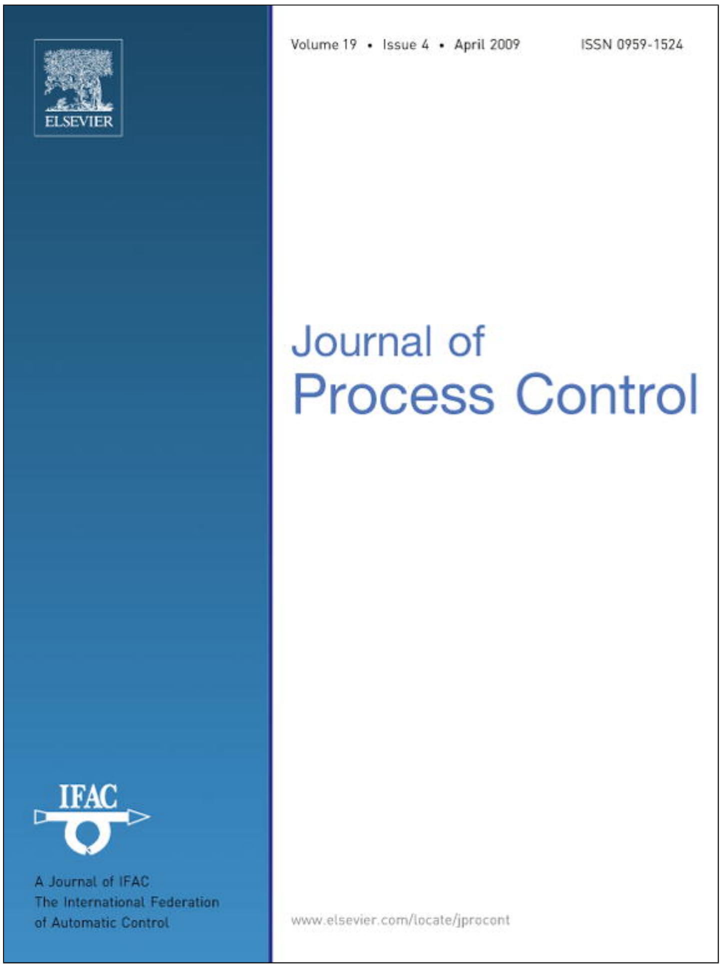
Volume 19 • Issue 4 • April 2009
ISSN 0959-1524
ELSEVIER
Journal of
Process Control
IFAC
A Journal of IFAC
The International Federation
of Automatic Control
www.elsevier.com/locate/jprocont

Journal of Process Control 19 (2009) 604–614

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Mixed product run-to-run process control – An ANOVA model with ARIMA disturbance approach

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

[a] Center for Control and Guidance Technology, Harbin Institute of Technology, Harbin 150001, China
[b] Department of Chemical Engineering, National Tsing-Hua University, Hsin-Chu 30013, Taiwan

ARTICLE INFO

Article history:
Received 16 August 2007
Received in revised form 26 July 2008
Accepted 6 September 2008

Keywords:
Run-to-run control
Mixed product system
Dynamic ANOVA
ARIMA disturbance model

ABSTRACT

A novel run-to-run control algorithm based on a dynamic analysis of variance (ANOVA) approach is proposed to deal with run-to-run (RtR) control of a high mixed operation, i.e., many different products are manufactured in many different tools. The conditions of different tools and products are identified based on the ANOVA analysis of the system output. A dynamic term in the form of an autoregressive integrated moving average (ARIMA) disturbance model is included in the process model to characterize the run-to-run disturbances such as drift, shift and/or some other unknown disturbances of different tools. It is shown from the study below that controller performance can be improved by introduction of the dynamic term, especially for products which are produced only occasionally. This makes it highly suitable for mixed product control system. An industrial example is also included to demonstrate superiority of this approach.

1. Introduction

Run-to-run (RtR) control has been identified as a key enabling technology of maintaining quality in semiconductor manufacturing. Active researches in this area have been summarized by many authors in books and review articles [1–3]. Most RtR control algorithms are 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 many different grades. The most common practice is to classify the situation of a specific tool used and a specific product manufactured as a "thread", and creates a RtR controller for each thread. Typically there will be thousands of threads for each operation. Several papers [4–6] have discussed the possibility of cross utilization of information of different threads to isolate changes of condition of tool and products. Zheng et al. [7] considered the stability of a single tool with different products. They demonstrated that exponentially weighted moving average (EWMA) control of the tool disturbance may not be stable if the effect of the tool is not stationary and the error of process gain estimates of different products are different.

Pasadyn and Edgar [5] noted that the absolute value of the product and tool disturbances cannot be estimated independently even if they are constant because each run must consist of a specific product manufactured on a specific tool. They proposed to use of monitoring wafers to condition of the tool. Firth et al. [6] proposed a method assuming that the tool noise and product noise are stationary. The resulting tool and product estimates may be biased, i.e. changes in condition of one tool may lead to changes in estimates of disturbance estimates of other tools and products. However, the estimates of different threads remained unbiased. Thus the model can be used for control but not fault detection and diagnosis. Bode et al. [8] recognized that specific regression techniques must be used to obtain unbiased estimates of tool and product states.

Analysis of variance (ANOVA) is a standard statistical tool in the area of linear modeling of multi-factor systems [9]. In [10], we have proposed a state estimation method of a mixed run plant based on analysis of variance. However, the method also assumed that the states of the tools are unchanged and a recursive Karman filter estimator is used. In this work, we shall relinquish the assumption that the condition of tool is unchanged and demonstrate that improved controller performance and diagnosis of tool conditions can be obtained.

2. Theoretical development

2.1. Plant

Fig. 1 shows the schematic plots of a "mixed run" manufacturing system. A number of products are manufactured on a number

* Corresponding authors. Tel.: +886 3 571 3697; fax: +886 3 571 5408.
E-mail addresses: ssjang@mx.nthu.edu.tw (S.-S. Jang), dshwong@che.nthu.edu.tw (D.S.H. Wong).

0959-1524/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2008.09.002

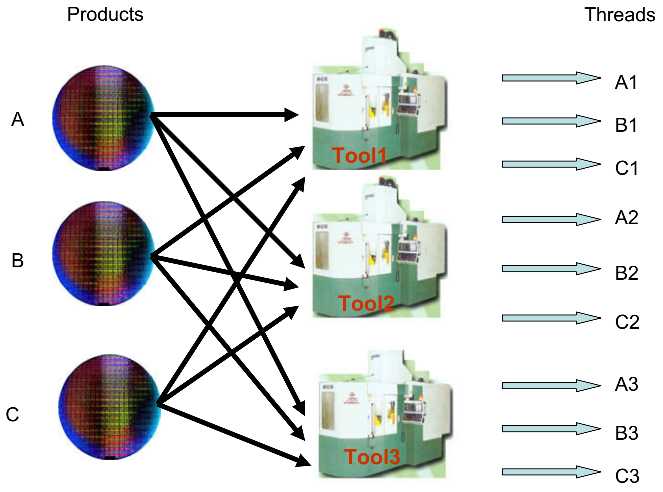

Fig. 1. Mixed product manufacturing system.

of tools. In each run, most operation variables follow a basic recipe for each product. After each run, an output y related to the quality of the product is measured. In run-to-run control, certain manipulated variable in the recipe will be adjusted based on measurement of output variables y of previous runs. Consider a simplified multi-tool and multi-product production system with a single (quality) output and a single (manipulated variable) input. A sequence of production data can be expressed as follows:

$$y(n) = f\left(x(n), c^T_{i(n)}(n), c^P_{j(n)}(n), c^{TP}_{i(n),j(n)}(n)\right) \begin{cases} i(n) = 1, \dots, I \\ j(n) = 1, \dots, J \\ n = 1, \dots, N \end{cases} \quad (1)$$

In Eq. (1), y_n and x_n denote the output and input of the nth run. f is a nonlinear function of plant behavior at the nth run which depends on the tool characteristics $c^T_{i(n)}(n)$ of the specific tool $i(n)$ used, the product characteristics $c^P_{j(n)}(n)$ of the specific product $j(n)$ being manufactured, as well as the interaction characteristics $c^{TP}_{i(n),j(n)}(n)$ of the specific pair tool $i(n)$ and product $j(n)$. In this work, it is assumed that there exist no interactions between tools and products, and the plant is operated in a very narrow range and hence a linear system can represent the plant behavior. Therefore, Eq. (1) can be simplified as

$$y(n) = bx(n) + c^T_{i(n)}(n) + c^P_{j(n)}(n) + \varepsilon(n) \quad (2)$$

where $\varepsilon(n) \in N(0, \sigma^2)$ being a zero mean white noise with variance σ^2, b is the process gain the dependency of y and x. However, the absolute value of tool and product characteristics can not be estimated independently because each run consists of a specific product manufactured on a specific tool.

2.2. Thread EWMA algorithm

Threaded EWMA controller is widely used for mixed run production. Consider a simple linear process performed in a mixed run situation on a single tool

$$y(n) = \alpha_{j(n)} + \beta x(n) + D(n), \qquad j(n) = 1, \dots, J; \quad n = 1, \dots, N \quad (3)$$

where $y(n)$ and $x(n)$ denote values of output and manipulated variable used on the nth run on the tool. $\alpha_{j(n)}$ is offset or bias term, and β is the static gain term associated with the product produced on the nth run on the tool. Let us assume that they are relatively independent of time. $D(n)$ is a stochastic noise process associated with the tool. In a threaded approach, a sequence of output and input for each specific product is re-sampled:

$$y(n_j) = \alpha_j + \beta x(n_j) + D_j(n_j) \quad (4)$$

where n_j is an index of the number of runs making the jth product that have been carried out. Given a process model $y = bx + a_j$ for each product, the offset term can be estimated by an EWMA filter

$$a_j(n_j) = \lambda(y(n_j) - bx(n_j)) + (1 - \lambda)a_j(n_j - 1) \quad (5)$$

The control action is

$$x(n_j + 1) = \frac{-a_j(n_j)}{b} \quad (6)$$

Here the quality target is assumed to be zero without loss of generality. 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 $D(n)$ from run to run, but a re-sampled series $D_j(n_j)$.

2.3. ANOVA model with ARIMA disturbance

To identify the disturbance states of the individual tools and product, the concept of ANOVA is introduced here. It is assumed that disturbance attributed to product $j(n)$ is a constant. Furthermore, in addition to having a constant offset, we assume that the condition of the tool changes from run to run. Therefore, if we sort the above data into I sequences according to the specific tool used, the production on the ith tool can be expressed as

$$y(k_i) = bx(k_i) + \mu + \tau_i + p_{j(k_i)} + \eta_i(k_i) \quad k_i = 1 \cdots K_i \quad (7)$$

In the above equation k_i is an index sequence that represents the order of runs that have been performed on the tool i. μ is the overall mean of all observed tool and product combinations. τ_i and p_j are averaged difference attributed to tool i and product j. They must satisfy the ANOVA constraints

$$\sum_{i=1}^{I} \tau_i = 0 \quad \sum_{i=1}^{J} p_j = 0 \quad (8)$$

$\eta_i(k_i)$ is a discrete-sampled stochastic dynamic disturbance attributed to tool i. The model is similar to the static ANOVA model proposed by Ma et al. [10], except for the additional run-to-run correlated stochastic term $\eta_i(k_i)$. In general, $\eta_i(k_i)$ can be represented by an ARIMA(p,d,q) process:

$$\begin{aligned} \Phi_i(B)\eta_i(k_i) &\equiv (1 + \phi_{i1}B + \cdots + \phi_{ip}B^p)(1 - B)^d \eta_i(k_i) \\ &= (1 + \theta_{i1}B + \cdots + \theta_{iq}B^q)\varepsilon_i(k_i) \equiv \Theta_i(B)\varepsilon_i(k_i) \end{aligned} \quad (9)$$

where B is the back shift operator, $\varepsilon_i(k_i) \in N(0, \sigma_i^2)$ being a zero mean white noise with variance σ_i^2. 1, φ_{i1} φ_{ip} and $\theta_{i1} \cdots \theta_{iq}$ are the coefficients of the polynomial $\Phi_i(B)$ and $\Theta_i(B)$.

2.4. Identification and run-to-run control based on the proposed model

Given a window of past operating data $y(k_1)$, $x(k_1)$, $k_1 = 1 \cdots K_1, \dots, y(k_I)$, $x(k_I)$, $k_I = 1 \cdots K_I$, by assume that $\varepsilon_i(0) = 0$, we have

$$\hat{\varepsilon}_i(k_i) = y(k_i) - \hat{b}x(k_i) - \hat{\mu} - \hat{\tau}_i - \hat{p}_{j(k_i)} - \frac{1 + \hat{\theta}_{i1}B + \cdots + \hat{\theta}_{iq}B^q}{\left(1 + \hat{\phi}_{i1}B + \cdots + \hat{\phi}_{ip}B^p\right)(1 - B)^d}\hat{\varepsilon}_i(k_i - 1) \quad (10)$$

the states of the tools and products can be estimated by minimizing the objective function

$$\begin{aligned} &\min_{\hat{b},\hat{\mu},\hat{\tau}_i,\hat{p}_j,\hat{\phi}_{i1}\cdots\hat{\phi}_{ip},\hat{\theta}_{i1}\cdots\hat{\theta}_{iq}} = \sum_{i=1}^{I}\sum_{k_i=1}^{K_i}[\hat{\varepsilon}_i(k_i)]^2 \\ &\text{s.t.} \sum_{i=1}^{I}\tau_i = 0 \quad \sum_{i=1}^{J}p_j = 0 \end{aligned} \quad (11)$$

The above optimization can be solved using standard optimization techniques such as fminsearch in MATLAB.The results of the observation window are applied to determine the control action in a control horizon $h_i = 1 \cdots H_i$ using the dead-beat control law:

$$x_i(h_i) = \frac{-\hat{\mu} - \hat{\tau}_i - \hat{p}_{j(h_i)} - \hat{\eta}_i(h_i|h_i - 1\,|)}{\hat{b}} \tag{12}$$

where $\hat{\eta}_i(h_i|h_i-1)$ is the one-step-ahead forecast of $\hat{\eta}_i(h_i-1)$. The estimated random input noise and the tool disturbance are updated when the measurement arrives

$$\hat{\varepsilon}_i(h_i) = y(h_i) - \hat{b}x(h_i) - \hat{\mu} - \hat{\tau}_i - \hat{p}_{j(h_i)} - \hat{\eta}_i(h_i|h_i-1) \tag{13}$$

$$\hat{\eta}_i(h_i) = \frac{\left(1 + \hat{\theta}_{i1}B + \cdots + \hat{\theta}_{iq}B^q\right)}{\left(1 + \hat{\phi}_{i1}B + \cdots + \hat{\phi}_{ip}B^p\right)(1-B)^d}\hat{\varepsilon}_i(h_i) \tag{14}$$

If the back shift operator shifts the error estimates beyond the start of the control horizon, the error estimates of the observation window will be used.

2.5. Estimation of missing product disturbance

It is possible that in the control horizon, products not found in the previous observation window are encountered. This product term $\hat{p}_j(h_i)$are estimated by one of the following methods:

(i) If the product 1 has appeared in one of the past windows H', but not the immediate previous window H, then find a product p_2 that appear in both windows, then

$$\hat{p}_1^H = \hat{p}_2^H + \hat{p}_1^{H'} - \hat{p}_2^{H'} \tag{15}$$

(ii) If product 1 has never appeared, the operator will determine which product in the current window is most similar to the unknown product and employ its product disturbance term.

2.6. Step disturbance (shift) compensation

For many semiconductor processes, preventive maintenance (PM) is often implemented due to degradation. This causes an abrupt change of tool conditions which can be seen as a shift or step disturbance in addition to the ARIMA(p,d,q) model. Step disturbance may be identified as "intervention event" [11]. The magnitude of the step disturbance can be estimated by maximum likelihood method. Consider a new tool disturbance

$$N_i(k_i) = \omega_i\xi(k_i) + \eta_i(k_i) \tag{16}$$

which includes an ARIMA noise $\Phi_i(B)\eta_i(k_i) = \Theta_i(B)\varepsilon_i(k_i)$ and an intervention event which is a step disturbance at time T:

$$\xi(k_i) = \begin{cases} 0 & k_i < T \\ 1 & k_i \geqslant T \end{cases} \tag{17}$$

The maximum likelihood estimator of ω_i is [11]

$$\hat{\omega}_i = \frac{\sum_{k_i=1}^N z_{k_i} w_{k_i}}{\sum_{k_i=1}^N z_{k_i}^2} \tag{18}$$

with $z_{k_i} = \pi(B)\xi(k_i)$, $w_{k_i} = \pi(B)N(k_i)$, and $\pi(B) = (\Theta(B))^{-1}\Phi(B)$. When the original tool noise η_i follows the IMA(1,1) model, $\pi(B) = (1-\theta B)^{-1}(1-B)$. The maximum likelihood estimator of ω_i becomes

$$\hat{\omega}_i = \frac{\sum_{k_i=1}^N z_{k_i} w_{k_i}}{\sum_{k_i=1}^N z_{k_i}^2} \simeq (1-\theta)\left(\sum_{s=0}^{\infty}\theta^s N_{T+1+s} - \sum_{s=0}^{\infty}\theta^s N_{T-s}\right) \tag{19}$$

In actual implementation, whenever there is a large spike at time T in error reported in the control horizon, $N_i(k_i) = y(k_i) - bx(k_i) - \hat{\mu} - \hat{\tau}_i - \hat{p}_{j(k_i)}$, the observed model errors in the control horizon, are used with intervention analysis, Eqs. (16)–(18) to estimate magnitude of the shift $\hat{\omega}_i$. Such information is then sent together with input and output data of the control horizon as new data of the next observation window. Parameter estimation is then performed with $\hat{\omega}_i\xi(k_i)$ added to the Eq. (10).

3. Simulation study

3.1. IMA(1,1) disturbance

To demonstrate the ability of the dynamic ANOVA control, a simulation example consisting of two tools and three products was used:

$$y(k_1) = x(k_1) + c_1^T(k_1) + c_{j(k_1)}^P + v_1(k_1) \qquad k_1 = 1\cdots K_1, \quad j(k_i) = 1,2,3 \tag{20}$$

$$y(k_2) = x(k_2) + c_2^T(k_2) + c_{j(k_2)}^P + v_2(k_2) \qquad k_2 = 1\cdots K_2, \quad j(k_2) = 1,2,3 \tag{21}$$

The metrology noise is normally distributed with zero mean and variance of 0.01, i.e. $v_i(k_i) \in N(0, 0.1^2)$. The product disturbances are constant with $[c_1^P, c_2^P, c_3^P] = [6, 10, 17]$. The tool disturbances are represented by a constant plus an IMA(1,1) process:

$$c_1^T(k_1) = 5 + \eta_1(k_1) \tag{22}$$

$$\eta_1(k_1) = \eta_1(k_1-1) + \varepsilon_1(k_1) - \theta_1\varepsilon_1(k_1-1) \quad \theta_1 = 0.5 \tag{23}$$

$$c_2^T(k_2) = 7 + \eta_2(k_2) \tag{24}$$

$$\eta_2(k_2) = \eta_2(k_2-1) + \varepsilon_2(k_2) - \theta_2\varepsilon_2(k_2-1) \quad \theta_2 = 0.2 \tag{25}$$

The stochastic parts of the two tool noises are also normally distributed with zero mean and variance of 0.04, i.e. $\varepsilon_1(k_1), \varepsilon_2(k_2) \in N(0, 0.2^2)$. The tool and product adopted for each run is randomly selected based on a given probability of occurrence. The two tools have equal probability distribution. The probability distributions of product A, B, and C for the examples of Section 3.1–3.3 are 0.3, 0.3 and 0.4, respectively.

In our simulation study, an observation window of 500 runs and a control horizon of 100 runs were used in the implementation of the proposed algorithm which is denoted as d-ANOVA. The observation window is used to estimate model parameters and the obtained model is adopted in the next control horizon. During the control horizon, μ, τ_i and p_j remain constant and η_i is updated recursively to capture the tool dynamics. As the control horizon moves on, the procedures are repeated which means that the proposed algorithm is implemented in a moving window approach.

The results are compared with the static ANOVA (s-ANOVA) approach proposed by Ma et al. [10], and threaded EWMA control, which is denoted by t-EWMA. It is expected a better model implies better control results. However, it should be pointed out that two main elements included our model: (i) the mixed run nature of the production and (ii) the presence of different tool disturbances, are common characteristics of real production processes. t-EWMA and s-ANOVA results are included to show how failure to such elements degrade controller performance.

Figs. 2–4 show the simulation results of three control windows for the d-ANOVA approach, s-ANOVA method and t-EWMA algorithm. For the t-EWMA method, the tuning parameter λ is 0.5 for threads on tool 1 and 0.8 for threads on tool 2. It can be seen that the d-ANOVA approach has the best performance. The variance of the output under the control of d-ANOVA approach is 0.057 which is very close to the minimum variance 0.053 [11, p. 128]. The variances of the output under the control of s-ANOVA and t-EWMA algorithm are 0.076 and 0.105, respectively.

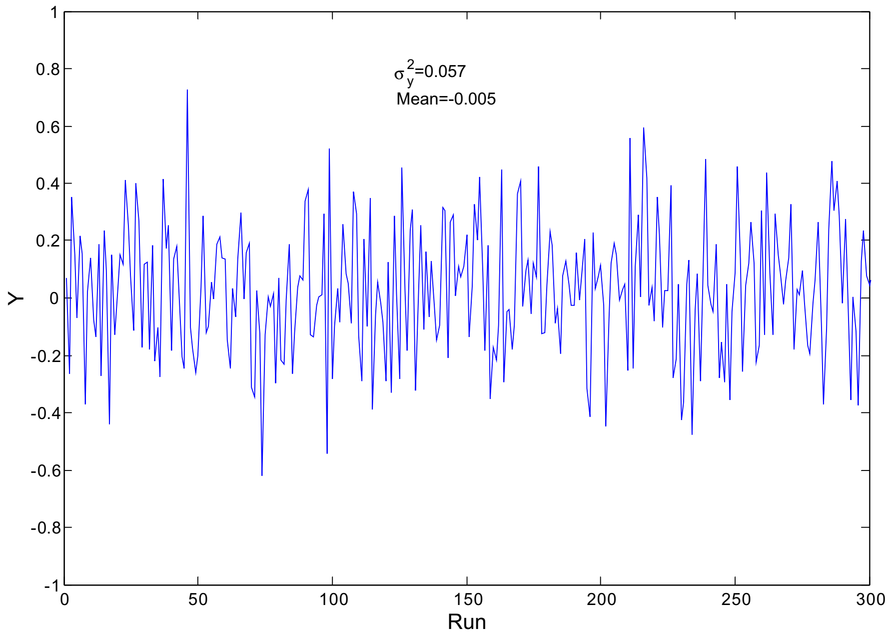

Fig. 2. Response under the control of d-ANOVA algorithm.

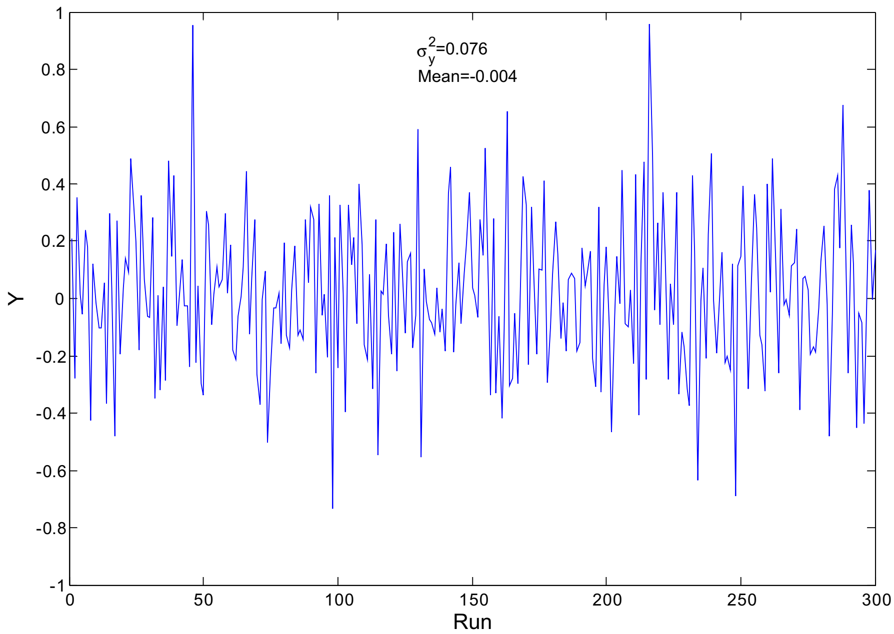

Fig. 3. Response under the control of s-ANOVA algorithm.

The actual (c_i^T) and estimated $(\hat{\mu}+\hat{\tau}_i+\hat{\eta}_i)$ dynamic trends of the two tools identified by the d-ANOVA approach are shown in Fig. 5. There is a constant offset between the real and estimated tool dynamics because only the relative states are available as we stated in above section. It can be seen that the tool dynamic trends can be correctly estimated.

3.2. Step change (shift)

When a tool undergoes a maintenance event, it would be recognized as a step disturbance in the tool disturbance. In the following example, there is a preventive maintenance for tool 1 at the 150th run which causes a shift disturbance of magnitude 4. The responses of the first 300 runs under the control d-ANOVA, s-ANOVA and t-EWMA algorithm are shown in Figs. 6–8, respectively. The mean and variances of the output under the control of d-ANOVA, s-ANOVA and threaded EWMA algorithm are (0.002, 0.069), (0.002, 0.091) and (0.013, 0.15), respectively. It should be noted that for threaded EWMA, 3 spikes are observed because each of the 3 threads on tool 1 has to encounter at least 1 run before the threaded EWMA controller begins to take effect. The s-ANOVA experiences one large spike and gradual convergence due to the recursive estimator used. The actual dynamic trends of the two tools and estimated relative dynamic trends of the two tools are shown in Fig. 9. Again it is found that tool dynamic can be correctly estimated.

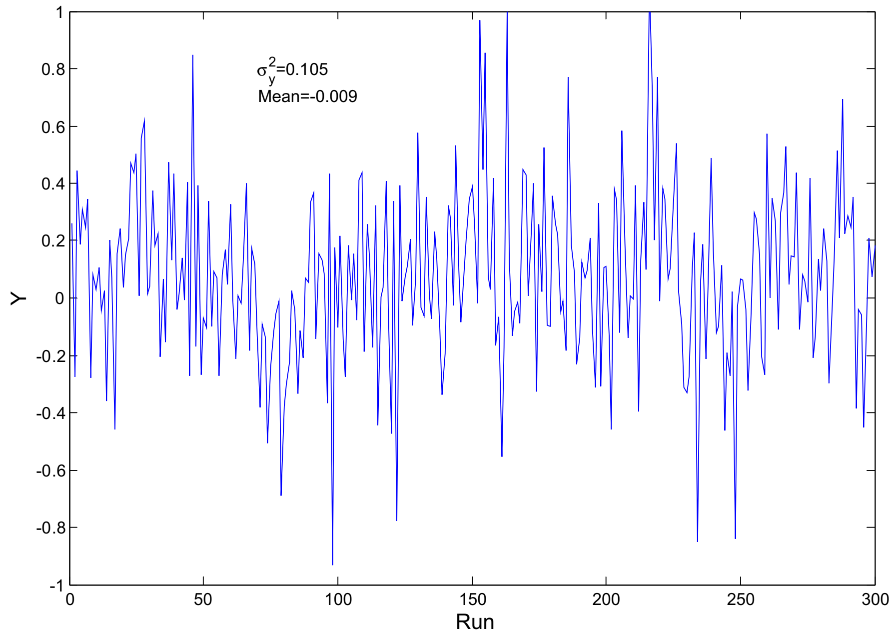

Fig. 4. Response under the control of t-EWMA algorithm.

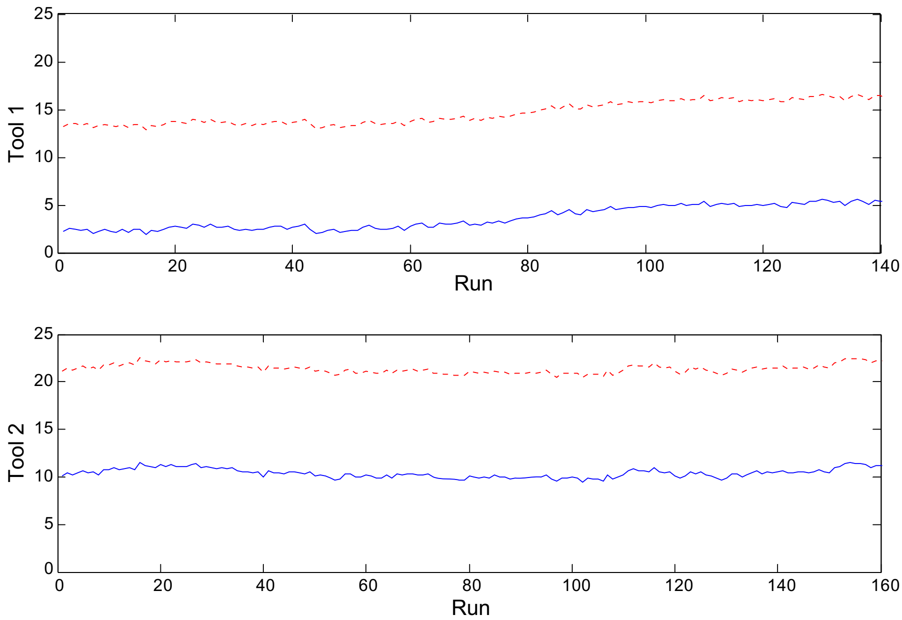

Fig. 5. Real and estimated tool effects (real —, estimated - - -).

The PM magnitude can be estimated using the method outlined in Section 2.6, an estimated shift of $\hat{\omega} = 3.99$ was found. Figs. 10 and 11 show the breakdown of the tool state $\hat{\mu} + \hat{\tau}_i + \hat{\eta}_i$ with and without this intervention analysis of 1000 runs. If we ignore this intervention event, then the disturbance term will show a shift during the window when it happens, then the d-ANOVA states $\mu + \tau_1$ will show a small shift as data after shift is included into the observation window. Later another large shift will occur when the last batch of data before the shift is excluded from the observation window. Similarly, the dynamic term η_1 will show corresponding shifts when it occurred, and regression window include data after shift and when data before shift move out of the regression window. If intervention analysis are included and the point of intervention identified, the data before and after shifts are regressed separately. As shown in Fig. 10, all shifts of the d-ANOVA states $\mu + \tau_1$ are found when data of the control window in which the shifts occur are included into the observation windows. It can be observed in Fig. 11 that no shifts were observed when the data before shift move out of the regression window when intervention analysis is implemented. The dynamic term η_1 will show shifts only when it occurred, and regression window include data after shift. There is no obvious difference between the controller performances whether intervention event analysis is included or not. However the use of intervention event analysis alerts operator to events of tool shifts near the time point only when it actually happens. Without intervention event analysis, the operator may misinterpret shifts of d-ANOVA states due to exclusion of a previous event from the regression window.

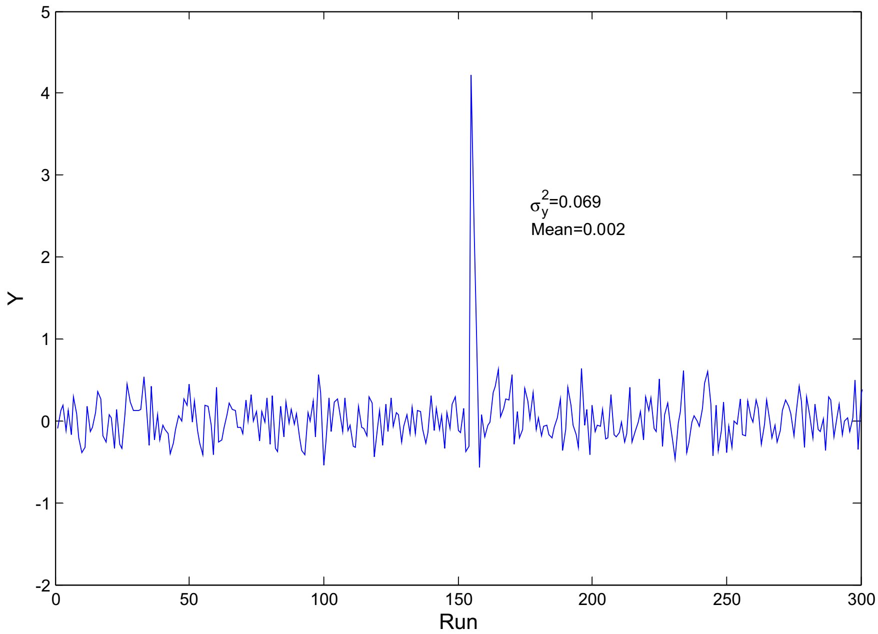

Fig. 6. Response under the control of d-ANOVA algorithm.

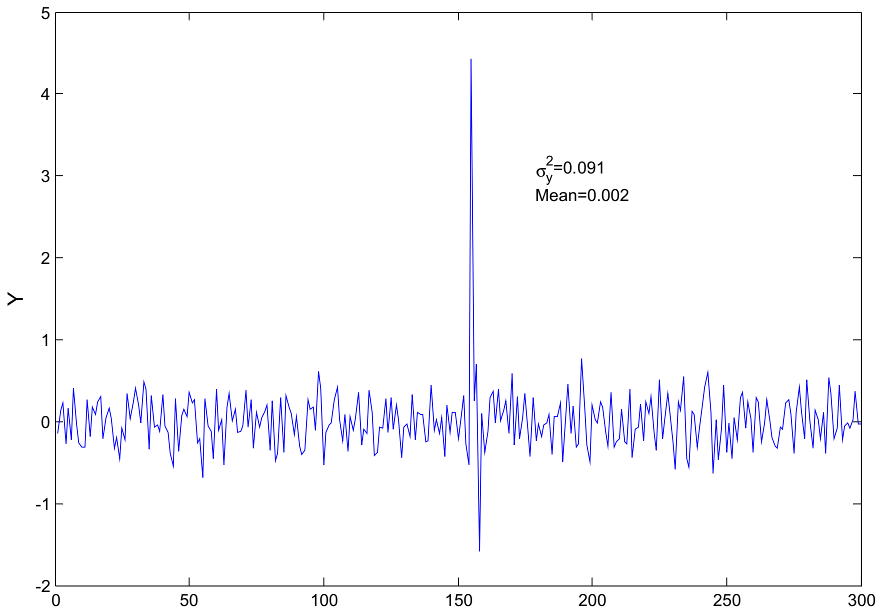

Fig. 7. Response under the control of s-ANOVA algorithm.

3.3. Product carry over

As stated above, some products may disappear for a period of time and come back on-line. When the off-line time is longer than the length of regression window, the corresponding parameter is not estimated in the last regression window as it comes back again. In this case, the product effect should be estimated by the method described in Section 2.5. In this example, product C was off-line after 100 runs and did not appear until 801st run. The relative ANOVA states of product A, B, and C estimated after the first regression window are −5, −1, and 6, respectively. Because of tool change, the ANOVA states of product A, B and overall mean estimated after the 7th regression window are −2.00, 2.00, respectively. When the product C came back at the 801st run, the product effect can be estimated by Eq. (15) a value of $-2+6-(-5)=9$ or $2+6-(-1)=9$ in the first control window. At the 900th run, a new set of product parameters are estimated. Note that s-ANOVA is not included in this particular case because the observation matrix will become ill-condition in case of some products are off-line (as shown in our previous work [10]).

The responses under the control of d-ANOVA and t-EWMA algorithm are shown in Figs. 12 and 13. The variance of the output under the control of d-ANOVA approach is 0.054. The variance of the output under the control of threaded EWMA approach is 0.091. It can be seen that for the t-EWMA control there are two prominent spikes when the product C come back on-line. It should be pointed out that when the product comes back on-line after a long period of absence, the conditions of the tools have changed. Threaded

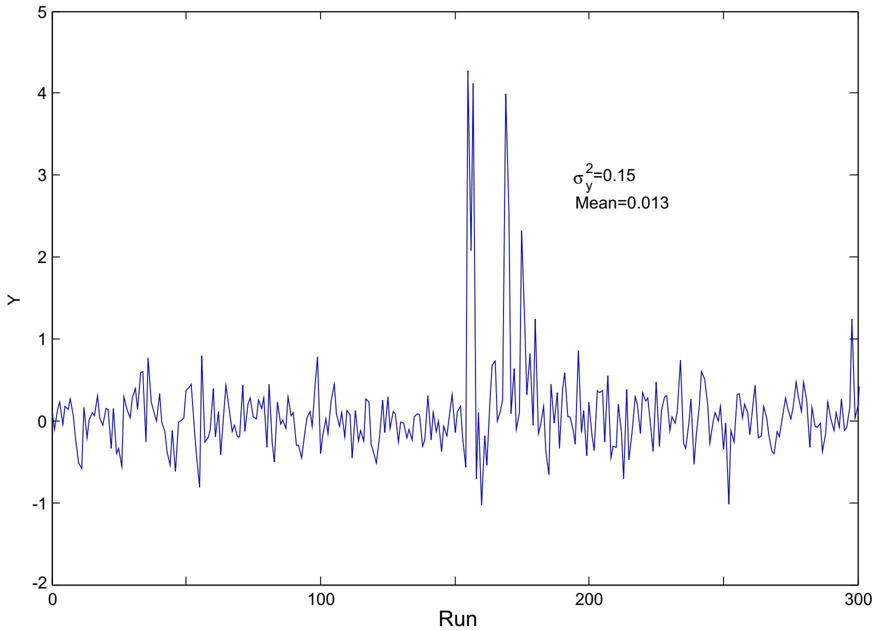

Fig. 8. Response under the control of t-EWMA algorithm.

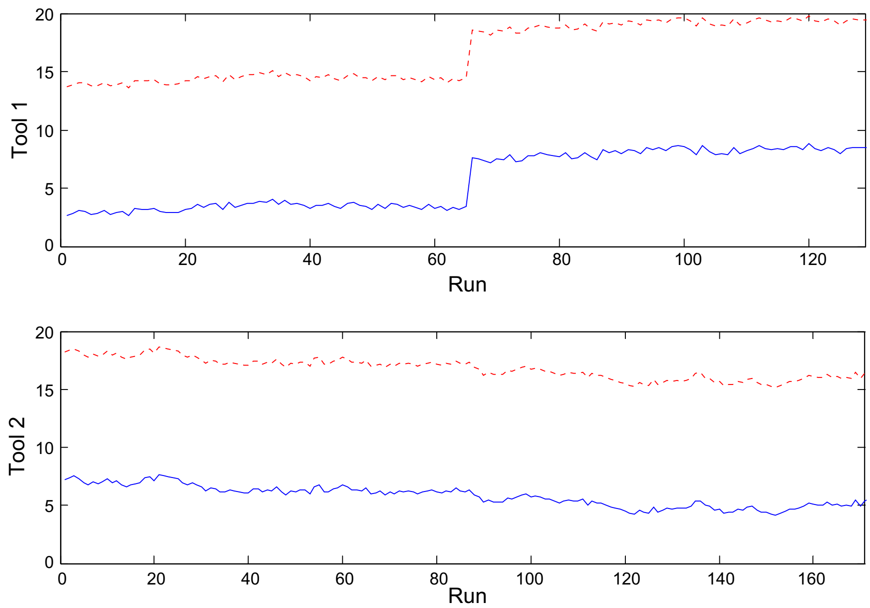

Fig. 9. The real and estimated tool effects (real —, estimated ---).

EWMA failed to account for this and must wait for new data with substantial offsets to come before controller took effect. The transition of d-ANOVA is much smoother. The performances of product C under the control of d-ANOVA and t-EWMA algorithm are 0.058 and 0.12, respectively. The real and estimated tool effect is shown in Figs. 14. It can be seen that there a change in value when the product C disappeared and come back from the regression window.

3.4. Performance comparison for products with different distribution levels

In this section, the performance of the three control algorithms, d-ANOVA algorithm, s-ANOVA method and t-EWMA algorithm for the "infrequently" products is investigated. The same simulation example of Section 3.1 is used, but the probability distributions of products A, B and C are adjusted to 0.6, 0.35 and 0.05, respectively. The results are shown in Table 1. Mean squared deviation (MSD) between the process output and the target is used to evaluate the performance of the three control algorithms. We can see that the d-ANOVA has the smallest MSD in all cases and there are no obvious differences among large, medium and small quantity products. The t-EWMA algorithm has comparable performance with the d-ANOVA algorithm for the products with large quantity (product A). However, it has poor performance for small quantity product (product C).

To see the effect of percentage of product distribution on the performance of products, a simulation example is implemented in which the percentage of C ranges from 1% to 30% and percentage

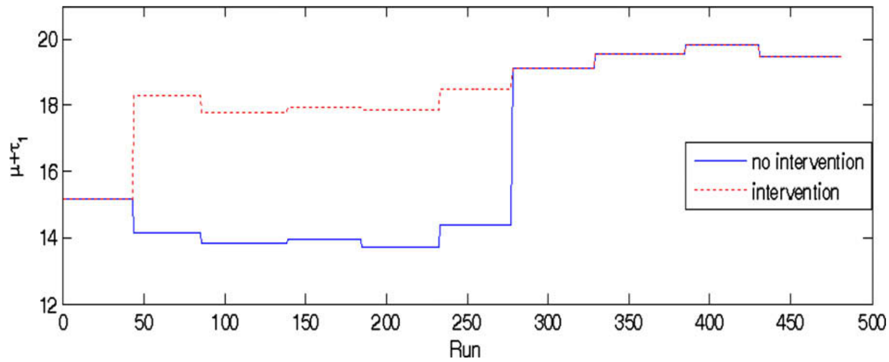

Fig. 10. ANOVA tool states estimated with and without intervention event analysis.

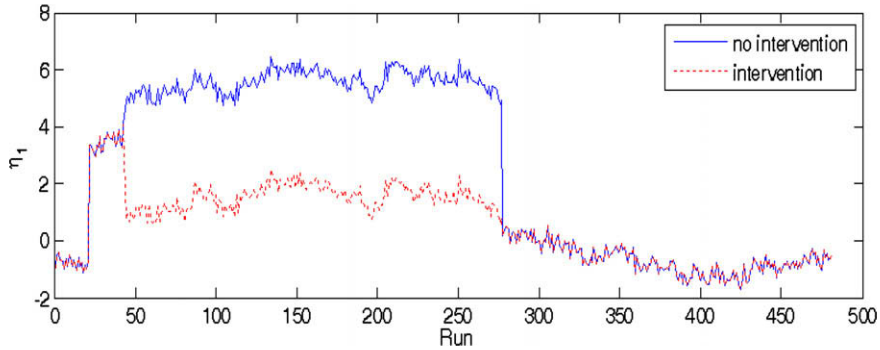

Fig. 11. Dynamic term estimated with and without intervention event analysis.

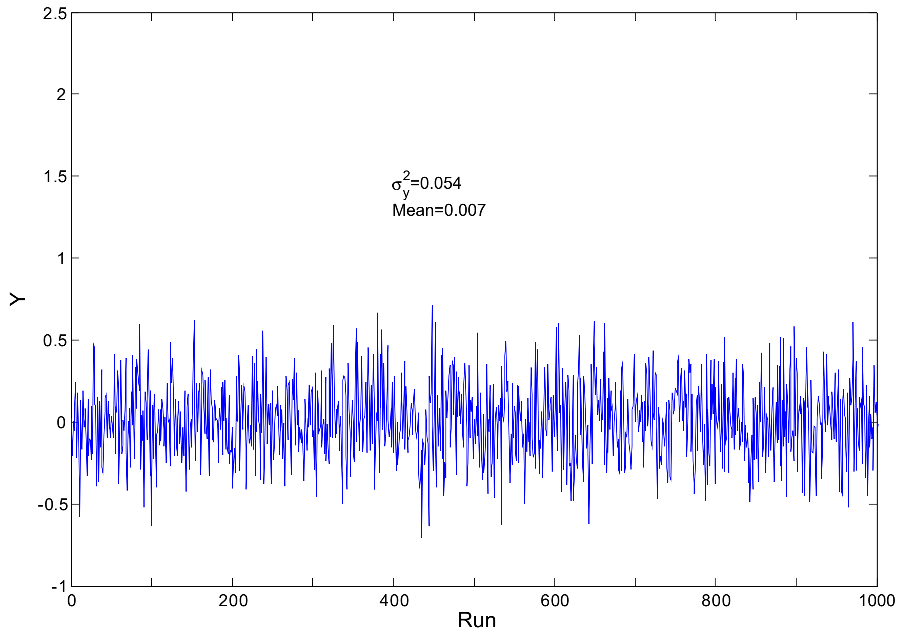

Fig. 12. Response under the control of d-ANOVA in the case of product coming off-line and back on-line.

of A changes from 64% to 35%. When percentage of C is small, simulation for each case was performed multiple times and an average value was taken. Fig. 15 shows the changes of MSD value with the percentage of C for the three control algorithm discussed in this paper. It is obvious that the performance of t-EWMA deteriorates substantially as the product decreases in quantity. However the mean square deviations remain relatively constant for the two ANOVA based methods.

3.5. Model structure error

In above sections, it is shown that good control performance can be obtained when the model is identified correctly. However, in real practice, the dynamics of the disturbance is very complex and the model structure adopted by the algorithm is simple for the easy implementation in most cases. Therefore, it is of interest

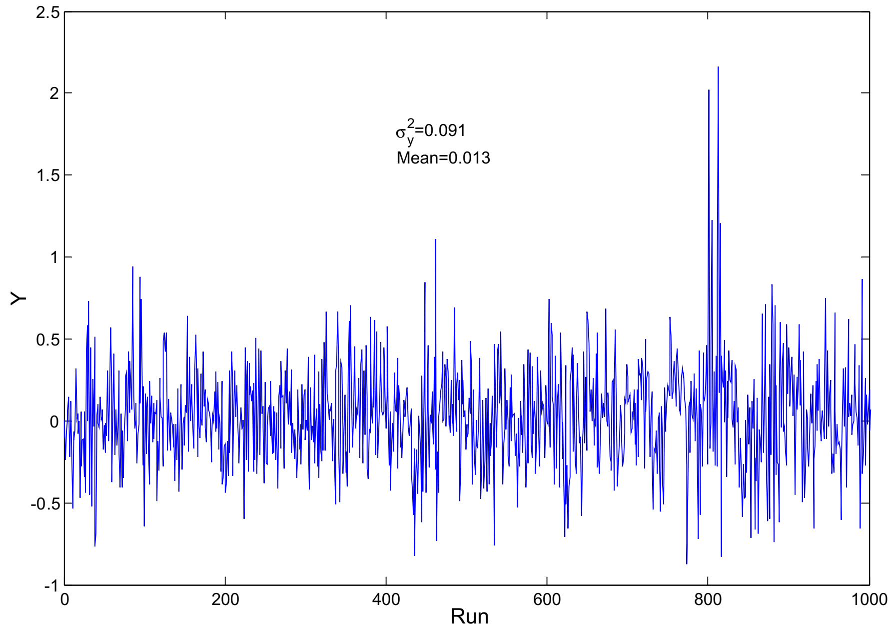

Fig. 13. Response under the control of t-EWMA algorithm in the case of product coming off-line and back on-line.

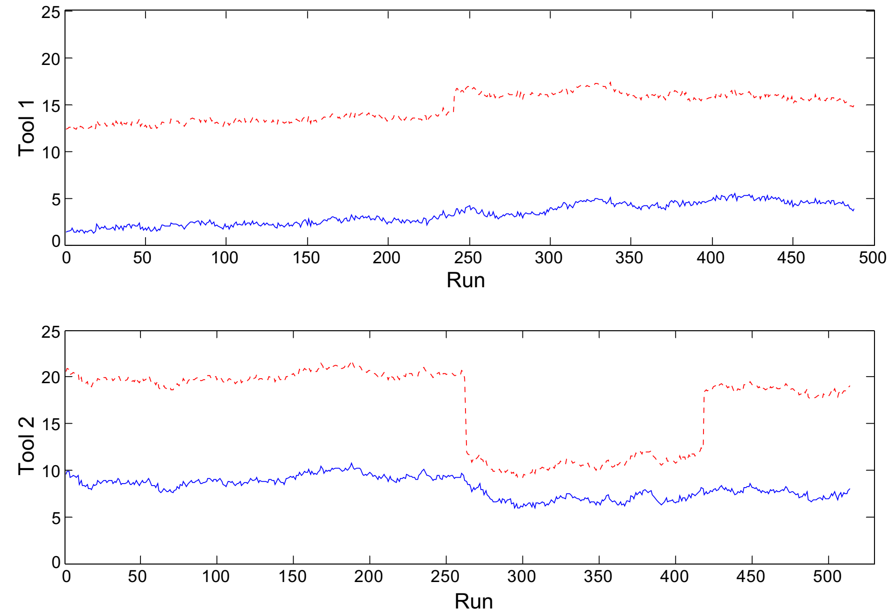

Fig. 14. Real and estimated tool effects (real —, estimated ---).

Table 1
Performance for products with large, medium and small quantity

Method	Measure			
	MSD_A	MSD_B	MSD_C	MSD
d-ANOVA	0.055	0.057	0.055	0.055
s-ANOVA	0.068	0.078	0.084	0.073
t-EWMA	0.072	0.099	0.364	0.098

to study how the control performance would be when there is model structure error exists.

Consider an example in which the two tool disturbances are represented by ARIMA(1, 1, 3) processes

$$(1 - 0.2B)(1 - B)\eta_1(k_1) = (1 - 0.1B - 0.41B^2 + 0.105B^3)\varepsilon_1(k_1)$$
$$(1 - 0.1B)(1 - B)\eta_2(k_2) = (1 - 0.1B - 0.22B^2 + 0.04B^3)\varepsilon_2(k_2)$$

The tool disturbance models adopted by the algorithm, $\hat{\eta}_i(k_i)$ $(i = 1, 2)$, are assumed to be IMA(1, 1) processes. The control performances of the three methods: EWMA, s-ANOVA and d-ANOVA at different levels of noises are shown in Table 2. It is found that the d-ANOVA approach has consistently the best performance among the three methods. In actual practice, various identification techniques can be used to ensure that the model structure assumed is a reasonable one. In this work, we merely wish to show that including a dynamic term can lead to improved performance.

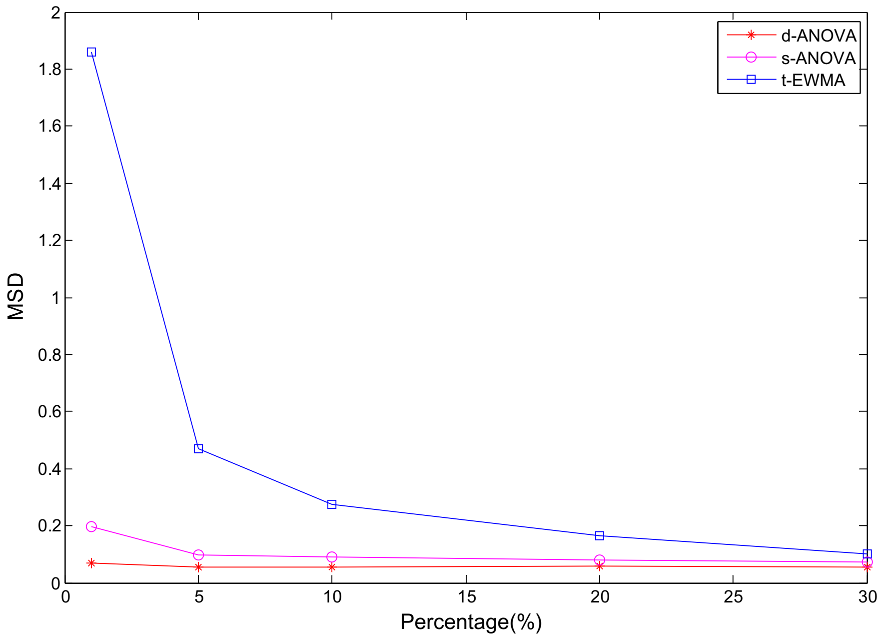

Fig. 15. Changes in mean square deviation versus percentage of product distribution.

Table 2
Performances for three methods under different noise level

Noise variance of input disturbance	Output variance		
	t-EWMA	s-ANOVA	d-ANOVA
$\sigma_\varepsilon^2 = 0.04$	0.118	0.078	0.061
$\sigma_\varepsilon^2 = 0.16$	0.465	0.313	0.214
$\sigma_\varepsilon^2 = 0.64$	1.75	1.05	0.72

4. Industrial example

In this section, wafer etching production data is used to test the effectiveness of the proposed algorithm. The collected wafer etching production data was originally under the control of s-ANOVA method. For such a process, it is known that aging effects such as the depletion of the etch solution or the degradation of the thermocouples in high temperature furnaces can induce trend or ramp disturbances. We use an IMA(1,1) process, the dynamic term, to characterize the disturbance. 2 tools, 6 product and 200 runs are picked up to implement the proposed algorithm. The data was standardized to have overall mean zero and overall variance 1. The controlled output is plotted in Fig. 16.

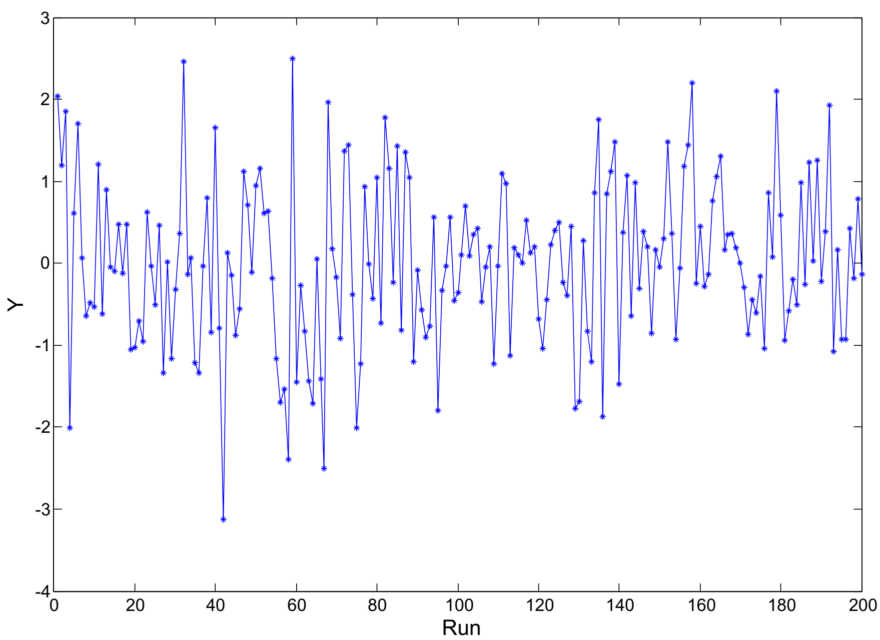

Fig. 16. Normalized production data.

Table 3
Performances of different products under the control of d-ANOVA, s-ANOVA and t-EWMA algorithms

Product(counts)	d-ANOVA		s-ANOVA		t-EWMA	
	STD	Mean	STD	Mean	STD	Mean
P1(5)	0.5	0.17	0.80	0.48	1.07	0.09
P2(4)	0.73	−0.05	1.20	0.17	1.15	−0.75
P3(35)	0.87	−0.03	1.03	−0.01	1.00	−0.06
P4(28)	0.96	0.19	0.88	0.12	0.93	−0.16
P5(51)	0.78	−0.24	0.89	0.05	0.86	−0.08
P6(77)	1.05	0.09	1.04	0.18	0.97	−0.06

The control results of dynamic ANOVA algorithm, s-ANOVA method and threaded EWMA method for different products are summarized in Table 3. The overall variances of the output under the control of d-ANOVA, and t-EWMA algorithm are 0.93 and 1.04, respectively. It should be noted that the three algorithms have almost equivalent performance for products with large lot counts, P3, P4, P5 and P6. t-EWMA algorithm guarantees asymptotic zero mean offset for individual products. Hence for products of large quantity, they have near zero offsets. ANOVA models results in overall zero offsets. Hence for large volume products only the overall mean of all products are maintained at zero offset. t-EWMA method has poorer performance for infrequent products, P1, P2. This is due to the fact that the tool condition may be quite different from the last run of the same thread. d-ANOVA is better than s-ANOVA, no intervention events have been detected. Hence inclusion of dynamic term to the ANOVA model is the main reason for better performance.

5. Conclusions

It is very important in RtR control of a mixed run plant to correctly identify the changes in condition of tool as well as the difference in behavior between tools and products. In this work, a novel mixed product run-to-run controller is proposed. The method of ANOVA is used to estimate the difference in behavior between tools and products and a dynamic term is included in the process model to characterize the run-to-run disturbance such as drift, shift and/or some other unknown disturbances. The effectiveness of the proposed algorithm is illustrated by different simulation examples in which scenarios tools drift and shift, product going off-line and coming back on-line are examined. The effect of product distributions on controller performances is also investigated. It is found that the performance of ANOVA based controller for products with small quantity is comparable to those with large quantity. The advantage of the d-ANOVA algorithm is also demonstrated using an industrial example.

References

[1] E.D. Castillo, A.M. Hurwitz, Run-to-run process control: literature review and extensions, Journal of Quality Technology 29 (1997) 184–196.
[2] J. Moyne, Run-to-Run Control in Semiconductor Manufacturing, CRC Press, Florida, 2001.
[3] E.D. Castillo, Statistical Process Adjustment for Quality Control, John Wiley and Sons, New York, 2002.
[4] A. Toprac, Y. Wang, Solving the High-mix Control Problem, AEC/APC Asia, Hsinchu, Taiwan, 2004.
[5] A.J. Pasadyn, T.F. Edgar, Observability and state estimation for multiple product control in semiconductor manufacturing, IEEE Transactions on Semiconductor Manufacturing 18 (2005) 592–604.
[6] S.K. Firth, W.J. Campbell, A. Toprac, T.F. Edgar, Just-in-time adaptive disturbance estimation for run-to-run control of semiconductor processes, IEEE Transactions on Semiconductor Manufacturing 19 (2006) 298–315.
[7] Y. Zheng, Q.H. Lin, D.S.H. Wang, S.S. Jang, K. Hui, Stability and performance analysis of mixed product run-to-run control, Journal of Process Control 16 (2006) 431–443.
[8] C.A. Bode, J. Wang, Q.P. He, T.F. Edgar, Run-to-run control and state estimation in high-mix semiconductor manufacturing, Annual Reviews in Control 31 (2007) 241–253.
[9] D.C. Montgomery, Design and Analysis of Experiments, fourth ed., John Wiley & Sons, New York, 1997.
[10] M.D. Ma, C.C. Chang, D.S.H. Wong, S.S. Jang, State estimation of a mixed run plant, Journal of Process Control (2008), doi:10.1016/j.jprocont.2008.07.009.
[11] G.E.P. Box, G.M. Jenkins, G. Reinsel, Time Series Analysis Forecasting and Control, third ed., Prentice-Hall, 1994.