PERFORMANCE ASSESSMENT OF RUN-TO-RUN EWMA CONTROLLERS

Amogh V. Prabhu*, Thomas F. Edgar* and Robert Chong†

**Dept of Chemical Engineering, The University of Texas at Austin, TX 78712*
†Advanced Micro Devices, Austin, TX 78751

Abstract: An iterative method is developed to determine a performance criterion for best achievable performance for discrete integral controllers. Using the performance criterion, optimal performance of the controller in place is also indicated. An analytical expression is derived so that a realistic assessment of the given integral controller is obtained. Using the theoretical equivalence of discrete integral and exponentially weighted moving average (EWMA) controllers, the method is then extended to performance assessment of EWMA controllers. A semiconductor manufacturing example is used to illustrate the utility of the method.

Keywords: performance assessment, feedback control, delay, integral control, single loop

1. INTRODUCTION

For any feedback control system in a manufacturing process, variation from the desired output can occur due to two reasons: Either the process state has changed or the controller performance has degraded. A change in process state occurs whenever any of the major process parameters change by an amount which cannot be corrected without a change in the controller tuning. But if the controller performance is degraded without any change in the state, then the controller itself must be analyzed to verify that it is behaving optimally under the given conditions.

1.1 Minimum variance control (MVC)

The first effort towards developing a performance index for feedback control systems was made by Harris (1989). This work proposed that minimum variance control represents the best achievable performance by a feedback system. All other kinds of control behave sub-optimally as compared to it. The method is applicable only to SISO systems and involves fitting a univariate time series to process data collected under routine control. This is compared to the performance of a minimum variance controller. However it has certain drawbacks:

– If controller performance is close to that of minimum variance, it indicates that it is behaving optimally. But if the deviation from minimum variance performance is large, it does not imply that the controller is sub-optimal. Under the given setup, it may the best that the controller can do. Therefore, a different benchmark may be required in such a case.
– The minimum variance index does a good job of indicating loops that have oscillation problems. Unfortunately it considers loops that are sluggish to be fine. This particularly happens when the controller has been detuned to a large extent, making controlling the loop slow.
– Minimum variance index is only a theoretical lower bound on the best possible performance. If applied in a real system, it can lead to large variations in input signals, and the closed loop often has poor robustness properties. Therefore it is not recommended to be applied to a system, but just serve as a benchmark.

1.2 Alternative methods

While the minimum variance control concept proposed by Harris (1989) was initially developed for feedback and feedforward-feedback controlled univariate systems (Desborough and Harris (1992, 1993)), the idea was further extended to multivariate systems. Stanfelj *et al.* (1993) have diagnosed the performance of single loop feedforward-feedback systems based on the MVC criteria. Eriksson and Isaksson (1994) have analyzed the MVC index and pointed out several drawbacks in the index similarly to those listed earlier. Huang *et al.* (1995) have introduced a useful method for monitoring of MIMO processes with feedback control, known as Filtering and Correlation (FCOR) analysis. This concept is further developed by Huang *et al.* (1997) to estimate a suitable explicit expression for the feedback controller invariant term of the closed-loop MIMO process from routine operating data. Harris *et al.* (1996a) have extended the MVC index to multivariable feedback processes in a manner similar to Huang *et al.* (1995) but without the filtering approach. Ko and Edgar (1998) have proposed a method to determine achievable PI control performance when the process is being perturbed by stochastic load disturbances. This is further extended to multivariable feedback control by Ko and Edgar (2000) using a finite horizon MV benchmark with specified horizon length. Salsbury (2005) has formulated statistical change detection procedures which can be used for processes subject to random load changes. The method is applicable to SISO feedback systems and uses a normalized index, which is similar to the damping ratio in a second order process.

Apart from these articles, Qin (1998) and Harris *et al.* (1999) have reviewed most methods up to 1998.

1.3 Performance Monitoring in Semiconductor Manufacturing

Most of the major processes involved in semiconductor manufacturing are done in a batch manner (Edgar *et al.*, 2000), so that any process change involves changes in the batch recipe. Run-to-run control is the most popular form of control wherein the controller parameters can be tuned after each lot, based on the data from the previous lot. Statistical process control is widely used, with most processes adopting an Exponential Weighted Moving Average (EWMA) algorithm. None of the above listed methods were developed for control systems used in the semiconductor industry. But a best achievable PID control performance bound was proposed by Ko and Edgar (2004). This was an iterative algorithm which optimized the controller parameters. Using the theoretical equivalence of EWMA controllers with discrete integral controllers, this iterative algorithm was adapted to run-to-run EWMA controllers, commonly used in semiconductor manufacturing.

In this article, we derive an iterative solution method for the calculation of achievable performance bound of a run-to-run EWMA controller, where the iterative solution uses the process input-output data and the process model. This iterative solution is based on an analytic solution for closed-loop output. A normalized performance index is then defined based on the best achievable performance. An example of a process controlled by such a controller is employed to illustrate the effectiveness of the proposed method.

2. THEORY DEVELOPMENT

The following theory explains in a step-wise manner how the performance monitoring method for a discrete integral controller (based on Ko and Edgar (2004)) can be used to monitor EWMA controllers.

2.1 Discrete Integral Controller

The process output is represented by the following discrete-time model

$$y_{k+1} = \overline{b_{k+1}} u_{k+1} + c_{k+1} \qquad (1)$$

Where y is the output, u is the input, b is the gain and c is the disturbance driven by white noise.

The feedback integral controller is given by

$$K = \frac{k_I}{1 - q^{-1}} \qquad (2)$$

The output u_k is obtained as

$$u_{k+1} = K\left(y_{sp} - y_k\right) = -\frac{k_I}{1-q^{-1}} y_k \qquad (3)$$

The above equation results from setting $y_{sp} = 0$. If there is no set-point change, the output of the process can now be simplified to

$$y_k = \frac{c_k}{1 + \overline{b_k} K} \qquad (4)$$

From the given data, we develop an ARMAX (Auto-Regressive Moving Average with eXogenous input) model. Using a prediction horizon p, we calculate the step response coefficients of the model (which is equivalent to the gain of the process in this case). Thus,

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_p \end{bmatrix} = -\begin{bmatrix} 0 & \cdots & \cdots & 0 \\ s_1 & 0 & \ddots & \vdots \\ \vdots & s_1 & 0 & \vdots \\ s_p & \cdots & s_1 & 0 \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_p \end{bmatrix} k_I + \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_p \end{bmatrix} \qquad (5)$$

or more simply put

$$Y = \left(I + Sk_I\right)^{-1} C \quad (6)$$

This forms the model of the given data, which can be used to calculate the optimal response. The output data impulse response is then determined, so that

$$y_k = \sum_{i=0}^{p} \Psi_i c_{k-i} \quad (7a)$$

$$\begin{bmatrix} \Psi_0 \\ \Psi_1 \\ \vdots \\ \Psi_p \end{bmatrix} = \left(I + Sk_I\right)^{-1} C \quad (7b)$$

Thus, knowing the impulse response coefficients, the disturbance vector C can be calculated.

2.2 Optimal Controller Gain

The variance of the output is given by

$$V = C^T \left(I + S^T k_I\right)^{-1} \left(I + Sk_I\right)^{-1} C \quad (8)$$

Then the optimal k_I can be obtained using Newton's method so that

$$k_{I_{new}} = k_{I_{old}} - \frac{\left(\partial V / \partial k_I\right)_{old}}{\left(\partial^2 V / \partial k_I^2\right)_{old}} \quad (9)$$

The first and second derivatives are given by

$$\frac{\partial V}{\partial k_I} = -2C^T \left(L^{-1}\right)^T SL^{-2} C = 0 \quad (10)$$

$$\frac{\partial^2 V}{\partial k_I^2} = 2C^T \left(L^{-2}\right)^T S^T S L^{-2} C + 4C^T \left(L^{-1}\right)^T S^2 L^{-3} C \quad (11)$$

The first derivative becomes zero for the optimal gain and $L = I + Sk_I$
The performance index is now given by the ratio of the variance of optimal and actual response

$$\zeta = \frac{Y_{opt}^T Y_{opt}}{Y^T Y} \quad (12)$$

and the optimal response is calculated by

$$y_{k_{opt}} = \left(\frac{1 + \overline{b_k} k_I}{1 + \overline{b_k} k_{I_{opt}}} \right) y_k \quad (13)$$

The normalized performance index has the range of $0 < \zeta \le 1$, and $\zeta = 1$ indicates the best performance under Integral Control. With this definition, $1-\zeta$ indicates the maximum fractional reduction in the output variance.

2.3 EWMA Controller

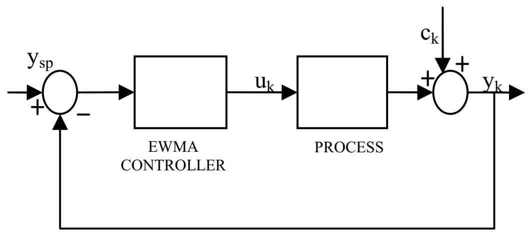

Fig. 1: EWMA controlled run-to-run process

The system shown above in Figure 1 is one controlled by a standard EWMA controller (Campbell *et al.*, 2002). The equations are as follows (with similar notations):

$$\overline{y_{k+1}} = \overline{b_{k+1}} u_{k+1} + c_{k+1} \quad (14)$$

The observer updates the disturbance c_{k+1} using an EWMA formula which is

$$c_{k+1} = \lambda \times \left(y_k - \overline{b_k} u_{k+1}\right) + \left(1-\lambda\right) \times c_k \quad (15)$$

The input is now given by (with y_{sp} as the target)

$$u_{k+1} = \frac{y_{sp} - c_{k+1}}{\overline{b}_{k+1}} \quad (16)$$

The actual gains are determined before the lot is processed using historical data.

$$b_k = \frac{y_k}{u_k} \quad (17)$$

For a pure gain system, the EWMA controller is equivalent to a discrete integral controller with gain k_I (Box, 1993), with

$$k_I = \frac{\lambda}{b_{mean}} \quad (19)$$

Thus by representing the process data as one controlled by a discrete integral process, the performance index of an EWMA controlled process may be obtained.

3. EXAMPLE

An etch process at AMD[1] was considered for analysis. The governing equations and input – output variables are also defined. The process model used for this process is as follows

EtchDepth = EtchRate*EtchTime + Bias (20)

The rate is updated by EWMA as given in the previous section. Accordingly, the manipulated variable is time, while the controlled variable is (EtchDepth – Bias). The algorithm for calculating the performance index essentially utilizes the moving window approach, i.e., considering only the last 'n' data points in time so that the performance index calculated represents the current state of the process. The data considered was for etch processes run with different equipment each time. Thus each type of etch process was analyzed separately to evaluate which equipment performed better than others. About 29 different etch processes at AMD were compared.

4. RESULTS

The following three types of results could be obtained by the above developed method. Not only can the method be used to compare different processes, the effect of delay is also demonstrated. Also the performance of a process can be tracked over time.

4.1 Distribution of performance indices

The etch processes showed a distribution of performance indices. The performance index usually lies between 0.8 and 1. Figure 2 shows the distribution of the processes considered in each range of performance index.

Although most processes were found to lie in the 0.9 to 1 range, the remaining processes were found to be uniformly distributed in the 0.1 to 0.8 region. Thus, majority of the processes were found to be operating sub-optimally.

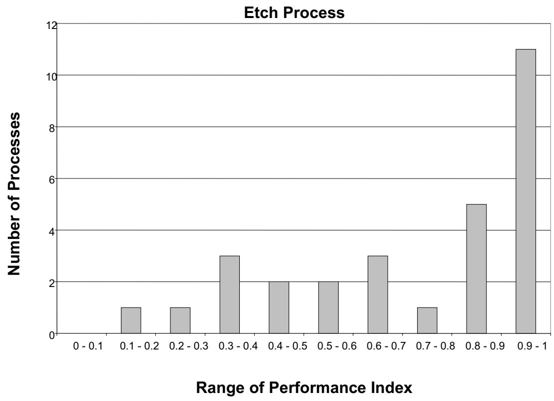

Fig. 2: Distribution of performance

4.2 Effect of delay

When no delay is considered in the calculation of the performance index, the algorithm assumes that the only reason for suboptimal performance is the controller itself. But if we do consider a delay of one or more, the algorithm takes into account that this delay is responsible for some degradation in performance. This is because the delay is considered during the selection of the ARMAX model for the data. Thus, with increasing delay, the performance index goes asymptotically to 1.0. This is because, as the metrology delay increases, it becomes the primary reason for suboptimal performance. In other words, the controller cannot work efficiently beyond a certain threshold. Consider the example shown in Figure 3.

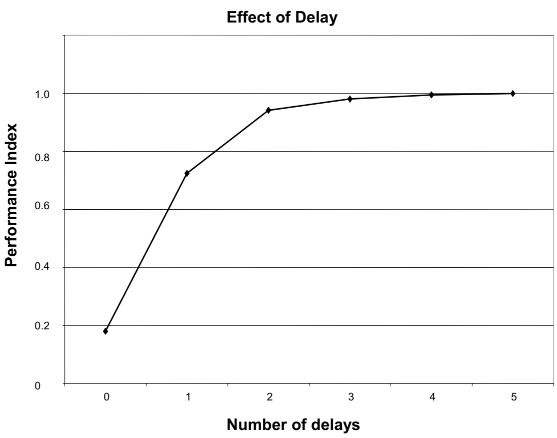

Fig. 3: Change in performance index with delay

[1] Advanced Micro Devices, Inc.

A performance index of one in this case does not indicate optimality but instead points to the delay in the process. Thus if the process has a significant amount of delay, expectations of optimal performance from the process must be greatly reduced.

4.3 Change in performance over time

Moving windows were used to study the change in performance of the process. Following is a sample chart which tracks the performance index over time for a moving window of 50 points. In Figure 4, the dots are the actual values of the index while the continuous line is the graphical trend for the thread with a 5-point moving average.

Figure 4 shows the decline in performance of the thread with time. A sudden degradation in performance is seen to have occurred mid-way in the process. Thereafter the performance is on the decline.

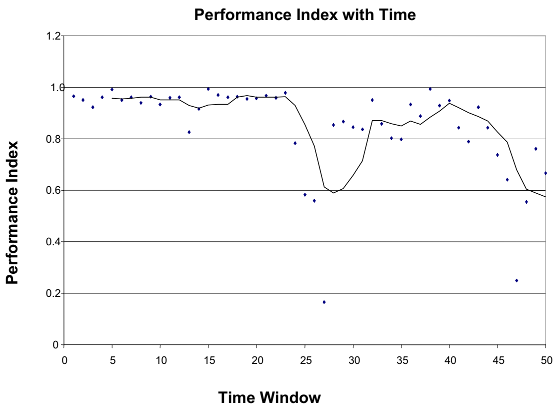

Fig. 4: Change in performance with time

5. CONCLUSIONS AND FUTURE WORK

The achievable performance bound was proposed for use in assessing and monitoring single-loop EWMA control loop performance. For this purpose, an iterative solution was derived that gives the best achievable performance in terms of the closed-loop input-output data and the process model. An explicit solution was derived as a function of EWMA settings. A performance index was defined based on the best achievable performance for use as a realistic performance measure in the single-loop EWMA control systems. An example showed the utility of the proposed method for the effective performance assessment of the existing controller as also for comparing the performance of different processes.

This work is one of the first applications of performance assessment techniques to run-to-run control systems. In the future, methods for non-EWMA processes can be developed. Also, most run-to-run processes in semiconductor manufacturing tend to have variable time delays. This aspect could be further explored and new techniques formulated to incorporate this variable delay. Also the next step in performance assessment needs to be suggested, viz. having determined which control loops perform sub-optimally, remedial steps must be outlined.

NOMENCLATURE

b_k	= Actual gain
$\overline{b_k}$	= Predicted gain
b_{mean}	= Average gain used
c_k	= Disturbance
C	= Vector of disturbance estimates
d_k	= Actual measurement
$\overline{d_k}$	= Predicted measurement
h_k	= Bias
I	= Identity Matrix
k_I	= Integral controller gain
$K(q^{-1})$	= Integral controller
L	= $I + Sk_I$
q^{-1}	= Backward shift operator
s_i	= Step response coefficient
S	= Matrix of step response coefficients
T	= Target
u_k	= Input used
V	= Variance
y_k	= Normalized output
y_{sp}	= Set-point for normalized output
Y	= Vector of normalized output values
Y_{opt}	= Vector of optimal output values

Subscripts

I	= Integral controller
k	= Time
mean	= Average from a set of given values
new	= Value for current iteration
old	= Value from previous iteration
opt	= Optimal value
p	= Prediction horizon
sp	= Set-point

Greek Symbols

λ	= EWMA weighting
ζ	= Performance Index
Ψ_i	= Impulse response coefficient

REFERENCES

Box G. E. P. (1993). Process Adjustment and Quality Control. *Total Quality Management*, **4**, 2.

Campbell, J. W., S. K. Firth, A. J. Toprac and T.F. Edgar (2002). A Comparison of Run-to-Run Control Algorithms. *Proceedings of the American Control Conference*, 2150.

Desborough, L., and T. J. Harris (1992). Performance Assessment Measures for Univariate Feedback Control. *Canadian Journal of Chemical Engineering*, **70**, 1186.

Desborough, L., and T. J. Harris (1993). Performance Assessment Measures for Univariate Feedforward/Feedback Control. *Canadian Journal of Chemical Engineering*, **71**, 605.

Edgar, T. F., S. W. Butler, W. J. Campbell, C. Pfeiffer, C. Bode, S. B. Hwang, K. S. Balakrishnan, and J. Hahn (2000). Automatic Control in Microelectronics Manufacturing: Practices, Challenges and Possibilities. *Automatica*, **36**, 1567.

Eriksson, P.-G., and A. J. Isaksson (1994). Some Aspects of Control Loop Performance Monitoring. *Proceedings of the IEEE Conference on Control Applications*, 1029.

Harris, T. J. (1989). Assessment of Control Loop Performance. *Canadian Journal of Chemical Engineering*, **67**, 856.

Harris, T. J., F. Boudreau, and J. F. MacGregor (1996a). Performance Assessment of Multivariable Feedback Controllers. *Automatica*, **32**, 1505.

Harris T. J., C. T. Seppala, and L. D. Desborough (1999). A Review of Performance Monitoring and Assessment Techniques for Univariate and Multivariate Control Systems. *Journal of Process Control*, **9**, 1.

Huang, B., S. L. Shah, and E. K. Kwok (1995). Online Control Performance Monitoring of MIMO Processes. *Proceedings of the American Control Conference*, 1250.

Huang, B., S. L. Shah, and E. K. Kwok (1997). Good, Bad or Optimal? Performance Assessment of Multivariable Processes. *Automatica*, **33**, 1175.

Ko, B.-S., and T. F. Edgar (1998). Assessment of Achievable PI Control Performance for Linear Processes with Dead Time. *Proceedings of the American Control Conference*, 1548.

Ko, B.-S., and T. F. Edgar (2000). Performance Assessment of Multivariable Feedback Control Systems. *Proceedings of the American Control Conference*, 4373.

Ko, B.-S., and T. F. Edgar (2004). PID Control Performance Assessment: The Single-Loop Case. *AIChE Journal*, **50**, 1211.

Qin, S. J. (1998). Controller Performance Monitoring – A Review and Assessment. *Computers and Chemical Engineering*, **23**, 173.

Salsbury, T. I. (2005). A Practical Method for assessing the Performance of Control Loops subject to random load changes. *Journal of Process Control*, **15**, 393.

Stanfelj, N., T. E. Marlin, and J. F. MacGregor (1993). Monitoring and Diagnosing Process Control Performance: The Single-Loop Case. *Industrial Engineering Chemistry Research*, **32**, 301.