_1 R_t + (1 - W_1)(a_{t-1} + p_{t-1}) \\ p_t &= W_2 (R_t - a_{t-1}) + (1 - W_2) p_{t-1} \\ \hat{R}_{t+1} &= a_t + p_t \end{aligned} \quad (30)$$

Results of applying the EWMA, PCC and d-EWMA schemes to the CMP process removal rate prediction are shown in Figure 11 and Table 1. As can be seen, the prediction improvement over the EWMA scheme is about 38% for PCC scheme and 57% for the d-EWMA scheme. Both schemes have chosen the optimal control parameters based on the minimum transient bias approach.

In summary, in this paper we present PCC and d-EWMA schemes for CMP process subject to deterministic drifts. Control performance of both schemes is analytically derived and the determination of the optimal control parameters is provided. Validation results using the filed CMP data demonstrate the control effectiveness of both control schemes.

ACKNOWLEDGMENTS

We would like to thank Dr. Jowei Dun and Mr. S.-A. Wu of TSMC who provided the CMP data for this study. This research was supported in part by the National Science Council of Republic of China under Grant NSC88-2416-H-002-055.

REFERENCES

- [1] G. Box and M. Jenkins, *Time Series Analysis - Forecasting and Control*. Oakland, CA:Holden-Day, 1974.
- [2] E. Sachs, A. Hu, and A. Ingolfsson, "Run by Run Process Control; Combining SPC and Feedback Control," *IEEE Transactions Semiconductor Manufacturing*, Vol. 8, No. 1, pp. 26-43, February 1995.
- [3] D. Boning, W. Moyne, T. Smith, J. Moyne, R. Telfeyan, A. Hurwitz, S. Shellman, and J. Taylor, "Run by Run Control of Chemical-Mechanical Polishing," *IEEE Transactions on Components, Packaging, and Manufacturing technology - Part C*, Vol.19, No.4, pp. 307-313, October 1996.
- [4] T. Smith, and D. Boning, "A Self-Tuning EWMA Controller Utilizing Artificial Neural Network Function Approximation Techniques," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, pp. 121-132, March 1997.
- [5] E.D. Castillo, and A. Hurwitz, "Run-to-Run Process Control: Literature Review and Extensions," *Journal of Quality Technology*, Vol. 29, No. 2, pp. 184-196, April 1997.
- [6] S. Bulter and J. Stefani, "Supervisory Run-to-Run Control of Polysilicon Gate Etch Using *In Situ* Ellipsometry," *IEEE Transactions Semiconductor Manufacturing*, Vol. 7, No. 2, pp. 193-201, May 1994.
- [7] T. Smith, D. Boning, J. Stefani, and S. Butler, "Run by Run Advanced Process Control of Metal Sputter Deposition," *IEEE Transactions on Semiconductor Manufacturing*, Vol.11, No.2, pp. 276-284, May 1998.
- [8] A. Ingolfsson and E. Sachs, "Stability and Sensitivity of an EWMA Controller," *Journal of Quality Technology*, Vol. 25, No. 4, pp. 271-287, October 1993.
- [9] E. D. Castillo, "Long Run and Transient Analysis of a Double EWMA Feedback Controller," *IIE Transactions*, Vol. 31, pp. 1157-1169, December 1999.

APPENDIX A

PCC controller:

$$\begin{aligned} a_t &= W_1(Y_t - bX_t) + (1 - W_1)a_{t-1} \\ &= W_1(\alpha + c\sigma + \varepsilon_t) + (1 - W_1)a_{t-1} \\ p_t &= W_2(Y_t - bX_t - a_{t-1}) + (1 - W_2)p_{t-1} \\ &= W_2(\alpha + c\sigma + \varepsilon_t - a_{t-1}) + (1 - W_2)p_{t-1} \end{aligned}$$

Using matrix expression, we obtain:

$$\begin{bmatrix} a_t \\ p_t \end{bmatrix} = \begin{bmatrix} 1 - W_1 & 0 \\ -W_2 & 1 - W_2 \end{bmatrix} \begin{bmatrix} a_{t-1} \\ p_{t-1} \end{bmatrix} + \begin{bmatrix} W_1\alpha + W_1\varepsilon_t + W_1c\sigma \\ W_2\alpha + W_2\varepsilon_t + W_2c\sigma \end{bmatrix}$$

$$\text{Let } U(t) = \begin{bmatrix} a_t \\ p_t \end{bmatrix}, A = \begin{bmatrix} 1 - W_1 & 0 \\ -W_2 & 1 - W_2 \end{bmatrix},$$

$$B(t) = \begin{bmatrix} W_1\alpha + W_1\varepsilon_t + W_1c\sigma \\ W_2\alpha + W_2\varepsilon_t + W_2c\sigma \end{bmatrix}$$

$$\text{Then, } U(0) = \begin{bmatrix} \alpha \\ 0 \end{bmatrix},$$

$$A = \begin{bmatrix} 1 - W_1 & 0 \\ -W_2 & 1 - W_2 \end{bmatrix} = \begin{bmatrix} 1 - W_1 & 0 \\ 0 & 1 - W_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{W_2}{W_2 - W_1} & 1 \end{bmatrix}$$

We obtain:

$$\begin{aligned} A^t &= \begin{bmatrix} 1 - W_1 & 0 \\ -W_2 & 1 - W_2 \end{bmatrix}^t \begin{bmatrix} 1 & 0 \\ \frac{W_2}{W_2 - W_1} & 1 \end{bmatrix} \\ &= \begin{bmatrix} (1 - W_1)^t & 0 \\ \frac{W_2}{W_2 - W_1} \left((1 - W_2)^t - (1 - W_1)^t \right) & (1 - W_2)^t \end{bmatrix} \end{aligned}$$

and, thus

$$\begin{aligned} U(t) &= AU(t-1) + B(t) \\ &= A[AU(t-2) + B(t-1)] + B(t) \end{aligned}$$

$$= A^t U(0) + \sum_{i=1}^t A^{t-i} B(i)$$

$$= \begin{bmatrix} (1 - W_1)^t & 0 \\ \frac{W_2}{W_2 - W_1} \left((1 - W_2)^t - (1 - W_1)^t \right) & (1 - W_2)^t \end{bmatrix} \begin{bmatrix} \alpha \\ 0 \end{bmatrix}$$

$$+ \sum_{i=1}^t \begin{bmatrix} (1 - W_1)^{t-i} & 0 \\ \frac{W_2}{W_2 - W_1} \left((1 - W_2)^{t-i} - (1 - W_1)^{t-i} \right) & (1 - W_2)^{t-i} \end{bmatrix}$$

$$\begin{bmatrix} W_1\alpha + W_1\varepsilon_i + W_1c\sigma_i \\ W_2\alpha + W_2\varepsilon_i + W_2c\sigma_i \end{bmatrix}$$

Finally,

$$a_t = \alpha(1-W_1)^t + \sum_{i=1}^t (1-W_1)^{t-i} (W_1\alpha + W_1\varepsilon_i + W_1c\sigma i)$$

$$p_t = \frac{\alpha W_2}{W_2 - W_1} \left((1-W_2)^t - (1-W_1)^t \right) + \sum_{i=1}^t \left\{ \begin{array}{l} \frac{W_2}{W_2 - W_1} \left((1-W_2)^{t-i} - (1-W_1)^{t-i} \right) \\ (W_1\alpha + W_1\varepsilon_i + W_1c\sigma i) \\ + (1-W_2)^{t-i} (W_2\alpha + W_2\varepsilon_i + W_2c\sigma i) \end{array} \right\}$$

$$Y_t = \alpha + bX_t + c\sigma + \varepsilon_t$$

$$= T + \frac{(1-W_1)^t - (1-W_2)^t}{W_1 - W_2} c\sigma + \varepsilon_t$$

$$+ \sum_{i=1}^{t-1} \frac{W_1^2 (1-W_1)^i - W_2^2 (1-W_2)^i}{W_1 - W_2} \varepsilon_{t-i}$$

As t approaches infinity:

$$\lim_{t \rightarrow \infty} E(a_t) = \alpha - \frac{c\sigma}{W_1} + c\sigma(t+1)$$

$$\lim_{t \rightarrow \infty} E(p_t) = \frac{c\sigma}{W_1}$$

APPENDIX B

$$a_t = W_1(e_t + p_{t-1}) + a_{t-1} = W_1 \sum_{i=1}^t e_i + W_1 \sum_{i=0}^{t-1} p_i + a_0$$

$$p_t = W_2 e_t + p_{t-1} = W_2 \sum_{i=1}^t e_i + p_0$$

$$a_t + p_t = (W_1 + W_2 - W_1 W_2) \sum_{i=1}^t e_i + W_1 W_2 \sum_{i=1}^t \sum_{j=1}^i e_j + a_0$$

$$= W_I \sum_{i=1}^t e_i + W_{II} \sum_{i=1}^t \sum_{j=1}^i e_j + a_0$$

APPENDIX C

$$a_t = W_1 e_t + p_{t-1} + a_{t-1} = W_1 \sum_{i=1}^t e_i + \sum_{i=0}^{t-1} p_i + a_0$$

$$p_t = W_2 e_t + p_{t-1} = W_2 \sum_{i=1}^t e_i + p_0$$

$$a_t + p_t = (W_1 + W_2) \sum_{i=1}^t e_i + \sum_{i=0}^{t-1} p_i + p_0 + a_0$$

$$= (W_1 + W_2) \sum_{i=1}^t e_i + \sum_{i=0}^{t-1} \left(w_2 \sum_{j=1}^i e_j + p_0 \right) + a_0 + p_0$$

$$= (W_1 + W_2) \sum_{i=1}^t e_i + W_2 \sum_{i=0}^{t-1} \sum_{j=1}^i e_j + (t+1)p_0 + a_0$$

$$= W_1 \sum_{i=1}^t e_i + W_2 \sum_{i=1}^t \sum_{j=1}^i e_j + (t+1)p_0 + a_0$$

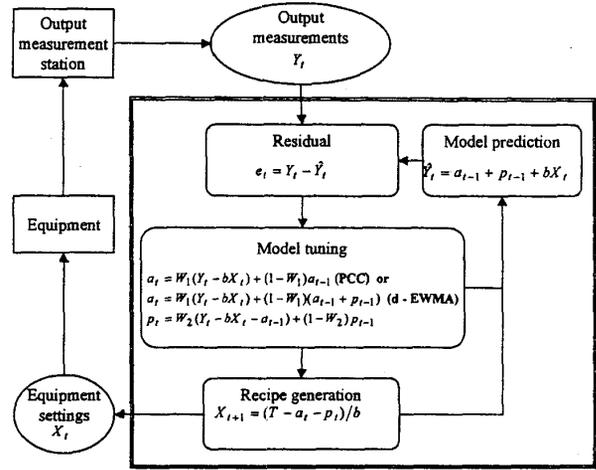


Figure 1. Functional block diagram of PCC and d-EWMA schemes

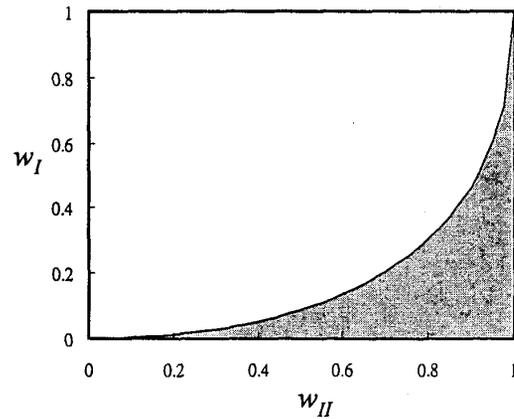


Figure 2. I-II controller's control space for PCC scheme

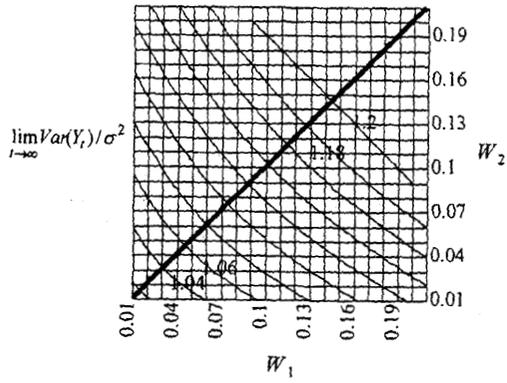


Figure 3. Performance characterization of PCC scheme.

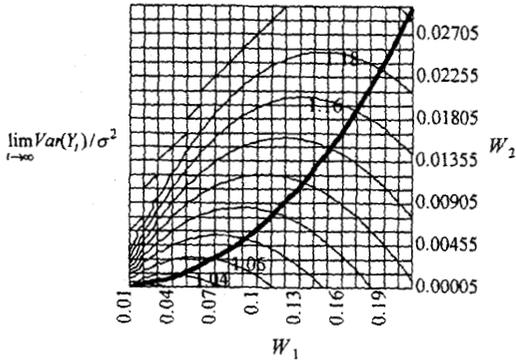


Figure 4. Performance characterization of d-EWMA scheme.

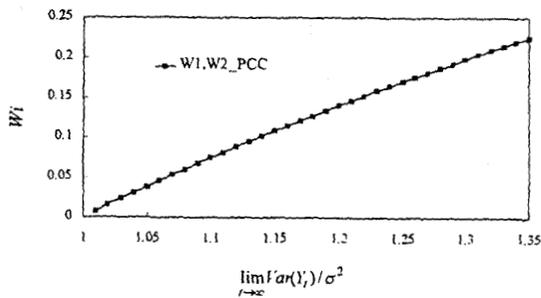


Figure 5. Optimal control parameters of PCC scheme under different steady state performance.

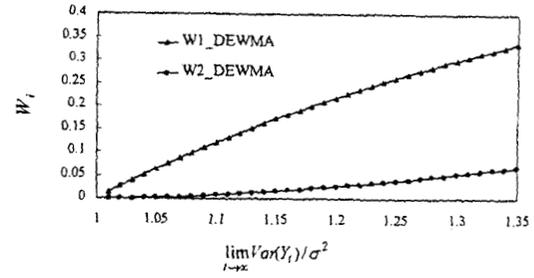


Figure 6. Optimal control parameters of d-EWMA scheme under different steady state performance.

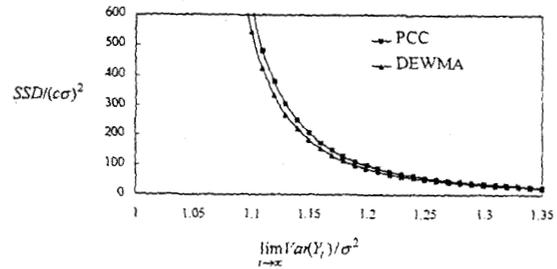


Figure 7. Transient bias performance versus steady state performance.

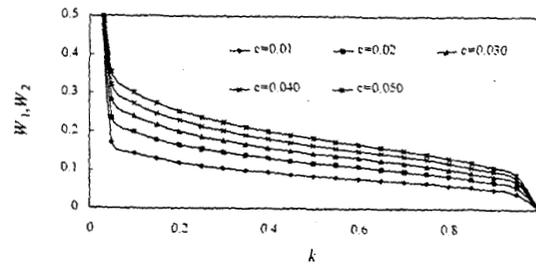


Figure 8. Optimal control parameters of PCC scheme under different k and c .

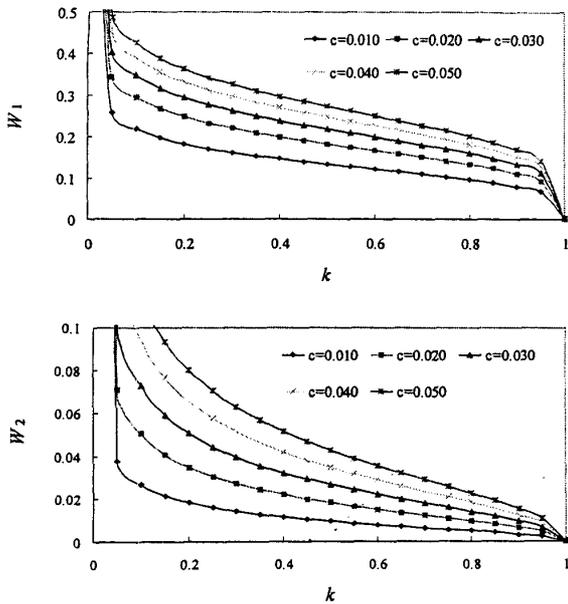


Figure 9. Optimal control parameters of d-EWMA scheme under different k and c .

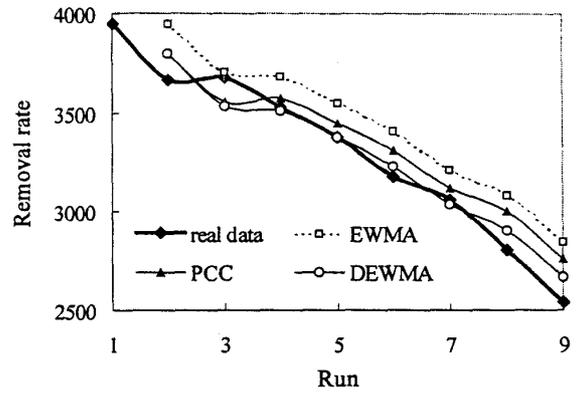


Figure 11. Comparison of CMP removal rate prediction.

	Prediction Root MSE	Improvement (%) over EWMA
EWMA	217.94	-
PCC	135.81	38%
d-EWMA	92.96	57%

Table 1. Comparison of CMP removal rate prediction.

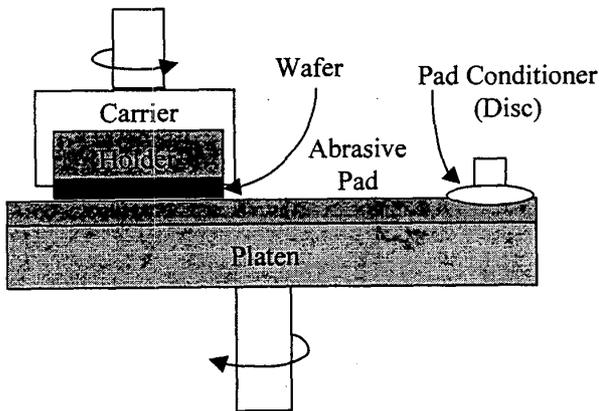


Figure 10. Schematic of CMP process.