Double EWMA Run-to-Run Controller for High Mixed Product Mode

Shu-Kai S. Fan†
Department of Industrial Engineering & Management
National Taipei University of Technology (Taipei Tech), Taipei, Taiwan
Tel: (+886) 2- 2771-2171 ext. 2382, Email: morrisfan@ntut.edu.tw
Yu-Ling Liao
Department of Industrial Engineering & Management
National Taipei University of Technology (Taipei Tech), Taipei, Taiwan
Email: a39815118@gmail.com

Abstract. In the semiconductor manufacturing industry, the mixed product mode has been a prominent characteristic and drawn considerable attention in academic recently. The most common practice is to classify the situation of a specific tool used and a specific product manufactured termed as a "thread", and creates a run-to-run controller for each thread. Typically, there will be "hundreds or thousands of threads" for each operation. The research scope of this article is to develop a new double exponentially weighted moving average run-to-run (RtR) control strategy that is able to effectively reduce various types of process disturbances quickly and keep the process means close to their desirable levels with relatively small effort for the high mixed product mode. A preliminary simulation study based on an equipment model from the literature was conducted to demonstrate the proposed control strategy.

Keywords: semiconductor manufacturing, run-to-run-control, high mixed product mode, double exponentially weighted moving average

1. INTRODUCTION

Most run-to-run (RtR) control algorithms are devised based on the assumption that there is only a single product fabricated in the manufacturing line. However, an actual semiconductor manufacturing facility is an assembly line consisting of a sequence of operations performed by parallel machines that manufacture products of different types and grades. The most common practice is to classify the situation of a specific tool used and a specific product manufactured termed as a "thread", and creates an RtR controller for each thread. Typically in common semiconductor practice, there will be "hundreds or thousands of threads" for each operation.

Ai et al. (2009) investigated the mixed-product drifted process and found that if the break length of a product is moderately large, then at the beginning runs of each cycle, the process output will far deviate from the target value. Towards this end, the drift-compensatory approach based on threaded exponentially weighed moving average (t-EWMA) control, or called threaded predictor-corrector control (t-PCC) was proposed to deal with large deviations at the beginning runs of specific cycle process in a mixed production. Zheng et al. (2010) proposed a cycled resetting algorithm for the discount factor, i.e., CR-EWMA algorithm, to reduce the large deviations as well as to achieve the minimum asymptotic variance control. The discount factor resetting fault tolerant (RFT) approach was used to handle the step fault. Ai et al. (2010) proposed a cycle forecasting EWMA (CF-EWMA) approach to deal with the large deviations in the first few runs of each cycle under drift disturbance. The CF-EWMA approach utilized the slop of estimations of disturbance and the length of break products to compensate for the deviations for the first run of campaign product in the next cycle. Based on the previous review, the main objective of this article is to present a feasibility study of the double EWMA strategy for mixed-product mode. The concept of product-based EWMA control will be elaborated in the next section.

2. MIXED PRODUCT MODEL

In Zheng et al. (2006), a simplified case is considered where only two products are produced on a single tool. The production schedule contains several cycles of i runs each, where j runs are for Product 1 and $(i-j)$ runs for Product 2. Here, j indicates the campaign length of Product 1, $(i-j)$ indicates the break length for Product 1 and vice versa for Product 2. The input-output relation for these two products on a given tool is assumed to follow a simple linear regression model with intercepts α_1, α_2 and slopes β_1, β_2. However, the two models share the same disturbance model η, as defined by

$$Y_{it+n} = \begin{cases} \alpha_1 + \beta_1 X_{it+n} + \eta_{it+n}, \ 1 \le n \le j, \\ \alpha_2 + \beta_2 X_{it+n} + \eta_{it+n}, \ j+1 \le n \le i, \end{cases} \qquad (1)$$

where t denotes the number of cycle, Y_{it+n} indicates the outputs of the two products, X_{it+n} indicates the control action at run $it+n$.

2.1 Tool Based EWMA Control

In a tool-based control, the variation between products is not taken into account. It implies that each run of different products on the same tool shares a single "noise" estimate $\tilde{\eta}_{it+n}$ which is calculated based on input-output data. Thus, both the EWMA filter in (2) and the deadbeat control in (3) are executed with the filtered noise disturbance obtained from the last run no matter which product has been produced. Consider a simple case with two products to be produced on a single tool using the tool-based control scheme as shown in Figure 1.

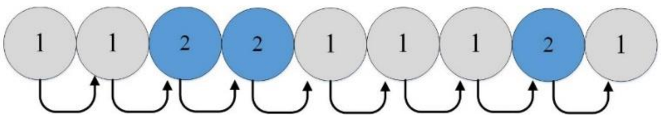

Figure 1. Tool-based control scheme.

$$\tilde{\eta}_{it+n} = \begin{cases} \lambda_1 (y_{it+n} - b_1 x_{it+n} - a_1) + (1-\lambda_1)\tilde{\eta}_{it+n-1}, \ 1 \le n \le j, \\ \vdots \\ \lambda_k (y_{it+n} - b_2 x_{it+n} - a_2) + (1-\lambda_k)\tilde{\eta}_{it+n-1}, \ j+1 \le n \le i. \end{cases} \qquad (2)$$

$$X_{it+n} = \begin{cases} \dfrac{T_1 - \hat{\eta}_{it+n-1} - a_1}{b_1}, \ 1 \le n \le j, \\ \dfrac{T_2 - \hat{\eta}_{it+n-1} - a_2}{b_2}, \ j+1 \le n \le i. \end{cases} \qquad (3)$$

Here, a_1, a_2, b_1, and b_2 are the fixed model parameters; λ_1 is the discount factor of the EWMA algorithm; X_{it+n} is the control action; T_1 and T_2 are the target values for Products 1 and 2.

2.2 Product Based EWMA Control

In "product-based" control, the EWMA filter action is performed with respect to the last run on which the same product is processed instead of the previous run in which a different product have been processed. Hence, for Product 1, the filtered plant noise can be expressed as in (3) and the deadbeat control actions for Product 1 can be formulated as in (4). Consider a simple case that two products are produced on a single tool and the product-based control scheme is demonstrated in Figure 2.

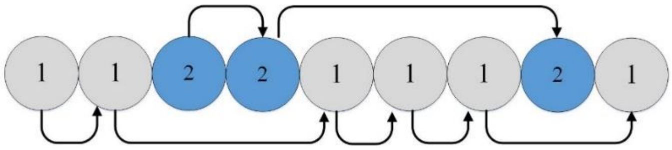

Figure 2. Product-based control scheme.

$$\hat{\eta}_{it+n} = \begin{cases} \lambda (Y_{it+1} - b_1 X_{it+1} - a_1) + (1-\lambda)\hat{\eta}_{i(t-1)+j}, \ n = 1, \\ \lambda (Y_{it+n} - b_1 X_{it+n} - a_1) + (1-\lambda)\hat{\eta}_{it+n-1}, \ n = 2, \ldots, j. \end{cases} \qquad (3)$$

$$X_{it+n} = \begin{cases} \dfrac{T - \tilde{\eta}_{(i-1)t+j} - a_1}{b_1}, \ n = 1, \\ \vdots \\ \dfrac{T - \tilde{\eta}_{it+n-1} - a_1}{b_1}, \ n = 2,3,\ldots, j. \end{cases} \qquad (4)$$

It is worth noting that in this control scheme, the output of Product 1 becomes independent of what is produced in other runs $it+n$, where $j+1 \le n \le i_t$ for different products.

2.3 Disturbance Model

A disturbance to a process can be defined as any occurrence or series of occurrences that result in a change in the processing conditions. This disturbance is often very hard to measure; therefore, the disturbance information may not be immediately available to the controller. At the same time, the controller must reduce the effects of the disturbance on the process output. In this paper, it is assumed the process disturbance follows an integrated first-order moving average series (i.e., $\text{IMA}(1,1)$) with deterministic linear drift δ frequently seen as disturbances in unstable processes. The disturbance model can be described as follows:

$$\eta_t = \eta_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1} + \delta \qquad (5)$$

where θ is the IMA parameter; ε_t is an independent and identically distributed white noise sequence with zero mean and constant variance; δ is the rate of deterministic drift.

3. PRODUCT-BASED DOUBLE EWMA CONTROL

In "threaded" RtR control, the EWMA filter is acted according to the last run on which the same product had been processed instead of the previous run on which the different product have just been processed. Butler and Stefani (1994) proposed the double EWMA controller (also called predictor-corrector controller). Now we apply the double EWMA control to the high mixed product mode serving as a "threaded" RtR control. The double EWMA controller is to recursively update the estimate of unknown parameters α and δ so as to provide adequate control action on X_{it+n}. In this light, for Product 1, the product-based double EWMA filter and deadbeat control law can be presented as follows:

$$a_{\sum_{c=0}^{t-1} i_c+n} = \begin{cases} \lambda_1 (Y_1 - b_1 x_1) + (1-\lambda_1) a_{1,0}, \\ \quad n=1 \text{ and } t=0 \\ \lambda_1 \left(Y_{\sum_{c=0}^{t-1} i_c+1} - b_1 x_{\sum_{c=0}^{t-1} i_c+1} \right) + (1-\lambda_1) a_{\sum_{c=0}^{t-2} i_c + j_{1,t-1}}, \\ \quad n=1 \text{ and } t \geq 1 \\ \lambda_1 \left(Y_{\sum_{c=0}^{t-1} i_c+n} - b_1 x_{\sum_{c=0}^{t-1} i_c+n} \right) + (1-\lambda_1) a_{\sum_{c=0}^{t-1} i_c+n-1}, \\ \quad n=2,3,\ldots, j_{1,t} \end{cases} \tag{6}$$

$$D_{\sum_{c=0}^{t-1} i_c+n} = \begin{cases} \lambda_2 (Y_1 - b_1 x_1 - a_{1,0}) + (1-\lambda_2) D_{1,0}, \\ \quad n=1 \text{ and } t=0 \\ \lambda_2 \left(Y_{\sum_{c=0}^{t-1} i_c+1} - b_1 x_{\sum_{c=0}^{t-1} i_c+1} - a_{\sum_{c=0}^{t-1} i_c} \right) + (1-\lambda_2) D_{\sum_{c=0}^{t-2} i_c + j_{1,t-1}}, \\ \quad n=1 \text{ and } t \geq 1 \\ \lambda_2 \left(Y_{\sum_{c=0}^{t-1} i_c+n} - b_1 x_{\sum_{c=0}^{t-1} i_c+n} - a_{\sum_{c=0}^{t-1} i_c+n-1} \right) + (1-\lambda_2) D_{\sum_{c=0}^{t-1} i_c+n-1}, \\ \quad n=2,3,\ldots, j_{1,t} \end{cases} \tag{7}$$

$$X_{\sum_{c=0}^{t-1} i_c+n} = \begin{cases} \dfrac{T_1 - a_{1,0}}{b_1}, & n=1 \text{ and } t=0 \\ \dfrac{T_1 - a_{\sum_{c=0}^{t-2} i_c + j_{1,t-1}} - D_{\sum_{c=0}^{t-1} i_c+n}}{b_1}, & n=1 \text{ and } t \geq 1 \\ \dfrac{T_1 - a_{\sum_{c=0}^{t-1} i_c+n} - D_{\sum_{c=0}^{t-1} i_c+n}}{b_1}, & n=2,3,\ldots, j_{1,t} \end{cases} \tag{8}$$

In (6-7), λ_1 and λ_2 are called discount factors that are used to recursively update the intercept and deterministic drift parameters via the EWMA filter, respectively. The parameter b_1 is the estimate of the slope β_1 of Product 1 and the parameter $a_{1,0}$ is the initial estimate of the intercept α_1, both of which can be obtained from design of experiments and analysis (DOE) during a pre-control stage.

4. SIMULATION AND RESULTS

In semiconductor manufacturing industry, mixed product production on the same tool is the ordinary practice. For simplicity, the following example will only discuss two products manufactured on the same tool with different cycle lengths. The production schedule consists of cycles of i_t runs in cycle *t*, in which $j_{1,t}$ and $j_{2,t}$ runs are scheduled to produce product 1 and product 2 in cycle *t*, respectively. Besides, i_t, $j_{1,t}$, $j_{2,t}$ may have different values for different cycle *t*. To be specific, $j_{1,t}$ and $i_t - j_{1,t}$ are defined as the campaign length and the break length with respect to product 1 in cycle *t*. Likewise, $j_{1,t}$ and $i_t - j_{1,t}$ are defined as the break length and the campaign length with respect to product 2 in cycle *t*.

4.1 The Performance Measures of RtR Control

In this article, we take five statistics associated with the controlled output as the performance measures to evaluate control consistency and stability in high mixed product manufacturing. Simulation is conducted to compare the performance of the presented product-based double EWMA control against the tool-based EWMA control under IMA(1,1) with the drift disturbance model. Performance measures of controlled outputs include mean, variance, mean squared error from target (MSE) and mean absolute deviation from mean (MAD), and mean absolute deviation from target (MAD_target).

4.2 Equipment Model for Simulation

Assume that the input-output relationship for the products on the given tool is linear with different intercepts, α_1 and α_2, and slopes, β_1 and β_2. At the same time, both products share the same tool disturbance and measurement noise. The disturbance model is collectively denoted by one variable by $\{\eta_t\}$, over the total cycles, which follows an IMA(1,1) series with deterministic linear drift δ. First, we need to construct the equipment model for simulation as follows:

$$Y_{\sum_{c=0}^{t-1} i_c+n} = \begin{cases} \alpha_1 + \beta_1 x_{\sum_{c=0}^{t-1} i_c+n} + \eta_{\sum_{c=0}^{t-1} i_c+n}, & 1 \leq n \leq j_{1,t}, \\ \alpha_2 + \beta_2 x_{\sum_{c=0}^{t-1} i_c+n} + \eta_{\sum_{c=0}^{t-1} i_c+n}, & j_{1,t} < n \leq i_t, \end{cases} \tag{9}$$

where Y denotes the controlled output. To model the change

$$\eta_{\sum_{c=0}^{t-1} i_c+n} = \begin{cases} \eta_{\sum_{c=0}^{t-1} i_c+n-1} + \delta_1 + (1-\theta_1 \mathbf{B}) \varepsilon_{\sum_{c=0}^{t-1} i_c+n}, & 1 \leq n \leq j_{1,t}, \\ \eta_{\sum_{c=0}^{t-1} i_c+n-1} + \delta_2 + (1-\theta_2 \mathbf{B}) \varepsilon_{\sum_{c=0}^{t-1} i_c+n}, & j_{1,t}+1 \leq n \leq i_t. \end{cases} \tag{10}$$

in tool condition, a noise disturbance $\{\eta_t\}$ obeying an IMA(1,1) series with deterministic linear drift δ is used as shown below:

where δ_1 and δ_2 are the rates of deterministic drift for Products 1 and 2, respectively.; $\mathbf{B}$ is the backward shift operator.

4.3 An Illustrating Example of High Mixed Product Mode

Consider two products (Product 1 and Product 2) are manufactured on the same tool for 5 cycles, and each cycle has the following parameters:

$(j_{1,0}, j_{2,0}, i_0) = (100,100,200)$ for Cycle 0,
$(j_{1,1}, j_{2,1}, i_1) = (50,80,130)$ for Cycle 1,
$(j_{1,2}, j_{2,2}, i_2) = (150,100,250)$ for Cycle 2,
$(j_{1,3}, j_{2,3}, i_3) = (100,100,200)$ for Cycle 3,
$(j_{1,4}, j_{2,4}, i_4) = (100,100,200)$ for Cycle 4.

Suppose that the true parameters of the equipment models for Product 1 and Product 2 are $(\alpha_1, \alpha_2) = (2,1)$ and $(\beta_1, \beta_2) = (2,1)$. To reflect the model mismatch, the initial parameter estimates are intentionally set to $(a_1, a_2) = (1,2)$ and $(b_1, b_2) = (1,2)$. The parameters of the disturbance and measurement noise are assumed to be $\theta = 0.5$, $\sigma^2 = 0.1^2$, and drift rate $\delta = 0.1$. The target value of the controlled outputs are set to $(T_1, T_2) = (0,5)$ and the discount factors are chosen to be $(\lambda_1, \lambda_2) = (0.6, 0.6)$ according to Ai et al. (2010). Note that the discount factors of (λ_1, λ_2) correspond to the intercept and drift update, respectively, in the proposed product based double EWMA controller. On the other hand, the discount factors of (λ_1, λ_2) correspond to Product 1 and Product 2, respectively, in the tool based EWMA controller.

The controlled outputs returned by using the proposed product-based double EWMA controller for the high mixed product mode from Cycle 0 to Cycle 4 are exhibited in Figure 3. From the figure, it is obvious that at the very beginning of each cycle, the process output deviates somewhat from target, while after several runs of oscillation, the process outputs converge to the target value gradually. Lastly, the target value is accurately attained. As time goes by, the deviation from target in the initial runs of late cycles becomes aggravated, especially Product 2 in Cycles 2, 3 and 4.

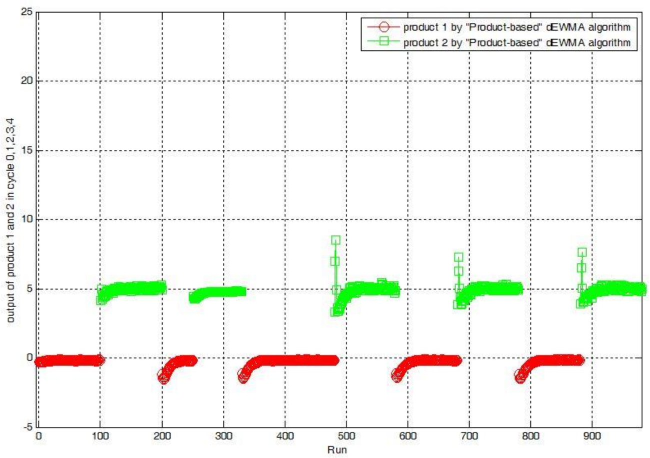

Figure 3. Controlled outputs of the product-based double EWMA control for Products 1 and 2.

The control performance based on the five performance measures generated by using the proposed product-based double EWMA controller is tabulated in Table 1. Basically, the controller maintains deviation from target for both products within ± 0.25. Average variances of Product 1 and Product 2 are about 0.0910 and 0.1619; average MSEs of product 1 and product 2 are about 0.2051 and 0.1838.

Measure \ Cycle	t=0		t=1		t=2		t=3		t=4	
	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2
Mean	-0.1928	4.9472	-0.5034	4.7110	-0.2754	4.8618	-0.3200	4.9185	-0.3278	4.9248
Variance	0.0014	0.0407	0.1734	0.0202	0.0732	0.3938	0.0920	0.1700	0.1149	0.1849
MSE	0.0386	0.0431	0.4233	0.1035	0.1486	0.4090	0.1935	0.1749	0.2213	0.1887
MAD	0.0257	0.1428	0.3328	0.1007	0.1641	0.3510	0.2121	0.2305	0.2386	0.2462
MAD_target	0.1928	0.1336	0.5034	0.2890	0.2754	0.3405	0.3200	0.2271	0.3278	0.2470

Table 1. Simulation results of the product-based double EWMA controller for Product 1 and Product 2 from Cycle 0 to Cycle 4.

For the comparison purpose, the controlled outputs returned by using the tool-based EWMA controller for the high mixed product mode from Cycle 0 to Cycle 4 are illustrated in Figure 4.

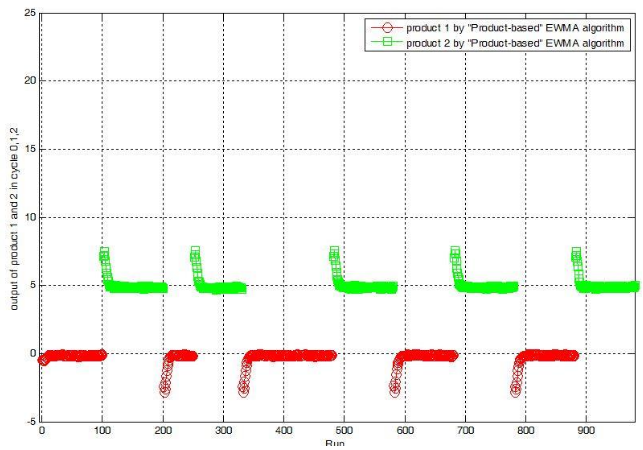

Figure 4. Controlled outputs of the tool-based EWMA control for Products 1 and 2.

In comparison of Table 4 to Table 3, it can be obviously seen that the tool-based controller requires much more transient runs than the proposed product-based double EWMA controller from the break phase to the campaign phase for both products except Product 1 in Cycle 0. Even so, after significant oscillation, the controlled outputs of both products are still driven towards the target values. The control performance based on the five performance measures generated by using the tool-based EWMA controller is listed in Table 2.

measure \ Cycle	**t=0**		**t=1**		**t=2**		**t=3**		**t=4**	
	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2
Mean	-0.1819	4.9877	-0.4315	4.9961	-0.2447	4.9994	-0.2880	4.9913	-0.2857	4.9907
Variance	0.0071	0.2439	0.4766	0.3146	0.1702	0.2591	0.2396	0.2440	0.2571	0.2359
MSE	0.0401	0.2416	0.6533	0.3107	0.2290	0.2565	0.3202	0.2416	0.3362	0.2336
MAD	0.0502	0.2470	0.4467	0.3070	0.1717	0.2544	0.2421	0.2439	0.2551	0.2324
MAD_target	0.1819	0.2566	0.4315	0.3099	0.2447	0.2549	0.2880	0.2505	0.2857	0.2398

Table 2. Simulation results of the tool-based EWMA controller for Product 1 and Product 2.

Basically, the controller maintains deviation from target for both products within ± 0.3. Average variances of Product 1 and Product 2 are about 0.231 and 0.2595; average MSEs of product 1 and product 2 are about 0.3157 and 0.2568. As evidenced by the comparison results between Table 1 and Table 2, it has been clearly demonstrated that the proposed product-based double EWMA controller greatly outperforms the tool-based EWMA controller in the mixed product mode study where two products are processed with different equipment models, different target values and different model mismatches.

An interesting observation that bears further scrutiny is that the proposed product-based double EWMA controller directly manipulates the output of Product 2 somewhat below the target value in a couple of runs in transient period and then bring the outputs towards the target value using the double EWMA filter. By contrast, the tool-based EWMA controller hinges purely on the single EWMA filter to gradually bring the outputs of Product 2 towards the target value.

5 CONCLUSION AND FUTURE RESEARCH

Recently, “threaded” RtR control has drawn considerable attention in the semiconductor camp, which attempts to maintain processes target and enhance yield for the high mixed product mode. The high mixed product mode has become a practicably prominent characteristic in the advanced semiconductor manufacturing practice. In typical production schedule, there are a variety of product types scheduled to be processed on the same tool. With the traditional RtR control algorithm, it is extremely difficult for practitioners to determine the current optimum recipe in connection with the last cycle in which the same product has been processed. The major reason is that there always has a long break length among different products in the high mixed product mode.

In this article, a new RtR control strategy for the high mixed product mode is addressed. A modified product-based double EWMA controller well suited to the parameter changes between runs, products and cycles, is developed. A simulation study of the high mixed product mode where two different products with different equipment models are processed on the same tool is conducted to illustrate the presented product-based double EWMA controller. Five different performance measures are designed to compare the control performance between the proposed controller and the benchmark controller, i.e., the tool-based EWMA controller. The simulation results show that the proposed controller performs much better than the benchmark controller in every performance measures.

Building upon this research, there still a number of topics deserving future research in this area. For instance, using design of experiment (DOE) or machine learning adaptively finds out the automated optimal discount factor of EWMA control for different products. Future research will also need to focus on the control mechanism under metrology delay or virtual metrology (VM), which has shown an important influence on advanced controller’s performance since it is a case of practical relevance.

REFERENCES

Ai, B., Zheng, Y., Jang, S. S., Wang, Y. W., Ye, L., and Zhou, C. J., "The optimal drift compensatory and fault tolerant approach for mixed-product run-to-run control," Journal of Process Control, vol. 19, no. 8, 2009, pp. 1401–1412.

Ai, B., Zheng, Y., Wang, Y., Jang, S. S., and Song, T., "Cycle forecasting EWMA (CF-EWMA) approach for drift and fault in mixed-product run-to-run process," Journal of Process Control, vol. 20, no. 5, 2010, pp. 689-708.

Butler, S. W., and Stefani, J. A., “Supervisory run-to-run control of a polysilicon gate etch using in situ ellipsometry,” IEEE Transactions on Semiconductor Manufacturing, vol. 7, no. 2, 1994, pp. 193-201.

Zheng, Y., Lin, Q. H., Wang, D. S. H., Jang, S. S., and Hui, K., "Stability and Performance Analysis of Mixed Product Run-to-Run Control," Journal of Process Control, vol. 16, no. 5, 2006, pp. 431-443.

Zheng, Y., Ai, B., and Wong, D. S. H., "An EWMA algorithm with a cycled resetting discount factor for drift and fault of high-mix run-to-run control," IEEE Transactions on Semiconductor Manufacturing, vol. 6, no. 2, 2010, pp. 229-242.