A Generalized Quasi-MMSE Controller for Run-to-Run Dynamic Models

Sheng-Tsaing Tseng and Pei-Yu Chen
Institute of Statistics, National Tsing-Hua University, Hsin-Chu, Taiwan

ABSTRACT
This study proposes a generalized quasi-minimum mean square error (qMMSE) controller for implementing a run-to-run process control where the process input–output relationship follows a general-order dynamical model with added noise. The expression of the process output, the long-term stability conditions and the optimal discount factor of this controller are derived analytically. Furthermore, we use the proposed second-order dynamical model to illustrate the effects of mis-identification of the process I-O model on the process total mean square error (TMSE). Via a comprehensive simulation study, the model demonstrates that the TMSE may inflate by more than 150% if a second-order dynamical model with moderately large carryover effects is wrongly identified as that of a first-order model. This means that the effects of mis-identification of the process I-O model on the process total mean square error (TMSE) is not negligible for implementing a dynamic run-to-run (RTR) process control. Supplementary materials for this article are available online.

1. Introduction
Run-to-run (RTR) process control techniques are frequently used in semiconductor manufacturing operations (Sachs, Hu, and Ingolfsson 1995; Hamby, Kabamba, and Khargonekar 1998; Chen and Guo 2001; Moyne et al. 2002). These techniques consist of the predicted model and the feedback control scheme. First, design of experiment (DOE) and response surface methodology (RSM) techniques are employed to construct the predicted model using off-line stage data. At the end of the production run, to adjust the process output to a desired target, the feedback control scheme operates and creates the new input recipe for the next run by comparing the output value with the target. RTR techniques have been successfully applied to several semiconductor manufacturing processes such as photolithography process, chemical mechanical polishing (CMP) process, reactive ion etching (RIE), etc. Figure 1 shows a flowchart of an online RTR feedback control of a single-input and single-output (SISO) process. The key idea of RTR control is to adjust the process output to its desired target.

Process dynamics and process disturbance are very common in a complex semiconductor manufacturing process. Specifically, the behavior of process output will be changed by the carryover. For example, in modeling a CMP process, the thickness removed, is dependent on the polish time of the current run $x_t-1$ as well as the polish times of previous runs $x_{t-2}$, $x_{t-3}$, ..., (or/and the output $Y_{t-1}$, $Y_{t-2}$, ...), since the degree of wear of the pad is dependent on the pass polish time as well as the hardness of path samples being processed. Therefore, we use $f(x_{t-1}, x_{t-2}, \ldots)$ to characterize the process dynamics, that is, carry over effects of the plant. Moreover, the disturbance may be due to the degradation of the health of machine that changes from run to run, for example, coagulation of particles in the polish fluids, etc. A color noise $N_t$ is adopted to characterize the effects of process disturbance. Once the process I-O model is specified, the corresponding predicted model and feedback controller can be obtained accordingly.

1.1 Related Literature on Static Model-Based Controller

Considering an SISO process, Ingolfsson and Sachs (1993) proposed a single EWMA (sEWMA) controller for adjusting the process output, where the process model is assumed to be deterministic and first-order static model. They derived the stability conditions of the sEWMA controller and demonstrated that the expected process output would not converge to a desired target if the process exhibits a linear drift. To overcome this weakness, Butler and Stefani (1994) proposed a double EWMA (dEWMA) controller, in which the offset term of the process output can be eliminated efficiently. Del Castillo (2001) and Tseng, Chou, and Lee (2002) also investigated long-term stability conditions and transient performance and determined the optimal discount factor for the dEWMA controller.

1.2 Related Literature on Dynamic Model-Based Controller

The aforementioned studies assumed that the dynamic behavior of the process output is only due to disturbance dynamics, not to process dynamics. The reasons may be because the allowable range for the adjustment of input recipe (in semiconductor manufacturing operation) may be very narrow. Therefore, a dynamic model was simply approximated by a static model and move all carryover terms to the disturbance with a color noise. Generally, the process dynamics has been proven and verified...
in the literature (Capilla et al. 1999; Del Castillo 2002; Fan et al. 2002; Jen and Jiang 2008; Tseng and Lin 2009; Qiu and Xiang 2014). More specifically, Pan and Del Castillo (2001) investigated the issue of identification and fine tuning of closed-loop dynamic models under discrete sEWMA and PI adjustments. Fan et al. (2002) proposed a triple EWMA (tEWMA) controller to achieve the goal of compensating the positive autocorrelation in the observed response. Those studies reveal the importance of process dynamics. Focusing on the case with the process input–output model following a combined first-order transfer function model with an ARIMA (p, d, q) process disturbance (Box, Jenkins, and Reinsel 1994), Tseng and Lin (2009) addressed the issue of stability conditions for the sEWMA controller. The result demonstrates that the conventional sEWMA controller may not ensure the stability conditions when the process dynamics are significant. Furthermore, even when the long-term stability of the sEWMA controller is valid, its transient behavior may suffer from a low convergent rate to the desired target. Recently, Tseng and Mi (2014) proposed a quasi-minimum mean square error (qMMSE) controller when the process I-O relationship follows a first-order dynamic model. The analytical results of the long-term stability conditions and control performance of this controller are addressed comprehensively. Nevertheless, the results rely on the assumption that the process I-O relationship follows a first-order dynamic model, which may be over-simplified to characterize a complicated dynamic model.

1.3 The Goal and Layout of This Study

To overcome the aforementioned difficulty, this study first proposed a generalized quasi-minimum mean square error (qMMSE) controller when the process I-O model follows a general-order dynamical model with added noise. An analytical expression of the controlled output and long-term stability conditions will be provided. Furthermore, we also derive the optimal discount factor of this controller and use some examples to compare the control performance of the proposed controller compared with the existing controllers.

This article is organized as follows. Section 2 presents the problem formulation. Section 3 addresses the analytical expression of the process output. Section 4 addresses the stability conditions of the proposed controller and determines the optimal discount factors that are useful to compare the control performance of the proposed controller with some competing controllers. Section 5 illustrates the proposed method using an illustrative example. Section 6 provides the comprehensive result of a comparison of the control performances among the proposed and existing popular controllers. Finally, Section 7 offers concluding remarks and future research topics.

2. A Generalized Quasi-MMSE Controller

Assuming that \( x_{t-1} \) and \( Y_t \) denote the input recipe and output response at time \( t \), respectively, the input–output relationship of an SISO process is postulated in terms of the following dynamic model (Del Castillo 2002):

\[
\delta_r(B) Y_t = \alpha + \gamma_r(B)x_{t-r} + \eta_t,
\]

where \( \delta_r(B) = (1 - \delta_1 B - \delta_2 B^2 - \cdots - \delta_r B^r) \), \( \gamma_r(B) = (\gamma_0 - \gamma_1 B - \gamma_2 B^2 - \cdots - \gamma_r B^r) \), \( B \) is a backshift operator and \( \alpha \) is the intercept parameter of the process; \( \delta_1, \ldots, \delta_r \) and \( \gamma_1, \ldots, \gamma_r \) are the carryover effects for the output responses and input recipes, respectively. Furthermore, we assume that \( \eta_t \) (process disturbance) follows an ARIMA \((p, d, q)\) series. In other words,

\[
\phi_p(B)(1 - B)^d \eta_t = \theta_q(B) \varepsilon_t,
\]

where \( \phi_p(B) = (1 - \sum_{i=1}^{p} \phi_i B^i) \) and \( \theta_q(B) = (1 - \sum_{j=1}^{q} \theta_j B^j) \). \( \phi_i, \theta_j \) are autoregressive and moving average parameters, respectively. The roots of \( \delta_r(B) = 0 \) and \( \phi_p(B) = 0 \) are outside the unit circle; \( \varepsilon_t \) is a white noise series with mean 0 and variance \( \sigma^2 \).

Assume that Equation (1) can also be expressed as a combined transfer function-noise model (Box, Jenkins, and Reinsel 1994) as follows:

\[
Y_t = (\delta_r(B))^{-1}(\alpha + \gamma_r(B)x_{t-r}) + (\delta_r(B))^{-1}\eta_t.
\]

Let \( \tau \) denote the desired target of process outputs. For \( d = 1, 2, \ldots \), the expression of MMSE controller is given in Equation (A1) of Appendix A, where the parameters\( \{\hat{\delta}_1\}_{i=1}, \{\hat{\gamma}_1\}_{j=1}, \{\hat{\phi}_i\}_{i=1}^p, \) and \( \{\hat{\theta}_j\}_{j=1}^q \) are unknown. Assume that these parameters are estimated by a combined transfer function-noise model (by using the off-line stage data), and denoted by \( \{\hat{\delta}_1\}_{i=1}, \{\hat{\gamma}_1\}_{j=1}, \{\hat{\phi}_i\}_{i=1}^p, \) and \( \{\hat{\theta}_j\}_{j=1}^q \). Plugging these estimated into Equation (A1), the generalized quasi-MMSE controller can be expressed as follows:

\[
(1 - B)^d x_{t-r+1} = \frac{1}{B} \left( \frac{\hat{\delta}_r(B)}{\hat{\gamma}_r(B)} - (1 - B)^d \left( \hat{\delta}_1 B + \cdots + \hat{\delta}_r B^r \right) \right)(Y_t - \tau),
\]

where \( \hat{\phi}_i(B) = 1 - \sum_{i=1}^{p} \hat{\phi}_i B^i \) and \( \hat{\theta}_j(B) = 1 - \sum_{j=1}^{q} \hat{\theta}_j B^j \). This controller can be directly applied to adjust a dynamic process with general order. Generally speaking, its control performance is not easy to address because the analytic expression of controlled output will be very complicated. Fortunately, a class of nonstationary ARIMA \((1,1,1)\) series is one of the most widely used models to describe the disturbance behavior of various run-to-run control problems, which has been appeared in the literature of RTR control (Capilla et al. 1999; Del Castillo 2002; Fan et al. 2002; Jen and Jiang 2008). In addition, for any time \( t \), the input recipe \( (x_{t-1}) \) will affect the process output.
(Y_t) immediately. Thereafter, by adopting (p, d, q) = (1, 1, 1), b = 1, and setting the discount factor \( \omega = (1 - \theta_t) \) in Equation (3), the controller can be expressed as

\[
(1 - B) x_t = \frac{\omega + \delta_t + \sum_{j=2}^{\infty} (\delta_j - \delta_{j-1}) B^{j-1} - \delta_j B^r + \phi_t (1 - B) (1 - \sum_{j=1}^{\infty} \delta_j B^r)}{(\gamma_0 - \gamma_1 B - \cdots - \gamma_r B^r) (1 - \phi_t B)} \times (Y_t - \tau).
\]

(4)

Hereafter, we will simply call this controller a generalized qMMSE controller with order (r, s). The proposed controller has the advantage of using all previous information in process outputs to adjust the new process input. Note that recursively iterating \((x_t, x_{t-1})\) in Equation (4) shows that the proposed controller uses not only the current data, \((Y_t - \tau)\), but also all of the previous data, \([(Y_j - \tau)]_{j=1}^{n=1}\), to adjust the input recipes. As a special case, the proposed controller in Equation (4) reduces to the conventional single EWMA controller if \(r = s = 0\). Under this simplified scenario, Equations (1) and (4) reduce to

\[
Y_t = \alpha + \gamma_0 x_{t-1} + \eta_t, \tag{5}
\]

and

\[
x_t = x_{t-1} - \frac{\omega}{\gamma_0} (Y_t - \tau). \tag{6}
\]

Furthermore, the proposed controller reduces to the qMMSE controller proposed by Tseng and Mi (2014) if \(r = s = 1\), and Equations (1) and (4) can be expressed as

\[
Y_t = \delta_1 Y_{t-1} + \alpha + \gamma_0 x_{t-1} - \gamma_1 x_{t-2} + \eta_t, \tag{7}
\]

and

\[
x_t = x_{t-1} - \frac{\omega + \delta_t - \delta_j B^r}{\gamma_0 - \gamma_1 B} (Y_t - \tau). \tag{8}
\]

3. Analytical Expression of Controlled Output

From Equation (4), and by setting \(t = t - 1\), we have

\[
\nabla x_{t-1} = -\frac{1}{\gamma_0} \left[ \omega + \delta_1 + \sum_{j=2}^{r} (\delta_j - \delta_{j-1}) B^{j-1} - \delta_j B^r + \phi_t (1 - B) \left( 1 - \sum_{j=1}^{r} \delta_j B^j \right) \right] (Y_{t-1} - \tau)

+ \frac{1}{\gamma_0} \left[ \gamma_1 + \sum_{j=1}^{s} \gamma_j B^{j-1} + \phi_t (\gamma_0 - \sum_{j=1}^{s} \gamma_j B^j) \right] \nabla x_{t-2}, \tag{9}
\]

By taking the difference of the process I-O model stated in Equation (1), we have

\[
Y_t - \tau = \left[ (1 + \delta_1) + (\delta_2 - \delta_1) B + \cdots + (\delta_r - \delta_{r-1}) B^{r-1} - \delta_j B^r \right] \nabla x_{t-1} + \nabla \eta_t, \tag{10}
\]

Note that Equation (9) can be further expressed as follows:

\[
\nabla x_{t-1} = \sum_{i=1}^{r+2} a_i (Y_{t-1} - \tau) + \sum_{l=r+3}^{r+s+3} a_j \nabla x_{t-2} - \tau, \tag{11}
\]

where

\[
a_j = \begin{cases} \frac{1}{\gamma_0} \left( \omega + \delta_1 + \phi_t \right), & j = 1, \\ \frac{1}{\gamma_0} (\delta_j - \delta_{j-1} - \phi_t \delta_1 - \phi_t \delta_1), & j = 2, \\ \frac{1}{\gamma_0} \left( \delta_j - \delta_{j-1} - \phi_t (\delta_j - \delta_{j-2}) \right), & 3 \leq j \leq r, \\ \frac{1}{\gamma_0} \left( \delta_r + \phi_t \delta_r - \phi_t \delta_{r-1} \right), & j = r + 1, \\ \frac{1}{\gamma_0} (\gamma_1 + \phi_t), & j = r + 3, \\ \frac{1}{\gamma_0} (Y_{t-1} + \phi_t \gamma_0), & j = r + 3, \\ \frac{1}{\gamma_0} \left( \gamma_j + \phi_t \gamma_j \right), & j = r + 2, r + s + 3, \\ \frac{1}{\gamma_0} \left( \gamma_{j-1} + \phi_t \gamma_{j-1} \right), & r + 4 \leq j \leq r + s + 2, \\ -\frac{1}{\gamma_0} \phi_t \gamma_{r+3}, & j = r + s + 3. \end{cases}
\]

In addition, substituting \(\nabla x_{t-1}\) in Equation (9) into Equation (10), we then have

\[
Y_t - \tau = \sum_{j=1}^{r+2} c_j (Y_{j-1} - \tau) + \sum_{l=r+3}^{r+s+3} c_j \nabla x_{t-2} - \tau + \nabla \eta_t, \tag{12}
\]

where

\[
c_j = \begin{cases} 1 + \delta_1 + \gamma_0 a_j, & j = 1, \\ \delta_j - \delta_{j-1} + \gamma_0 a_j, & 2 \leq j \leq r, \\ -\delta_r + \gamma_0 a_{r+1}, & j = r + 1, \\ \gamma_0 a_j, & j = r + 2, r + s + 3, \\ -\gamma_{j-1} + \gamma_0 a_j, & r + 4 \leq j \leq r + s + 2. \end{cases}
\]

The proof of Equation (12) is given in Appendix B. Let

\[
S_4^0 = \left( \begin{array}{ccccc} Y_t - \tau, & \ldots, & Y_{t-r-1} - \tau & \nabla x_{t-1}, & \ldots, & \nabla x_{t-(r+1)} \end{array} \right)^t,
\]

\[
S_4^0 = \left( \begin{array}{cccc} \alpha + \gamma_0 x_t - \tau, & \ldots, & -\tau & \nabla x_{t-1}, & \ldots, & 0 \end{array} \right)^t,
\]

and \(Q\) is a square matrix of order \((r+s+3)\) as follows:

\[
Q = \begin{bmatrix} c_1 & \ldots & c_{r+1} & c_{r+2} & \ldots & c_{r+s+2} & c_{r+s+3} \\ 1 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \ldots & 1 & 0 & \ldots & 0 & 0 \\ a_1 & \ldots & a_{r+1} & a_{r+2} & \ldots & a_{r+s+2} & a_{r+s+3} \\ 0 & \ldots & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & 1 & 0 \end{bmatrix}.
\]
Note that \( Z_1^*, Z_2^*, \) and \( S_1^* \) are the column vectors with order \((r+s+3)\). Combining Equations (11) and (12), we have the following recursive formula:

\[
Z_t^* = QZ_{t-1}^* + S_t^*,
\]

By using an iterative procedure, we have

\[
Z_t^* = Q^{t-1}Z_1^* + \sum_{i=0}^{t-1} Q^i S_{t-i}.
\]

Hence, the deviance between the process output and the target can be partitioned into two parts: a nonrandom part \( \Gamma_{t-1} \) and a random part \( W_t \); and we have the following lemma.

**Lemma 1.** The process output controlled by a generalized qMMSE controller with order \((r,s)\) can be expressed as follows:

\[
Y_t - \tau = \Gamma_{t-1} + W_t, \quad t \geq 1,
\]

where

\[
\Gamma_{t-1} = \left( \begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \end{array} \right) Q^{t-1} Z_1^*,
\]

and

\[
W_t = \left( \begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \end{array} \right) \left( \sum_{i=0}^{t-1} Q^i S_{t-i} \right).
\]

This result is very useful for studying the stability conditions of the generalized qMMSE controller in the next section.

### 4. Stability Conditions and Optimal Discount Factor

A process is said to be asymptotically stable if the expectation of the process output converges to the desired target and its asymptotic variance remains bounded:

\[
\lim_{t \to \infty} E(Y_t) = \tau,
\]

and

\[
\lim_{t \to \infty} \text{Var}(Y_t) < \infty.
\]

Equations (18) and (19) are known as the stability conditions.

From the definition of stability conditions stated in Equations (18) and (19), it is revealed that the stability conditions can be attained when \( \lim_{t \to \infty} \Gamma_{t-1} = 0 \) and \( \lim_{t \to \infty} \text{Var}(W_t) < \infty \).

Define the maximum absolute values of the eigenvalues of matrix \( Q \) as follows:

\[
\rho(Q) = \max \{|\lambda| \mid Qx = \lambda x, \ x \neq 0\}.
\]

Then, the stability conditions of the proposed controller can be obtained as follows:

**Result 1.** Assuming that the process I-O model follows Equation (1) with \( b = 1 \), the controller stated in Equation (4) is asymptotically stable if

\[
d \leq 1 \quad \text{and} \quad \rho(Q) < 1.
\]

Appendix C provides the proof of Result 1. Furthermore, Appendix D shows that the eigenvalues of \( Q \) can be obtained by

\[
\det (Q - \lambda I) = (-1)^{r+s+1} \lambda^{(r+s+3)} \left( 1 - \sum_{i=1}^{r+s+3} \pi_i \lambda^{-i} \right) = 0,
\]

where

\[
1 - \sum_{i=1}^{r+s+3} \pi_i \lambda^{-i} = \left[ \left( 1 - \sum_{i=1}^{r+2} c_i \lambda^{-i} \right) \left( 1 - \sum_{j=1}^{s+2} a_j \lambda^{-j} \right) \right] - \left( \sum_{i=1}^{s+1} c_i \lambda^{-i} \right) \left( \sum_{j=1}^{r+2} a_j \lambda^{-j} \right).
\]

Note that when \( r = s = 1 \), and \( \phi_1 = \hat{\phi}_1 = 0 \), \( Q \) can be reduced as follows:

\[
Q = \begin{pmatrix}
1 + \delta_1 - \frac{\gamma_0}{\gamma_0} (\omega + \delta_1) & -\delta_1 + \frac{\gamma_0}{\gamma_0} \delta_1 & 0 & -r_1 + \frac{\gamma_0}{\gamma_0} \hat{\gamma}_1 & 0 \\
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
-\frac{1}{\gamma_0} (\omega + \delta_1) & \frac{\gamma_1}{\gamma_0} & 0 & \frac{\gamma_1}{\gamma_0} & 0 \\
0 & 0 & 0 & 1 & 0
\end{pmatrix}.
\]

Then, the stability conditions of the qMMSE controller proposed by Tseng and Mi (2014) can be directly obtained from Equation (21). By setting \( \delta_1 = \hat{\gamma}_1 = 0 \) and \( \delta_1 = \gamma_1 = 0 \), \( Q \) can be further reduced to:

\[
Q = \begin{pmatrix}
1 - \frac{\gamma_0}{\gamma_0} \omega & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
-\frac{1}{\gamma_0} \omega & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0
\end{pmatrix}.
\]

Again, from Equation (21), we can obtain the stability conditions of the sEWMAs controller directly. In Section 5, we will use an example to demonstrate how to apply Theorem 1 to find the feasible region of generalized qMMSE controllers. Furthermore, if the stability conditions are held, we have the following result:

**Result 2.** When \( t \) is large,

\[
W_t \sim \text{ARMA}(r + s + p + 3, s + q + 2 - d).
\]

Appendix E provides the proof of this result. This result demonstrates that the asymptotic distribution of the process output (adjusted by the proposed controller) will follow a stationary ARMA process with order \((r + s + p + 3, s + q + 2 - d)\) when the stability conditions are held.

In relevant literature, the TMSE of the process outputs is usually used to evaluate the control performance of controllers, and is defined as follows:

\[
\text{TMSE} = \sum_{t=1}^{N} E(Y_t - \tau)^2.
\]
where the production run size $N$ is assumed to be moderately large. Based on Lemma 1, Equation (22) can be expressed as follows:

$$\text{TMSE} = \sum_{t=1}^{N} \Gamma_{t-1}^2 + \sum_{t=1}^{N} \text{var}(W_t).$$

(23)

The first term of Equation (23) comes from the nonrandom part, which is irrelevant to the process disturbances. Then, Equation (23) can be approximately expressed as follows:

$$\text{TMSE} \approx \sum_{t=1}^{N} \Gamma_{t-1}^2 + N \times \nu,$$

(24)

where $\nu = \text{var}(W_t)$ can be easily obtained by the Yule–Walker Equations (Box, Jenkins, and Reinsel 1994).

Therefore, the optimal discount factor, $\omega^*$, as well as its corresponding TMSE (denoted by $\text{TMSE}^*$), can be obtained directly by

$$\text{TMSE}^* = \min_{\omega} \left( \sum_{t=1}^{N} \Gamma_{t-1}^2 + N \times \nu \right),$$

(25)

and

$$\omega^* = \arg\min_{\omega} \text{TMSE}.$$  

(26)

5. Some Illustrative Examples

In the following, we use an example to illustrate how to obtain the stability conditions and the optimal discount factor of the proposed controller.

5.1 An Illustrative Example and the Off-Line Prediction Model

In this example, we use an example to illustrate how to obtain $Y_t$ and $x_{t-1}$ denote the removal rate (unit: Angstrom/minute) and the platen speed (original scale in revolutions per minute (rpm) and are normalized to the range of $(-3, 3)$) of CMP process for run $t$, respectively. Assume that $r = 1800$. By considering the second-order carryover effects on both input and output variables, and slightly modifying the model proposed by Fan et al. (2002) and Tseng and Mi (2014), this study adopts the following dynamic model as the plant I-O model:

$$Y_t = 0.53 Y_{t-1} + 0.3 Y_{t-2} + 633 + 96.3 x_{t-1} + 32.1 x_{t-2} + 23.5 x_{t-3} + \eta_t,$$

(27)

where $\eta_t$ is an ARIMA(1,1,1) series with parameters, $\phi_1 = -0.1$, $\theta_1 = 0.44$, and $\sigma_\epsilon = 40$. That is, $(1 + 0.1B)(1 - B)\eta_t = (1 - 0.44B)\epsilon_t$, where

Chemical Mechanical Polishing (CMP) is an important processing step in the manufacture of semiconductors. Let $Y_t$ and $x_{t-1}$ denote the removal rate (unit: Angstrom/minute) and the platen speed (original scale in revolutions per minute (rpm) and are normalized to the range of $(-3, 3)$) of CMP process for run $t$, respectively. Assume that $r = 1800$. By considering the second-order carryover effects on both input and output variables, and slightly modifying the model proposed by Fan et al. (2002) and Tseng and Mi (2014), this study adopts the following dynamic model as the plant I-O model:

$$Y_t = 0.53 Y_{t-1} + 0.3 Y_{t-2} + 633 + 96.3 x_{t-1} + 32.1 x_{t-2} + 23.5 x_{t-3} + \eta_t,$$

(27)

where $\eta_t$ is an ARIMA(1,1,1) series with parameters, $\phi_1 = -0.1$, $\theta_1 = 0.44$, and $\sigma_\epsilon = 40$. That is, $(1 + 0.1B)(1 - B)\eta_t = (1 - 0.44B)\epsilon_t$, where

For example, when $\omega = 0.05$, then we have $\rho(Q) = 0.948 < 1$. From Theorem 1, the process output adjusted by the generalized qMMSE controller with discount factor $\omega = 0.05$ will satisfy the stability conditions. Furthermore, let $F = \{\omega|\rho(Q) < 1\}$ denote the feasible region for the discount factor of the generalized qMMSE controller. Via numerical computation, it can be shown that $F = (0, 1)$. The result demonstrates that the proposed controller will achieve the stability conditions regardless of how large the discount factor ($0 < \omega < 1$) that is adopted. Furthermore, the asymptotic distribution of the controlled output will follow a stationary ARMA (8,4) process.

5.3 Optimal Discount Factor of the Controller

Set $N = 30$ and under the same settings and estimators of these parameters stated in Section 5.1, TMSE can be computed directly from Equation (24). Figure 3 shows the plot of TMSE...
under various values of $0 < \omega < 1$. The plot indicates that the optimal TMSE$^*$ = $1.97 \times 10^6$ and the corresponding optimal discount factor is $\omega^* = 0.30$. Furthermore, another discount factor with $\omega_1 = 0.05$ is also shown in Figure 3. The corresponding TMSE$(\omega_1) = 3.64 \times 10^6$. Comparing these two TMSE values, the proposed controller with the optimal discount factors has the ability to reduce TMSE by as much as 45.8%. Figure 4 also shows the sequences of process outputs controlled by $\omega^*$ and $\omega_1$. The results demonstrate that the generalized qMMSE controller with the optimal discount factor regulates the process output more efficiently than with the nonoptimal discount factor.

5.4 The Effect of Non-Gaussian Error Distribution on the Control Performance of the Proposed Controller

In Sections 5.1–5.3, we assume that $\eta_t$ is an ARIMA(1,1,1) series and $\epsilon_t$ is a Gaussian white noise with mean 0 and standard deviation $\sigma_\epsilon = 40$. In practical applications, however, $\epsilon_t$ may not follow a Gaussian white noise. Therefore, we need to investigate the effect of non-Gaussian noise on the control performance of the proposed controller comprehensively. We will mainly compare the performance of the proposed controller under the cases when $\epsilon_t = c_\epsilon \ast t(\nu)$, and $\epsilon_t$ follows a skewed normal (SN) distribution, $\text{SN}(\xi, w, \alpha)$, where $\xi$, $w$, and $\alpha$ are the location, scale, and shape parameters, respectively. Note that $c_\epsilon$, $\xi$, $w$, and $\alpha$ are chosen in such a way that $E(\epsilon_t) = 0$, and $\text{Var}(\epsilon_t) = 40^2$. Five non-Gaussian noises for $\epsilon_t$ are chosen in such a way that $\epsilon_t$ is a skewness to the left distribution. The parameters in the part of transfer function are not so serious. Note that, however, when $\epsilon_t$ has a larger skewness to the left, with $\alpha = -5$ and $\alpha = -20$, it will cause a very serious bias (even with the reverse sign) on the estimate of parameter $\phi_1$.

Now, based on the estimates in Table 1, we can check their stability conditions via Equation (21). Table 2 summarizes the values of $\rho$ ($\textbf{Q}$) when $\omega = 0.05$ and the stability regions of $\textbf{F}$ for these five non-Gaussian noises. The results demonstrate that the proposed controller still achieves the stability conditions for all discount factors in $(0, 1)$, no matter when $\epsilon_t$ is seriously skewness to the right or to the left.

For each non-Gaussian noise, Table 3 also summarizes the optimal discount factor, the optimal TMSE$^*$, and its proportion of TMSE reduction (with respect to the TMSE under $\omega = 0.05$). From Table 3, it demonstrates that optimal $\omega^*$ for the cases of $\text{SN}(\nu, 50.25, 64.23, 5)$ and $\text{SN}(\nu, 52.76, 66.21, 20)$ are larger than that of the other three cases. It means that a larger discount factor may be needed when $\epsilon_t$ is a skewness to the right distribution. Furthermore, it also demonstrates that TMSE$^*$ will be enlarged when $\epsilon_t$ is a skewness to the left distribution. The partial reasons may be because the proposed controller uses an estimate with a reverse sign for $\phi_1$ (that is far away from the true value with $\phi_1 = -0.10$), which leads the TMSE to be extremely large in comparing with other cases. We also show the plots of all the sequences of process output, which are controlled by $\omega^*$ and $\omega_1 = 0.05$, for all non-Gaussian noises and the results are given in Appendix H. Roughly speaking, the proportion of

![Figure 3](image-url) The plot of TMSE versus the discount factor $\omega$.

![Figure 4](image-url) The sequences of process outputs controlled by the generalized qMMSE controller with two different discount factors $\omega^*$ and $\omega_1$. 
In the case of high-order dynamic models. Instead, we use a probability. Generally, it is not easy to obtain an analytical result if the non-Gaussian noise is a serious skewness to the right distribution.

### 5.5 How Large Is the Sample Size Needed to Achieve Global Stability?

In the previous Section 5.2, the predicted model is developed based on the off-line data. Therefore, the sample size and the strength of the I-O relationship play major roles in determining the stability conditions of the proposed controller. For the static SISO model, Tseng and Hsu (2005) derived simple formulas for the adequate sample size to achieve stability with a guaranteed probability. Generally, it is not easy to obtain an analytical result in the case of high-order dynamic models. Instead, we use a simulation study to address this decision problem. In the following study, we assume that the parameters \( \delta_1, \delta_2, \gamma_0, \gamma_1, \gamma_2, \phi_1 \) (stated in Section 5.1) are all proportional to its original scale with a uniform random variable. That is, \( \delta_i(\text{new}) = f \times \delta_i \), for \( i = 1, 2, \phi_1(\text{new}) = f \times \phi_1 \), and \( \gamma_j(\text{new}) = f \times \gamma_j \), for \( j = 0, 1, 2 \), where \( f \sim \text{unif}(0.9, 1.1) \). Under various settings of the sample size in the offline stage \( T = 30, 40, \ldots, 120 \), we implement 1000 simulation trials from the process I-O model with these new parameter settings. Figure 5 shows the relationship of the proportions of achieving global stability (among 1000 simulation trials) with respect to sample size. For example, when \( T = 60 \), we obtain 973 trials that achieved global stability among 1000 simulation trials. Figure 5 also demonstrates a monotonic pattern whereby the higher the proportion of achieving global stability, the larger the sample size that is required. In approximate terms, we will have a 98.9% confidence to achieve global stability when \( T \geq 70 \).

### 5.6 The Effects of Model Misspecification on Control Performance

The control performance of the sEWMA controller has been comprehensively addressed with that of the qMMSE controller with order (1, 1) in Tseng and Mi (2014) and the results demonstrate that the sEWMA controller, in general, is worse than that of the qMMSE controller under a dynamic process. Therefore, we will not include the performance comparison of the sEWMA controller in the following study.

Revisiting Section 5.1, we assume that the off-line observations truly come from Equation (27) but are wrongly misspecified as its submodels. The corresponding estimators of these parameters in the submodels are:

1. \( \hat{\delta}_1, \hat{\delta}_2, \hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\phi}_1 \) (0.427, 0.408, 91.892, 
   \[ -31.846, 0.149 \);  
2. \( \hat{\delta}_1, \hat{\gamma}_0, \hat{\gamma}_1, \hat{\phi}_1 \) (0.769, 94.720, 
   \[ -6.677, -28.181, -0.679 \); and  
3. \( \hat{\delta}_1, \hat{\gamma}_0, \hat{\gamma}_1, \hat{\phi}_1 \) (0.821, 81.095, 9.120, 
   \[ -0.678 \).

Substituting these estimates into matrix \( Q \) and from Theorem 1, it can be easily shown that all these submodels are still
global stability. That is, the stability conditions of controlled outputs (under these submodels) can be achieved for all discount factors, $\omega \in [0, 1]$. Furthermore, by using these estimated parameters, the optimal discount factors for these submodels are computed via Equation (26), and the values are 0.07, 0.54, and 0.28; while the corresponding $\text{TMSE}$ for these submodels are $3.30 \times 10^6$, $2.14 \times 10^6$, and $3.52 \times 10^6$, respectively. Compared to $\text{TMSE}^* = 1.97 \times 10^6$ under $(r, s) = (2, 2)$ stated in Section 5.3, this demonstrates that the effect using submodel with $(r, s) = (1, 2)$ on TMSE is not so serious, while the submodels with $(r, s) = (2, 1)$ and $(1, 1)$ will increase the proportions of TMSE up to 40.3% and 44.0%, respectively. These results demonstrate that the model misspecification effects on the TMSE are not negligible when the process I-O model was wrongly misidentified as $(r, s) = (2, 2)$ and $(1, 1)$. Figure 6 also shows the outputs controlled by four generalized qMMSE controllers (with their individual optimal discount factors). From the plot, all of the process outputs converge to the desired target, but the convergent rates to the desired target under $(r, s) = (2, 2)$ and $(1, 2)$ are faster than those in the other two cases.

6. A Comprehensive Study on Performance Comparisons

In the above CMP example, we compare the control performances of the proposed controller with respect to some competing controllers based on 100 observations in the off-line stage. In practice, the control performance strongly depends on the estimated accuracy of the process I-O prediction model, which clearly depends not only on the observation in the offline stage, but also on the dynamic parameters for both input and output variables. Therefore, we will further discuss these issues via comprehensive simulation studies.

6.1 The Effect of Off-Line Observations on the Control Performance

Case 1: Process I-O plant model is

\[ Y_t = 0.5Y_{t-1} + 0.3Y_{t-2} + 4 + x_{t-1} + 0.45x_{t-2} + 0.33x_{t-3} + \eta_t, \quad (28) \]

where $\eta_t$ is an IMA(1,1) series with parameter $\theta = 0.25$ and $\sigma_\eta = 1$. For simplicity, we assume that the target for the output response is $\tau = 0$, and $N = 50$. In the offline stage, suppose that $T$ observations are taken from this I-O model, and the total number of simulation trials is $m$. Let $\text{TMSE}^*_r(j, T)$ denote the minimum TMSE of the generalized qMMSE controller with order $(r, s)$ under $j$th simulation trial, where $j = 1, \ldots, m$. Now, we define the inflation factor (IF$_1$) as follows:

\[
\text{IF}_1(T) = \frac{1}{m} \sum_{j=1}^{m} \left( \frac{\text{TMSE}^*_r(1, T)}{\text{TMSE}^*_r(2, T)} \right). \quad (29)
\]

This index can be used to measure the inflation (or penalty) of TMSE when the true order $(2, 2)$ of process I-O dynamical model, although wrongly treated as order $(1, 1)$. Set $m = 100$. The plots of IF$_1$ under various combinations of $T = 50, 60, \ldots, 120$ are shown in Figure 7.

The results demonstrate that the value of IF$_1$ is around 1.590 ($\pm 0.225$). This means that the effects of model misidentification on IF$_1$ are not negligible.

Case 2: Process I-O plant model is

\[ Y_t = 0.5Y_{t-1} + 4 + x_{t-1} + 0.45x_{t-2} + \eta_t, \quad (30) \]

where $\eta_t$ is an IMA(1,1) series with parameter $\theta = 0.25$ and $\sigma_\eta = 1$. In addition, we assume that the target for the output response is $\tau = 0$, and $N = 50$. Similar to Equation (28), we define another inflation factor (IF$_2$) as follows:

\[
\text{IF}_2(T) = \frac{1}{m} \sum_{j=1}^{m} \left( \frac{\text{TMSE}^*_r(2, T)}{\text{TMSE}^*_r(1, T)} \right). \quad (31)
\]

The plots of IF$_2$ under various combinations of $T = 50, 60, \ldots, 120$ are shown in Figure 8, and the results demonstrate that the value of IF$_2$ is around 1.285 ($\pm 0.043$). This means that the effects of model misspecification on IF$_2$ are also not negligible but slightly smaller than those of IF$_1$.
6.2 Second-Order Carryover Effects on the Control Performance

In this subsection, the following plant model is considered:

\[ Y_t = 0.5Y_{t-1} + \delta_2 Y_{t-2} + 4 + x_{t-1} + 0.45x_{t-1} + \gamma_2 x_{t-2} + \eta_t, \]

where

\[ \delta_2 = c_1 \times 0.5 \quad \text{and} \quad \gamma_2 = c_2 \times 0.45. \]

The constants \(c_1\) and \(c_2\) can be appropriately used to describe the second-order carryover effects on both the output and input recipe, respectively. Furthermore, we assume that \(\eta_t\) follows a general ARIMA \((1, 1, 1)\) model with parameters \((\phi, \theta)\) and \(\sigma_\eta = 1\). The total run size is \(N = 50\), and the desired target for the output response is \(\tau = 0\). Prefix a pair of \((\phi, \theta)\), where \(-0.4 \leq \phi \leq 0.4\) and \(-0.25 \leq \theta \leq 0.25\), and Figure 9 shows the contours of \(IF_1\) (where 50 simulation trials are conducted) under various combinations of \((c_1, c_2)\), where \(0 \leq c_1 \leq 0.8\), and \(0 \leq c_2 \leq 0.8\).

Typically, under the cell of \((\phi, \theta) = (0, 0.25)\), and \((c_1, c_2) = (0.6, 0.73)\), we found that \(IF_1 \approx 1.52\). This result is very close to \(IF_1 \approx 1.59\), which was obtained from Figure 7 under \(T = 100\). Furthermore, with a fixed pair of \((\phi, \theta)\), the value of \(IF_1\) becomes larger and larger when \(c_1\) and \(c_2\) are increasing. More specifically, when \(c_1 \geq 0.5\) and \(c_2 \geq 0.5\), the average value of \(IF_1\) is approximately 1.55. This means that the penalty of using the wrong control scheme with \((r, s) = (1, 1)\) will be serious when the second-order carry-over effects are significant.

7. Conclusions

Most of the studies on run-to-run feedback control assume that the process I-O relationship follows a static or first-order dynamic model. Generally speaking, the process relationship can be a more complex dynamic model, rather than these simplified models. Under an input–output dynamical model, this study first proposed a generalized qMMSE controller for adjusting the process output. Then, the expression of process output,
the long-term stability conditions and the optimal discount factor of this controller are also derived analytically. Furthermore, under \((r, s) = (2, 2)\), we also address the effects of model identification in the offline stage and second-order carryover effects on the TMSE performance. Based on our simulation study stated in Section 6.2, the penalty of TMSE will inflate approximately 1.55 times larger when the true carryover parameters in Section 6.2, the penalty of TMSE will inflate approximately on the TMSE performance. Based on our simulation study stated herees for many insightful suggestionsthat significantly improve the quality of this article. The authors also thank Dr. I-C Lee and Miss Z-H Wu for her useful programming and suggestions. The work was partially supported by the Ministry of Science and Technology (Grant No: NSC-100-2221-E-007-060-MY3) of Taiwan, Republic of China (ROC).

**References**

Box, G. E. P, Jenkins, G. M., and Reinsel, G. C. (1994), *Time Series Analysis, Forecasting and Control* (3rd ed.), Englewood Cliffs, NJ: Prentice Hall. [382,385]

Butler, S. W., and Stefani, J. A. (1994), "Supervisory Run-to-Run Control of Polysilicon Gate Etch using in Situ Ellipsometry," *IEEE Transactions on Semiconductor Manufacturing*, 7, 193–201. [381]

Capilla, C., Ferrer, A., Romero, R., and Hualda, A. (1999), "Integration of Statistical and Engineering Process Control in a Continuous Polymerization Process," *Technometrics*, 41, 14–28. [382]

Chen, A., and Guo, R. S. (2001), "Age-Based Double EWMA Controller and its Application to CMP Processes," *IEEE Transactions on Semiconductor Manufacturing*, 14, 11–19. [381]

Del Castillo, E. (2001), "Some Properties of EWMA Feedback Quality Adjustment Schemes for Drifting Disturbances," *Journal of Quality Technology*, 33, 153–166. [381]

—— (2002), *Statistical Process Adjustment for Quality Control*, New York: Wiley-Interscience. [382]

Fan, S. K. S., Jiang, B. C., Jen, C. H., and Wang, C. C. (2002), "SISO Run-to-Run Feedback Controller using Triple EWMA Smoothing for Semiconductor Manufacturing Processes," *International Journal of Production Research*, 40, 3093–3120. [382,385]

Hamby, E. S., Kabamba, P. T., and Khargonekar, P. P. (1998), "A Probabilistic Approach to Run-to-Run Control," *IEEE Transactions on Semiconductor Manufacturing*, 11, 654–669. [381]

Ingolfsson, A., and Sachs, E. (1993), "Stability and Sensitivity of an EWMA Controller," *Journal of Quality Technology*, 25, 271–287. [381]

Jen, C. H., and Jiang, B. C. (2008), "Combining On-line Experiment and Process Control Methods for Changes in a Dynamic Model," *International Journal of Production Research*, 46, 3665–3682. [382]

Lin, C. H., Tseng, S. T., and Wang, S. F. (2013), "Modified EWMA Controller Subject to Metrology Delay," *IIE Transactions*, 45, 409–421. [390]

Meeker, W. Q., and Escobar, L. A. (1998), *Statistical Methods for Reliability Data*, New York: Wiley. [390]

Moyne, J., Solakhian, V., Yershov, A., Anderson, M., and Mockler-Hebert, D. (2002), “Development and Deployment of a Multi-Component Advanced Process Control System for an Epitaxy Tool,” in *Advanced Semiconductor Manufacturing 2002 IEEE/SEMI Conference and Workshop*, pp. 123–130. [381]

Pan, R., and Del Castillo, E. (2001), "Identification and Fine Tuning of Closed-Loop Processes Under Discrete EWMA and PI Adjustments," *Quality and Reliability Engineering International*, 17, 419–427. [382]

Qiu, P., and Xiang, D. (2014), "Univariate Dynamic Screening System: An Approach for Identifying Individuals with Irregular Longitudinal Behavior," *Technometrics*, 56, 248–260. [382]

R Development Core Team (2013), R: A Language and Environment for Statistical Computing, Vienna, Austria: R Foundation for Statistical Computing. Available at [http://www.R-project.org]. [385]

Sachs, E., Hu, A., and Ingolfsson, A. (1995), "Run by Run Process Control: Combining SPC and Feedback Control," *IEEE Transactions on Semiconductor Manufacturing*, 8, 26–43. [381]

Tseng, S. T., Chou, R. J., and Lee, S. P. (2002), "Statistical Design of Double EWMA Controller," *Applied Stochastic Models in Business and Industry*, 18, 313–322. [381]

Tseng, S. T., and Hsu, N. J. (2005), "Sample-Size Determination for Achieving Asymptotic Stability of a Double EWMA Control Scheme," *IEEE Transactions on Semiconductor Manufacturing*, 18, 104–111. [387]

Tseng, S. T., and Lin, C. H. (2009), "Stability Analysis of Single EWMA Controller Under Dynamic Models," *IEEE Transactions*, 41, 654–663. [382]

Tseng, S. T., and Mi, H. C. (2014), "Quasi-Minimum Mean Square Error Run-to-Run Controller for Dynamic Models," *IEEE Transactions*, 46, 185–196. [382,383,384,385,387]

Tseng, S. T., Mi, H. C., and Lee, I. C. (2016), "A Multivariate EWMA Controller for Linear Dynamic Processes," *Technometrics*, 58, 104–115. [390]

**Supplementary Materials**

All the proofs are given in Appendices (Appendices.pdf).

**Acknowledgments**

The authors are grateful to the editor, associate editor, and anonymous referees for many insightful suggestions that significantly improve the quality of this article. The authors also thank Dr. I-C Lee and Miss Z-H Wu for her