Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

EWMA Controller Tuning and Performance Evaluation in a High Mixed System

Ming-Da Ma[a], Chun-Cheng Chang[b], David Shan-Hill Wong[b]
Shi-Shang Jang[b*]

[a]Center for Control and Guidance Technology,Harbin Institute of Technology, Harbin, China
[b]Department of Chemical Engineering, National Tsing-Hua University, Hsin-Chu, Taiwan

Abstract: The exponentially weighted moving average (EWMA) controller is a very popular run-to-run (RtR) control scheme in semiconductor industry. However, in any typical step of semiconductor process, many different products are produced on parallel tools. RtR control is usually implemented with a "threaded" control framework, i.e.: different controllers are used for different combinations of tools and products. In this paper, the problem of EWMA controller tuning and performance evaluation in a mixed product system is investigated by simulation and time series analysis. It was found that as the product frequency changed, the tuning guidelines of a threaded EWMA controller were different for different types of tool disturbances. For a stationary ARMA(1,1) noise, the tuning parameter should be increased as product frequency decreases. If the tool exhibits non-stationary tool dynamics, e.g. ARIMA(1,1,1) noise, the tuning parameter should increase as the product frequency decreases. Copyright © 2008 IFAC

Keywords: mixed product system, EWMA, performance evaluation, time series

1. INTRODUCTION

In the last two decades, run-to-run (RtR) control which combines statistical process control (SPC) and feedback control has been widely used in semiconductor manufacturing industry (Moyne et al., 2001; de Castillo, 2002). RtR control adjusts process recipes or inputs from run to run to compensate for various process disturbances to maintain the process output close to a given target. RtR control can efficiently improve the product yield, throughout and reduce scrap, rework and cycle time.

In semiconductor manufacturing industry, production resembles an automated assembly line in which many similar products with slightly different specifications are manufactured step-by-step on a number of different tools. This constitutes a high-mix production system. Implementation of RtR is commonly implemented with a "threaded" framework. In this approach, each specific combination of tool and product is called a thread. Each thread has its own controller. Threaded exponentially weighed moving average (EWMA) control is probably the most popular control architecture although other frameworks were proposed (Pasadyn and Edgar, 2005; Firth et al., 2006; Ma et al., 2007). One key advantage of this approach is that stability is guaranteed (Zheng et al, 2006). On the other hand, some commonly encountered questions are: whether the controller performance is optimal, and how should each thread be tuned differently as the production frequency changes. Controller performance evaluation has seen active development in research and gradual acceptance in industry since the seminal work of Harris (1989). Recent development in this subject was reviewed by Qin (1998). Prabhu et al. (2006) has proposed a performance evaluation for a EWMA controller of a single thread. In this paper, we shall use the performance evaluation technique and time series analysis to investigate the optimal of a threaded exponentially weighted moving average (EWMA) controller in a high mixed system.

2. THREADED EWMA CONTROL OF A MIXED PRODUCT PLANT

Consider a static simple linear single input-single-output process performed in a mixed run situation on a single tool

$$y[t]=\alpha_{k[t]}+\beta u[t]+N[t] \qquad (1)$$

where y[t] and u[t] denote values of output and manipulated variable used on the t^{th} run on the tool. $\alpha_{k[t]}$ is offset or bias term associated, β is process gain, N[t] is a noise process associated with the tool. In a threaded approach, a sequence of output and input for each specific product is resampled:

$$y[t_k] = \alpha_k + \beta u[t_k] + N_k[t_k] \qquad (2)$$

where t_k is an index of the number of runs making the k^{th} product that have been carried out. Given a process model $y=bu+a_k$ for each product, the offset term can be estimated by EWMA filter

$$a_k[t_k] = \lambda(y[t_k] - bu[t_k]) + (1-\lambda)a_k[t_k - 1] \qquad (3)$$

The control action is

$$u[t_k + 1] = \frac{T_k - a_k[t_k]}{b} \qquad (4)$$

Note that the threaded EWMA control is similar to a single product EWMA control, except the disturbance experienced is not the actual change in tool condition N[t] from run to run, but a re-sampled series $N_k[t_k]$.

3. PERFORMANCE EVALUATION USING SIMULATIONS

Performance assessment is widely implemented in process control now. Usually, minimum variance (MV) is adopted as the benchmark. However, MV may not be easily achieved. Therefore, best achievable performance (BAP) under the current controller is often used. Prabhu et al. (2006) proposed a performance assessment method for EWMA control to obtain BAP estimate and optimal λ using closed loop response data. We shall use this method to evaluate the performance of the above mixed product system under threaded EWMA control. A simulation example consisting of one tool and three products is used. Each product is assigned a unique bias. Three disturbances used are ARMA(1,1) (first order autoregressive-moving average), IMA(1,1) (first order integrated moving average) and ARIMA(1,1,1) (first order autoregressive integrated moving average) processes respectively.

$$\text{Noise I}: \ (1-\phi B)N[t] = (1-\theta B)\varepsilon[t], \phi = 0.8, \theta = 0.4$$

$$\text{Noise II}: \ (1-B)N[t] = (1-\theta B)\varepsilon[t], \theta = 0.5$$

$$\text{Noise III}: \ (1-\phi B)(1-B)N[t] = (1-\theta B)\varepsilon[t]$$
$$\phi = 0.2, \theta = 0.6$$

Here B stands for backward shift operator. In the simulation, the percentage of product 1 changes from 1% to 50%, Product 2 and 3 makes up the rest of the products. For each type of disturbance, 400 sets of disturbance sequences, each with 400 runs are generated. The production schedule is random. A value of $\lambda = 0.2$ is used to obtain the closed loop data. The changes in estimated optimal λ_{opt} are shown in Figures 1. It is interesting to find out that the tuning rules for the above disturbances are different. For noise I, which is a stationary ARMA(1,1) noise, the optimal tuning parameter λ decreases as the percentage of the product decreases. However, for the other two nonstationary noises II and III, the optimal tuning parameter λ increases as the percentage of the product decreases. It is interesting therefore to see if differences in trends are general.

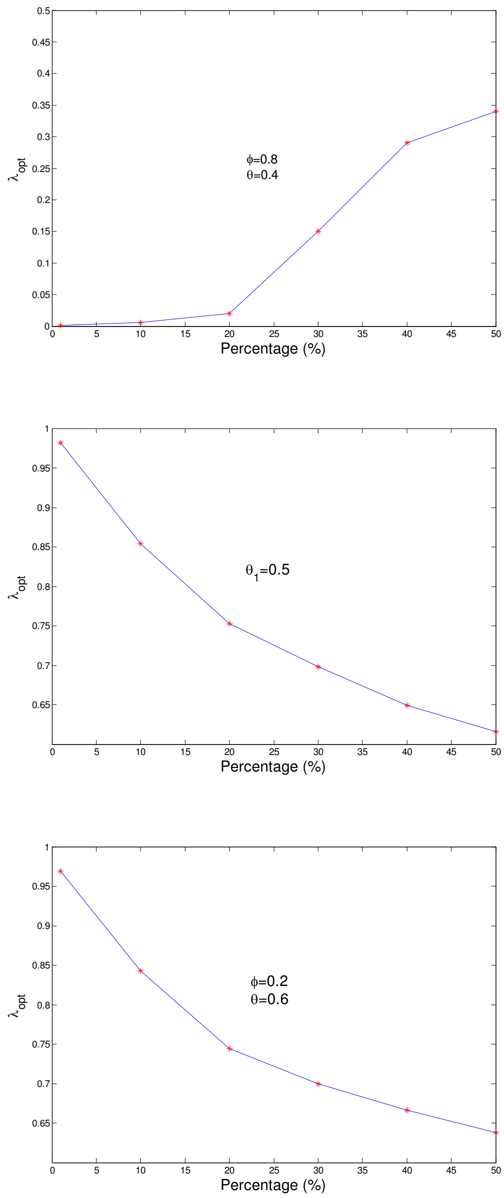

Fig. 1. Effect of production frequency on optimal tuning parameter λ

4. TIME SERIES ANALYSIS

For a simple gain process controlled by a EWMA controller with a single product, the output variance for an ARMA(1,1) disturbance is given by (see Appendix)

$$\sigma_y^2 = \left(\frac{2}{1+\kappa} - 2\rho_1 \frac{1-\kappa}{(1-\phi\kappa)(1+\kappa)} \right) \sigma_\varepsilon^2 \qquad (5)$$

where $\kappa = 1 - \lambda\xi$, $\xi = \beta/b$, and ρ_1 is the autocovariance of the disturbance N[t]. The optimal λ, given ϕ and ρ_1 can be obtained by differentiating the above equations with respect to κ. Figure 2 is contour plot of optimal λ at different values of ϕ and ρ_1.

If the product is manufactured at regular intervals of every h runs, h being a positive integer, then its threaded EWMA controller will face a disturbance which is a resampled sequence of the original sequence at at every h^{th} time points. If an ARMA(1,1) process with parameters (ϕ , θ) is resampled, the resulting process, $M_h[t]$, is also an ARMA(1,1) process (MacGregor, 1976)

$$(1-\phi_h B) M_h [t] = (1-\theta_h B)\varepsilon[t] \qquad (6)$$

with

$$\phi_h = \phi^h \qquad \rho_1(h) = \phi^{h-1}\rho_1 \qquad (7)$$

where $\rho_1(h)$ is the autocorrelation of $M_h[t]$. Since both ϕ_h and $\rho_1(h)$ decreases exponentially with h, from Figure 2, we can infer that optimal λ must decrease with production frequency when the tool disturbance is an ARMA(1,1) noise. When the process is stationary and the samples are taken less frequently in time, the autocorrelation of the sampled data will decrease. When the sampling interval is sufficiently large, the data will appear to be uncorrelated. It is well known that when the disturbance is white noise, no control action should be added, i.e., $\lambda = 0$. Therefore, we can draw the conclusion that for stationary disturbances, the optimal EWMA controller gain decreases as the decreasing of the percentage of the products.

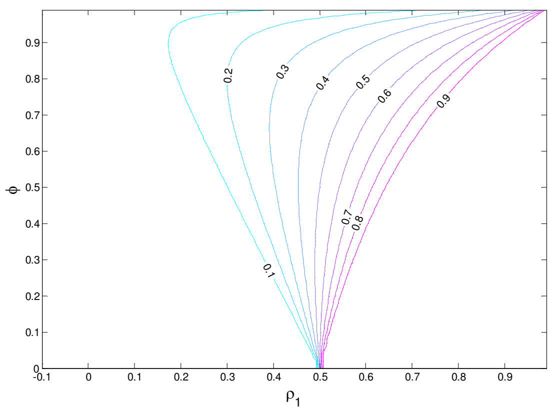

Fig. 2. Contour plots for optimal ARMA(1,1) disturbance processes.

If an IMA(1,1) process N[t] is sampled at every h time points, the resulting process is denoted by M[t], is also an IMA(1,1) process with

$$(1-B) M_h [t] = (1-\theta_h B)\eta_h [t] \qquad (8)$$

With

$$\frac{h(1-\theta_1)^2}{\theta_1} = \frac{(1-\theta_h)^2}{\theta_h} \qquad (9)$$

and $\eta_h[t]$ is a white noise process with σ_h (Box, et al., 1994). Figure 3 shows the changes of θ_h with the sampling interval h. It can be seen that θ_h decreases as h increases. Box (et al., 1994) showed that for an IMA(1,1) disturbance and the optimal EWMA controller gain is $\lambda = 1\text{-}\theta$. Therefore, the optimal EWMA controller tuning parameter increases as the production frequency decreases.

Given any arbitrary covariance or correlation sequence with only a finite number of nonzero elements there is a finite moving average (MA) process corresponding to the sequence (Andersen, 1971). Hence an ARIMA(p,1,q) series can be approximated by an IMA(1,q') series.

$$(1-B) N[t] = \left(1-\theta_1 B - \theta_2 B^2 - \cdots - \theta_{q'} B^{q'}\right)\varepsilon[t] \qquad (10)$$

If this series is sampled at intervals of h units and $h > q'\text{-}1$, then the resulting process can be represented by an IMA(1,1) series (McGregor, 1976):

$$(1-B) M_h [t] = (1-\theta_h B)\varepsilon_h (t) \qquad (11)$$

Hence, our conclusion on IMA(1,1) processes can be extended to other non-stationary disturbances.

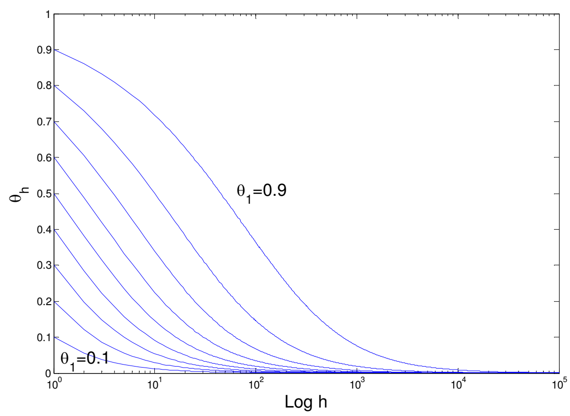

Fig. 3. Changes in parameter θ_h with h when resampling an IMA(1,1) process with parameter θ_1

5. CONCLUSION

In this paper, the effect of production frequency on optimal tuning of threaded EWMA controller in a high mixed production system was studied. Simulation shows that for stationary disturbances, the optimal EWMA controller gain decreases as the production frequency decreases. For non-stationary disturbances, the EWMA controller gain increases as the production frequency decreases. The conclusions are supported by re-sampled time series analysis at fixed intervals. Different semiconductor manufacturing processes will have different tool dynamics. Processes with rapid tool wear, such as chemical-mechanical-polishing is likely to have a tool disturbance non-stationary process. Processes such as photolithography and overlay may be more stationary. According to our study, the tuning guidelines for products with small lot count will be different for different processes.

REFERENCES

Sachs, E., A. Hu and A. Ingolfsson (1995). Run by run process control: combining SPC and feedback control. IEEE Trans. Semiconduct. Manufact., 8, 26-43.

Patel, N.S. and S.T. Jenkins (2000). Adaptive optimization of run-to-run controllers: the EWMA example. IEEE Trans. Semicond. Manuf., 13, 97-107.

Tseng, S.T., A.B. Yeh, F. Tsung and Y.Y. Chan (2003). A study of variable EWMA controller. IEEE Trans. Semicond. Manuf., 16, 633-643.

Smith, T.H. and D.S. Boning (1997). Artificial neural network exponentially weighted moving average controller for semiconductor processes. J. Vacuum Science Tech., 15, 236-239.

Ying Z., Q.H. Lin, D.S.H. Wang, S.S. Jang and K. Hui (2006). Stability and Performance Analysis of Mixed Product Run-To-Run Control, J. of Process Control, 16, 431-443.

Moyne, J. (2001). Run-to-Run Control in Semiconductor Manufacturing. CRC Press, Florida.

Castillo, E.D. (2002). Statistical Process Adjustment for Quality Control. John-Wiley and Sons, New York.

Firth, S.K., W.J. Campbell, A. Toprac and T. F. Edgar (2006). Just-In-Time Adaptive Disturbance Estimation for Run-to-Run Control of Semiconductor Processes. IEEE Trans. Semicond. Manuf., 19, pp. 298-315.

Pasadyn, A.J. and T. F. Edgar (2005). Observability and State Estimation for Multiple Product Control in Semiconductor Manufacturing. IEEE Trans. Semicond. Manuf., 18, pp. 592-604.

Ma, M.D., C.C. Chang, D.S.H. Wong and S.S. Jang (2007). State estimation of a mixed run plant. Submitted to J. of Process Control.

Prabhu, A.V., T.F. Edgar and R. Chong (2006). Performance assessment of run-to-run EWMA controllers. ADCHEM.

Box, G.E.P., G.M. Jenkins and G. Reinsel (1994). Time series analysis forecasting and control (3rd Ed.). Prentice-Hall.

MacGregor, J.F. (1976). Optimal choice of the sampling interval for discrete process control. Technometrics, 18, pp. 151-160.

Anderson, T.W. (1971). The statistical analysis of time series. John Wiley, New York.

Harris, T.J. (1989). Assessment of control loop performance. Can. J. Chem. Eng., 67, 856-861.

Qin, S.J. (1998). Control performance monitoring – A review and assessment. Computers and Chem. Eng., 23, pp.173-186.

APPENDIX

For the control system described in equations (2) to (4), the EWMA controller is equivalent to a discrete integral controller with gain $k_i = \lambda/b$ from the viewpoint of control engineering. If the process target is set equal to zero, the output y can be expressed as:

$$y[t] = \frac{1-B}{1-\kappa B} N[t] \tag{A1}$$

where $\kappa = 1 - \lambda\beta/b$. Taking long division, we get:

$$y[t] = \sum_{j=0}^{\infty} \psi_j N_{t-j} \tag{A2}$$

with

$$\psi_0 = 1, \quad \psi_j = \kappa^{j-1}(\kappa - 1), \quad j \geq 1$$

Hence the variance of the output is

$$\sigma_y^2 = \left(\sum_{j=0}^{\infty} \sum_{j'=0}^{\infty} \psi_j \psi_{j'} \rho_{|j-j'|} \right) \sigma_\varepsilon^2 \tag{A3}$$

For ARMA(1,1) process, we have

$$\rho_j = \phi^{j-1} \rho_1$$

(A4)

By substituting (A4) into (A3) we have

$$\sigma_y^2 = \left(\frac{2}{1+\kappa} - 2\rho_1 \frac{1-\kappa}{(1-\phi\kappa)(1+\kappa)} \right) \sigma_\varepsilon^2 \tag{A5}$$