A NOVEL MIXED PRODUCT RUN-TO-RUN CONTROL ALGORITHM – DYNAMIC ANCOVA APPROACH

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 150001, China*
[b]*Department of Chemical Engineering, National Tsing-Hua University, Hsin-Chu 30043, Taiwan*

Abstract

In semi-conductor manufacturing industry, production resembles an automated assembly line in which many similar products with slightly different specifications are manufactured step-by-step, with each step being a complicated physiochemical batch process performed by a number of tools. This constitutes a high-mix production system for which effective run-to-run control (RtR) and fault detection control (FDC) can be carried out only if the states of different tool and different products can be estimated. However, since in each production run, a specific product is performed on a specific tool, absolute individual states of products and tools are not observable. In this work, a novel state estimation method based on analysis of variance (ANOVA) is developed to estimate the relative states of each product and tool to the grand average performance of this station in the fab. The method is formulated in a form of recursive state estimation using Kalman filter. The advantages of this method are demonstrated using simulations to show that the correct relative states can be estimated in production scenarios such as tool-shift, tool-drift, product ramp-up, tool-offline and preventive maintenance. Furthermore, application of this state estimation method in a minimum variance based RtR control scheme shows that substantial improvements in process capabilities can be gained, especially for products with small lot counts.

Keywords: state estimation, ANOVA, Kalman filter, run-to-run control

1. Introduction

The semi-conductor manufacturing industry is one of the fastest evolving industries in the world. As feature sizes shrink and wafer sizes increase, sophisticated control methods are needed to improve product yield, throughput, and overall equipment effectiveness.

The run-to-run (RtR) controller is a model-based process control system that integrates concepts in statistical quality control (SPC) and engineering process control (EPC). It is achieved by adjusting process inputs (recipes) at the beginning of each run based on information obtained from previous runs. In the last decade, RtR control has been extensively deployed in the semi-conductor industry. Research and development in this area have been summarized by many authors (Moyne, 2001; Castillo, 2002).

Most of RtR control algorithms are based on the assumption that there is only a single product fabricated in the manufacturing line. This is, however, far from reality. In semi-conductor manufacturing industry, production resembles an automated assembly line in which many similar products with different specifications are manufactured step-by-step, with each step being a complicated physiochemical batch process carried out by a number of tools. A specific combination of product and tool is known as a "thread" (Firth, et al., 2006). Single product RtR control algorithms can be applied to a thread. However, the number of threads can be very large, upto thousands in a foundry fab. It is cumbersome to maintain so many controllers. Moreover, controller performance will be degraded for those infrequent threads since condition of the tool may be quite different from the last run of the same thread. Zheng et al. (2006) showed that even if the actual root cause is the change in condition of the tool, a single tool-based EWMA controller is unstable if the model uncertainties of different products are different. Firth et al. proposed a least square method known as just-in-time adaptive disturbance estimation (JADE) which includes additional constraints that the product and tool states remained unchanged from run-to-run and a proprietary weighting method.

In statistics, the problem of identification of different bias factors has been described as the analysis of variance (ANOVA) (Montgomery, 1997). In this work, a novel state estimation method based on ANOVA is developed to estimate the relative states of each product and the relative states of each tool to the grand average performance of this station in the fab. This method is formulated in a form of recursive state estimation using Kalman filter. The advantages of the proposed method are demonstrated using simulations to show that the correct ANOVA states can be estimated in production scenarios such as tool/product-shift, tool-drift, product ramp-up, product /tool-offline and preventive maintenance. Furthermore, application of this state estimation method in a minimum variance based RtR control scheme shows that substantial gains in process capability for specialized small lot counts products.

2. STATE ESTIMATION BASED ON ANOVA

2.1 Analysis of Variance (ANOVA)

According to ANOVA, the effects of different factors are expressed as:

$$y_k - bu_k = \mu + \tau_{n_k} + p_{m_k} + \varepsilon_k \qquad (1)$$

where μ is the overall mean of all observed tool and product combinations, τ_n ($n=1,\ldots,N$) represent the difference between the average results of all possible products on n^{th} tool and the overall mean, and p_m ($m=1,\ldots,M$) represent the difference between the average results on all possible tools of the m^{th} product and the overall mean. Unlike absolute states of the particular tool and product, τ_m and p_m are relative contributions subject to the constraints (Montgomery, 1997)

$$\sum_{n=1}^{N}\tau_n = 0 \qquad \sum_{m=1}^{M} p_m = 0 \qquad (2)$$

Here it is assumed that there exist no interactions between tools τ_n (n=1, 2,…,N) and parts p_m (m=1, 2,…,M).

If we assumed that the ANOVA states are stationary over several periods of time, then ANOVA model of the multi-tool and multi-product plant can be expressed in the following state space form:

$$\begin{aligned}[\alpha_{t+1}] &= [T][\alpha_t]+[\omega_t] \\ \left[\hat{Y}'_t\right] &= [Z'][\alpha_t]+[\upsilon_t]\end{aligned} \qquad (3)$$

where [ω_t] and [υ_t] are independent, zero-mean, Gaussian noise processes of covariance matrices [Q] and [R], respectively. [T] is the transition matrix.

The observability matrix for the above ANOVA state space model is

$$[O]=\left([Z'],[Z'][T]\cdots[Z'][T]^{N+M}\right)^T \qquad (4)$$

The system is observable if the observability matrix is full rank. In this case, the transition matrix [T] is an identity matrix, the observability matrix $[O]$ is equal to the output matrix $[Z']$ which is of full rank $N+M+1$.

2.2 Recursive Observation

Estimation can then be carried out in a recursive manner from intervals to intervals. At the start of any time interval t, given an estimated ANOVA state vector $[\tilde{\alpha}_{t-1}]$ and a estimated covariance matrix of the ANOVA states $\left[\tilde{P}_{t-1}\right]$, then the predicted values of the ANOVA state vector $\left[\tilde{\alpha}_{t\,|\,t-1}\right]$ and predicted the covariance matrix for this period $\left[\tilde{P}_{t|\,t-1}\right]$ are given by

$$\left[\tilde{\alpha}_{t|t\text{-}1}\right] = [T][\tilde{\alpha}_{t\text{-}1}]$$
$$\left[\tilde{P}_{t|\,t-1}\right] = [T]\left[\tilde{P}_{t\text{-}1}\right][T]^T + [Q] \qquad (5)$$

After the operating records $\left(y_{t,k}, u_{t,k}, n_{t,k}, m_{t,k}\right)$, $k = 1 \cdots K_t$ of this period are collected, the minimum mean square estimator of the ANOVA states and the covariance matrix can be updated by the following equations

$$[\hat{\alpha}_t] = \left[\tilde{\alpha}_{t|t\text{-}1}\right] + \left[\tilde{P}_{t|t\text{-}1}\right][Z'_t]^T [\Phi_t]^{-1}\left(\left[\hat{Y}'_t\right] - [Z'_t]\left[\tilde{\alpha}_{t|t\text{-}1}\right]\right)$$
$$\left[\tilde{P}_t\right] = \left[\tilde{P}_{t|t\text{-}1}\right] - \left[\tilde{P}_{t|t\text{-}1}\right][Z'_t]^T [\Phi_t]^{-1}[Z'_t]\left[\tilde{P}_{t|t\text{-}1}\right] \qquad (6)$$
$$[\Phi_t] = [R] + [Z'_t]\left[P_{t|\,t-1}\right][Z'_t]^T$$

It is not possible to guarantee that $[Z'_t]$ contains all the threads and of full rank $N{+}M{+}1$ during the data collection interval. For example, in the extreme case, the model can be updated whenever the result of a single run is reported. However it is important to ensure that $[Z']$ over an extended history is of full rank $N{+}M{+}1$. This assumption may be invalid if some products are terminated or a certain tool is offline for an extended period of time. The assurance of the above full rank assumption can be achieved by monitoring the condition number of the matrix $\left[\tilde{P}_t\right]$.

3. SIMULATION RESULTS

In this section a series of simulation tests are designed to investigate the effectiveness of the proposed algorithm in various operation scenarios. A simple two-tool-three-product example is used in the following simulation studies. For each operation scenarios, three different comparisons were made. First estimated ANOVA states $[\tilde{\alpha}]$ are compared with the actual ANOVA states $[\alpha]$. The actual biases of each tool a_n^{tool} and products a_m^{prod} are compared with $\mu + \tau_n$, the average performance of all products on the n^{th} tool, and $\mu + p_m$, the average performance of the mth products on all tools. Note that it is easy to identify the changes of tools or products using the ANOVA estimator proposed in this paper. If the ANOVA states of the tools or products remain stable, the corresponding conditions of the tools or products are confirmed to be unchanged. In case of a change of the ANOVA state are observed, the change of $\tilde{\mu} + \tilde{\tau}_n$ or $\tilde{\mu} + \tilde{p}_m$ confirms the change of the tool *n* or the product *m*. Furthermore, the changes of $\tilde{\mu} + \tilde{\tau}_n$ or $\tilde{\mu} + \tilde{p}_m$ are consistent with the actual states a_n^{tool} and a_m^{prod}.

3.1 Tool shift

There are many events which can result in an apparent immediate shift in the operating conditions. For instance, such a disturbance might occur when a tool undergoes a maintenance event. This event would be seen by the process as a step disturbance in the output variable. The disturbance is often not measurable. Therefore, the controller should learn from the process output and compensate for the effect of the disturbance.

The comparisons of actual and estimated ANOVA states, the actual and estimated states of each thread, and actual biases and relative bias of the each tool and products are shown in Figure 1. In this case, there is an abrupt change for the bias value of tool 1 at the 100th run. In Figure 1a, we observed that all the threads on tool 1 experience an abrupt change while the states of all the threads associated with tool 2 remain unchanged. In Figure 1b, we found that the ANOVA states $\tilde{\mu}$ and $\tilde{\tau}_1$ experience positive shifts while the ANOVA tool state $\tilde{\tau}_2$ experience a shift in the other direction. In Figure 1c, it is found that the average performance of all products on tool 1 $\tilde{\mu}+\tilde{\tau}_1$ experienced a shift; the average performance $\tilde{\mu}+\tilde{\tau}_2$ of all products on tool 2 remained unchanged while the average performance $\tilde{\mu}+\tilde{p}_{A,B,C}$ of all products on tool 2 experienced a shift too.

Figure 2 illustrated the results of JADE estimates. In the simulation, an identity weighting matrix is used for comparison with ANOVA estimates. It is interested to note that when a shift is induced to tool 1, shifts are also observed to other factors as shown in Figure 2b. The biases obtained are not true estimates of these factors. However, the estimated states of all the threads are correct as shown in Figure 2a since the method is a least square fit of all the threads. Hence use of the recombined thread states for controlling existing threads is not a problem. Application of these individual factor states to estimate new thread may lead to errors.

3.2 Tool drift & PM

If a manufacturing process is known to drift due to equipment aging, then a deterministic drift exists in the system. Aging can be found in wafer etching process and chemical mechanical polishing. A drift persisted for a long period would normally be followed by a maintenance event and corresponding process reset, resulting a saw-tooth pattern in an uncontrolled quality characteristic.

The simulation results are shown in Figure 3. The two tools experienced deterministic drifts of slopes 0.1 and 0.2 respectively. Tool 1 was reset at the 115th run and tool 2 was reset at the 200th run. As shown in Figures 3a and 3b, the states of each thread and the ANOVA states of each tool and product can be estimated correctly throughout the simulation. Furthermore, the ANOVA states of the products remained unchanged. Figure 3c illustrated that the average performances $\tilde{\mu}+\tilde{\tau}_{1,2}$ of tool 1 and tool 2 show patterns that are consistent with the actual states. The average performances $\tilde{\mu}+\tilde{p}_{A,B,C}$ of the products 1 to 3 show saw-tooth patterns that are consistent with the changes in average performance of the whole plant. Since there are no changes in the relative performance of different product, the ANOVA product states $\tilde{p}_{A,B,C}$ remained unchanged as shown in Figure 3b.

3.3 Controller performance

In a high mixed foundry, some products are fabricated infrequently in small quantity. However, it should be noted that products which are produced with large quantity are usually of marginal profits and the products which are produced infrequently are often high-value added and contribute a substantial portion of the profit. Therefore, it is highly desirable for a control algorithm to have comparable performance for products with different run counts.

In this section, the performance for "infrequent" products for the three control algorithm, threaded EWMA algorithm, JADE and the ANOVA method proposed in this paper are investigated. A deadbeat control for ANOVA method is used. Runs are evenly distributed between the two tools, and the probability distributions of products A, B, C are 60%, 35% and 5% respectively. Two IMA (1, 1) time series tool disturbances are injected into the system which variances are $\sigma^2=0.16$ respectively.

The simulation results are shown in table 1. For all three methods, product A, which is the most frequent product produced, had the best performance. Product C, which is the produced in small quantity, had the worst performance. However, the performance of threaded EWMA is extremely poor for product C, while performances of JADE and ANOVA for product C are still acceptable. The ANOVA result is the best among the three methods for all three products with different run counts.

Table 1: Comparison of RtR control performances for different products

Method	MSE_A	MSE_B	MSE_C	MSE
ANOVA	0.2400	0.2833	0.4124	0.2642
JADE	0.2557	0.3182	0.5423	0.2964
EWMA(Threaded)	0.2746	0.4103	2.2348	0.4152

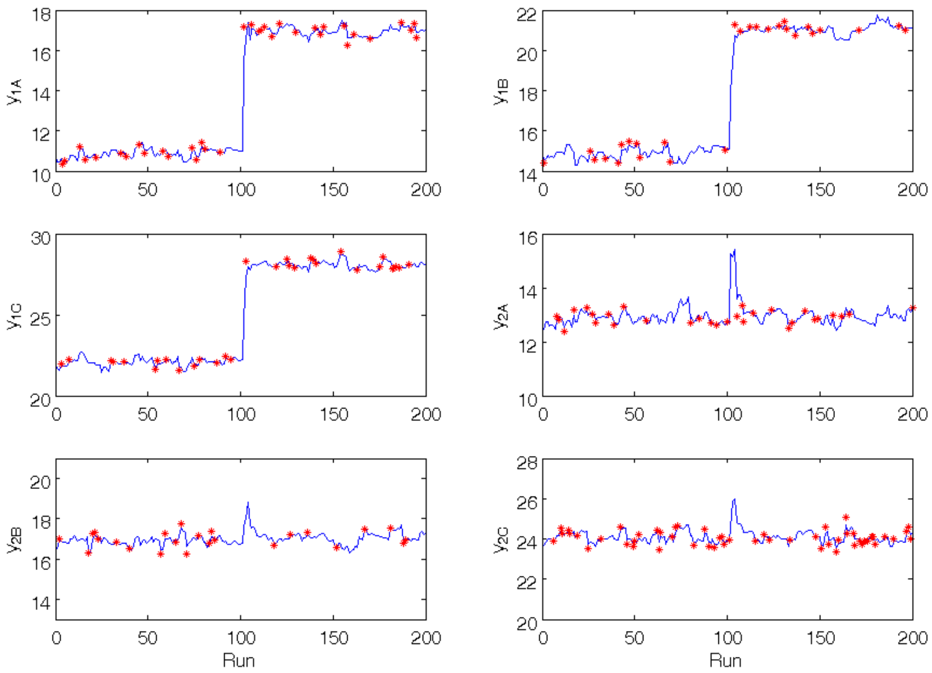

Figure 1a

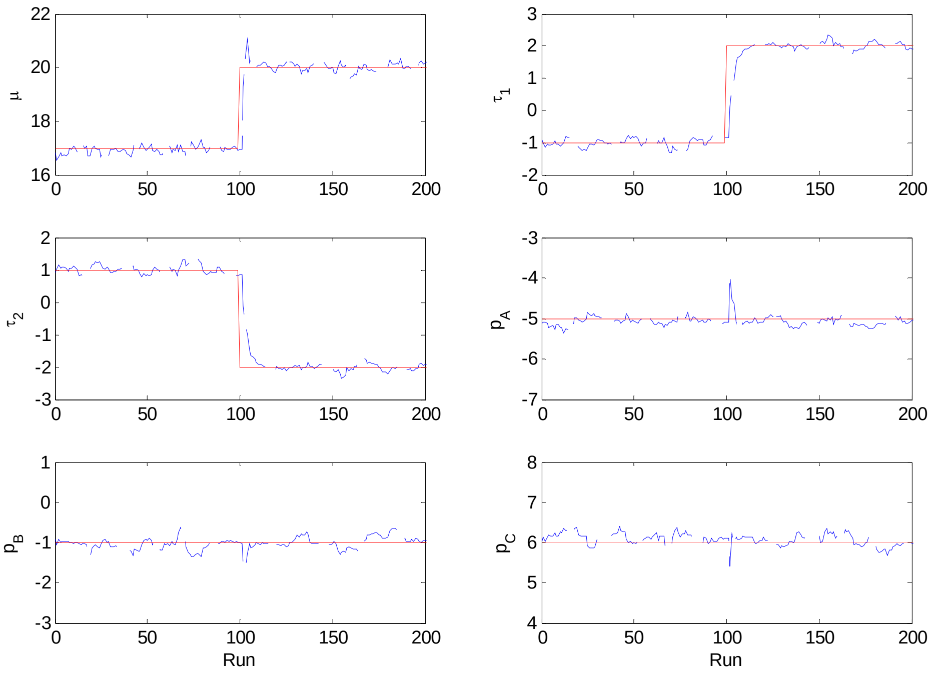

Figure 1b

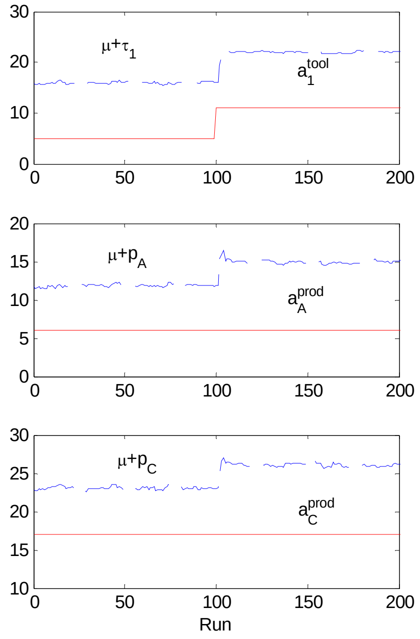

Figure 1c

Figure 1: Changes in actual and estimated (a) states of thread, (b) ANOVA parameters and (c) absolute and relative factor biases in a tool shift. (— actual values, --- estimated values, * observed values)

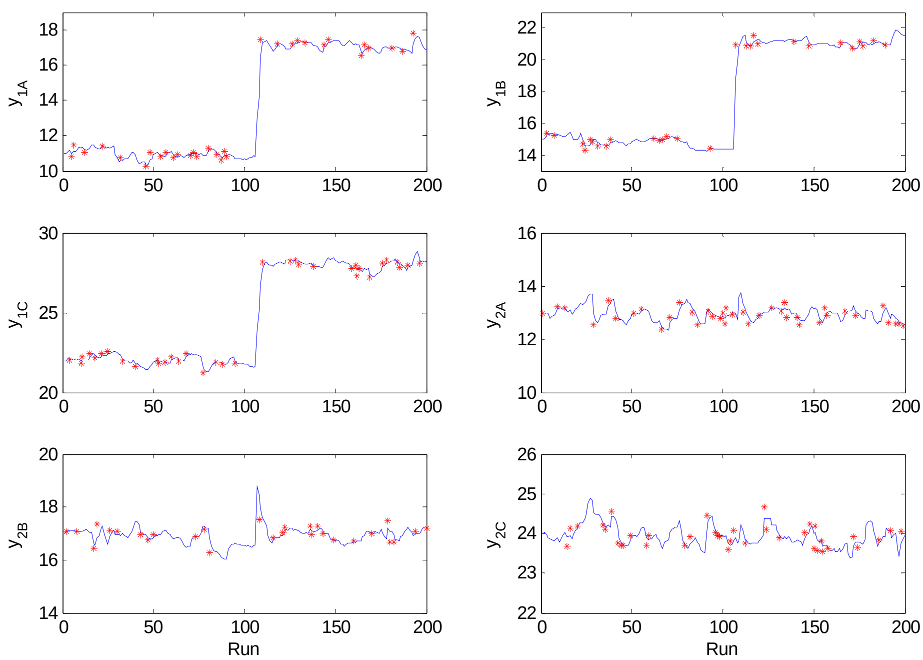

Figure 2a

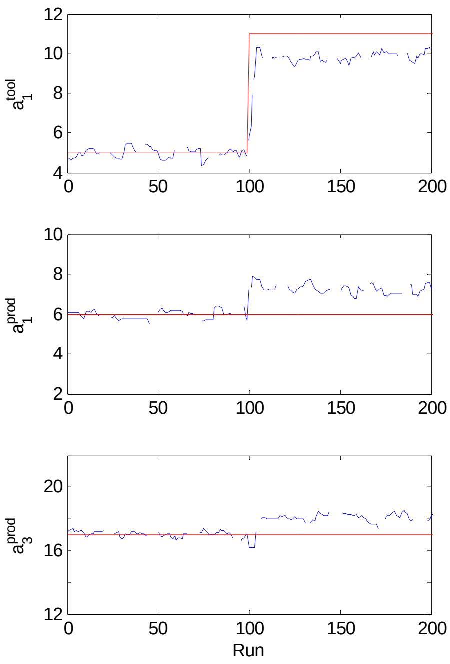

Figure 2b

Figure 2: Changes in actual and estimated (a) states of thread and (b) absolute factor biases using least square method in a tool shift. (— actual values, --- estimated values, * observed values)

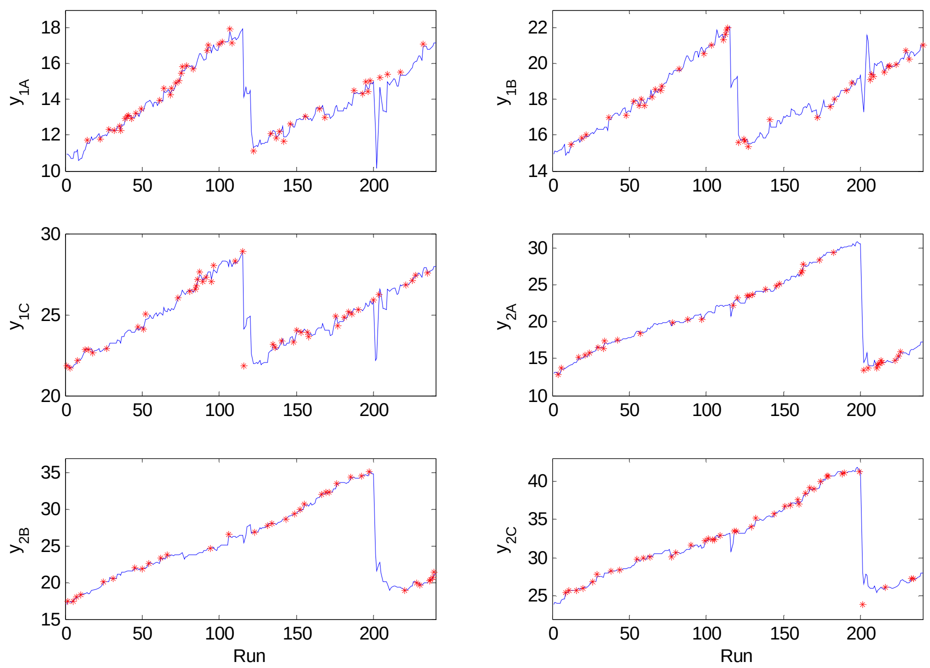

Figure 3a

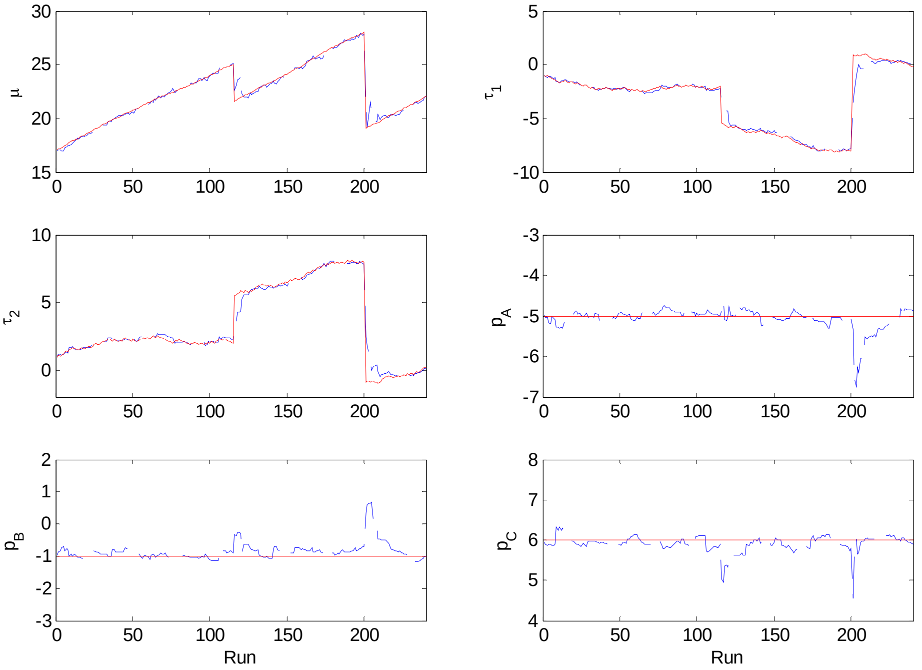

Figure 3b

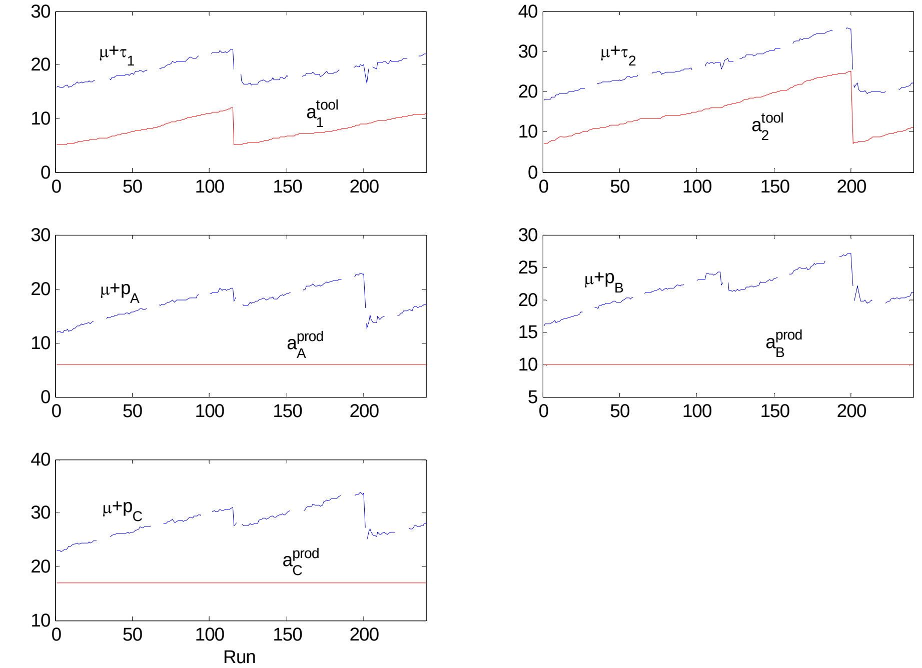

Figure 3c

Figure 3: Changes in actual and estimated (a) states of thread, (b) ANOVA parameters and (c) absolute and relative factor biases with tool drifts and preventive maintenances (— actual values, --- estimated values, * observed values)

4 CONCLUSIONS

In this paper a novel state estimation method based on statistics method ANOVA is developed to estimate the relative states of each product and the relative states of each tool to the grand average of this station in the fab. The method is formulated in a form of recursive state estimation using Kalman filter. Simulation results show that the correct ANOVA states can be estimated in production scenarios such as tool-shift, tool-drift, product ramp-up and offline. Furthermore, application of this state estimation method in a minimum variance based RtR control scheme shows that substantial improvement of quality of products with small run counts. This makes the proposed method highly suitable for mixed product control system.

References

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

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

Firth, S. K., Campbell, W. J., Toprac, A., & Edgar, T. F., (2006) *IEEE Transactions on Semiconductor Manufacturing*, 19, 298-315.

Zheng Y., Lin, Q. H., Wong, D. S. H., Jang, S. S., & Hui, K., (2006) *Journal of Process Control*, 16, 431-443.

Montgomery, D. C., *Design and Analysis of Experiments*, John Wiley & Sons, New York (1997).