EWMA/SD : AN END-OF-LINE SPC SCHEME TO MONITOR SEQUENCE-DISORDERED DATA

Chih-Min Fan*, Ruey-Shan Guo**, Shi-Chung Chang*
*Department of Electrical Engineering, National Taiwan University
**Department of Industrial and Business Administration, National Taiwan University
e-mail: min@ac.ee.ntu.edu.tw

Abstract

In this paper, we focus on the design issues of applying SPC control chart to end-of-line wafer acceptance test (WAT) data. Since the sequence of end-of-line data is not the same as the sequence in each process step, an abnormal trend in any of the process steps is more difficult to detect based on the end-of-line data than based on single process data (if available). To overcome this deficiency, we propose an exponentially weighted moving average method for sequence-disordered data (EWMA/SD). The basic idea is that moving average can smooth out the sequence-disordered effect and weighting factors allow us to choose an effective window size so that the underlying trend can be popped out. It is different from the traditional EWMA method as it has the capability to handle the sequence-disordered data. An end-of-line trend detection system has been developed for validation of the method, which consists of three modules: run length distribution generator, optimal parameter generator, and EWMA/SD control chart. Based on process characteristics, the corresponding run lengths of an EWMA/SD control chart for different parameter vectors are derived by the Markov chain approach. The optimal parameter vector is chosen as the one which meets the requirement for maximum false alarm rate and maximizes the detection speed at the same time. Results of simulation and field data validation show that EWMA/SD is able to smooth out the sequence-disordered data, be sensitive to process changes, and be robust to nature noise.

1. Introduction

Statistical process control (SPC) charts, such as Shewhart chart, exponentially weighted moving average (EWMA) chart and CUSUM chart, have been widely applied to semiconductor manufacturing to detect abnormal trends such as process drift and shift. In this paper, we focus on the special design issues when applying SPC techniques to monitor the end-of-line data. End-of-line data mostly refers to the WAT data taken after completing the whole fabrication process. It provides important integration status and facilitates early detection of abnormal trends to prevent series yield crisis.

However, control chart techniques cannot be applied to end-of-line data without caution. Since the sequence of end-of-line data is not the same as the sequences in each processing machine, our empirical data characterization indicates that an abnormal trend in any of the processing machines is more difficult to detect using end-of-line data. To overcome this deficiency, we propose an EWMA/SD method. It is different from the traditional EWMA method as it has the capability to smooth out sequence-disordered data and to choose an optimal parameter vector so that the underlying trend can be popped out.

2. Current Practice

A current fab practice of the end-of-line SPC adopts a heuristic, biweekly SPC review to monitor the "sequence-disordered" data. Under this method, the process capability metric Cpk [1] values of key end-of-line parameters are monitored every two weeks. If any one of the Cpk values is less than the corresponding threshold values specified by engineers, a problem might have occurred and the corresponding control charts must be reviewed to see if there is any significant trend. The philosophy of this method is to reduce the sequence-disordered effect by plotting the data in a two-week time period so that the trend pattern could be identified more easily. Several problems arise with the use of this method:

- the threshold value of each end-of-line item is empirical,
- a trend pattern is identified by engineers subjectively, and
- a two-week time period for trend detection is too long.

3. Sequence-disordered Effect

To overcome this deficiency, we first characterize the sequence-disordered effect by analyzing the end-of-line data generation process. Capturing the problem features, we then propose an EWMA/SD method to enhance the detection resolution and speed.

Let $\{X_i\}$ be a random sequence representing wafer lot averages of a measurement item taken at process step P, and $\{Y_k\}$ be a random sequence representing wafer lot averages of a physically related measurement item (to X) taken at the end-of-line. For example, $\{X_i\}$ may correspond to an in-line measurement such as dopant concentration and deposited layer thickness of lot i at step P, and $\{Y_k\}$

0-7803-3737-9/97/ $5.00©1997 IEEE

may correspond to an electrical test measurement such as a threshold voltage or a saturation current. Based on the physical laws, $\{Y_k\}$ can be expressed as a function of $\{X_i\}$.

In general, affected by different process flows and dispatching polices, the cycle time from process step P to the end-of-line step is not constant among lots. As a result, the lot with sequence label i at step P very likely has a different lot sequence label k at the end-of-line step. Define the sequence-disordered grade R of the lot from step P to the end-of-line as $k-i$. Then the range of R can characterize the sequence-disordered effect, which is noted as sequence-disordered range R^*. If we further define $\{Z_i\}$ as a sequence that reorders $\{Y_k\}$ according to lot sequence at step P, the end-of-line data generation process can be illustrated as that in figure 1.

To monitor the trend in process step P, we can exploit the sequences $\{X_i\}$, $\{Z_i\}$ or $\{Y_k\}$. Usually, the in-line sequence $\{X_i\}$ might be missing or not available. There is also a limitation in tracking back $\{Z_i\}$ from the historical data base because there are too many process steps to monitor. The required computer memory and processing time would be too much if we trace all process steps. Therefor, our approach is to detect the trend by exploiting end-of-line sequence $\{Y_k\}$ directly. If a trend is identified, we then trace the most possible step (fault) to diagnose the root cause (figure 2).

In order to demonstrate the sequence-disordered effect, a simulation is performed and results are shown in figure 3. As can be seen, the shift pattern is significant in the in-line sequence $\{Z_i\}$ but not so obvious in the end-of-line sequence $\{Y_k\}$. Theoretically, the sequence-disordered effect increases both the mean and variance of $\{Y_k\}$ during the transient phase, which results in a larger fluctuation in the transient phase as compared to that of a non-disordered sequence $\{Z_i\}$. As a result , the trend pattern is destroyed and the detection becomes more difficult. Our approach is to smooth out the sequence-disordered fluctuation and to estimate the process mean based on the exponentially weighted moving average of $\{Y_k\}$ as described in the next section.

4. EWMA/SD

4.1 Moving Average

The fluctuation in the transient phase caused by sequence-disordered effect can be viewed as a type of noise (figure 3). Usually, the moving average (MA) is adopted to smooth out the noise of observation data and estimate the process mean [2]. The MA with size m of a sequence $\{Y_k\}$ is defined by

$$B_k = \frac{1}{m}\sum_{n=0}^{m-1} Y_{k-n} \,. \qquad (1)$$

The larger the variance of noise, the larger the MA size is needed to smooth it out.

To apply MA technique to end-of-line data, we have to study the correlation among MA size m, sequence-disordered range R^* and the trend detection capability. If R^* is large, the interval length of transient phase is increased. Then a larger m is needed to smooth out the sequence-disordered fluctuation. However, the detection speed will be slow because a large window size is used. On the other hand, if m is too small, the MA will be too sensitive to the last few data points, which not only results in too many false alarms when process is under control but also fails to smooth out the sequence-disordered noise when a trend occurs. Therefor, the choice of MA size should be carefully studied to optimize its trend detection capability.

4.2 EWMA

EWMA is another approach preferred for trend detection. Because it can be implemented more efficiently and has the capability to resemble various MA with different m by changing the weighting factor flexibly [3][4].

Assume that the end-of-line sequence $\{Y_k\}$ defined in Section 3 follows a normal distribution with mean μ_0 and variance σ^2. The successive values generated by applying EWMA to $\{Y_k\}$ are:

$$\begin{aligned} A_k &= \lambda Y_k + (1-\lambda)A_{k-1} \\ &= \lambda Y_k + \lambda(1-\lambda)Y_{k-1} + \ldots + \lambda(1-\lambda)^{k-1}Y_1 \,, \end{aligned} \qquad (2)$$

where $A_0 = \mu_0$ and $0 < \lambda \le 1$. The control limits of $\{A_k\}$ are derived as

$$\mu_0 \pm h\sqrt{\lambda / (2-\lambda)}\sigma \,, \qquad (3)$$

where h is an adjustable control limit gain usually specified as 3 for a compromise between high detection speed and low false alarm rate.

When EWMA is applied to trend detection of sequence-disordered data, we have to choose an appropriate window size to smooth out the sequence-disordered data. As shown in figure 4, we can see that the effective window size is determined by the weighting factor λ. Therefor, the objective of an EWMA/SD method is to choose an optimal parameter vector (λ , h) so that it can handle the sequence-disordered data and maximize the trend detection capability.

4.3 Application

To evaluate the performance of EWMA/SD method, the concept of average run length is used, where run length is a random variable characterizing the number of observations from the time a level shift occurs until the shift is clearly indicated by a control chart [5]. The appropriate EWMA/SD parameter vector (λ , h) is the one which results in a large average run length, ARL_1, when the process is under control so that the false alarm rate can be reduced and a

short average run length, $ARL2$, when a shift has occurred so that the trend detection can speed up.

In order to overcome current practices' deficiency as mentioned in section 3, an EWMA/SD trend detection system has been developed. As shown in figure 5, it consists of three modules: run length distribution generator, optimal parameter generator, and EWMA/SD control chart. Inputs to this system include end-of-line data and process characteristics. Here, process characteristics are defined as follows:

- M_p : total number of machines in step p
- R_p^* : sequence-disordered range from step p to the end-of-line step.
- δ : potential magnitude of the shift

Based on these process characteristics, many pairs of $ARL1$ and $ARL2$ corresponding to different parameter vectors (λ, h) are derived in the run length distribution generator [4], which are used as the inputs to the optimal parameter generator. The optimal parameter vector is generated by two steps:

1. Find all the feasible parameter vectors that can meet the requirement of minimum $ARL1$ (maximum false alarm rate).
2. Choose the one in the feasible parameter vectors that can minimize $ARL2$ (maximize the detection speed) as the optimal parameter vector.

Once the optimum parameter vector is generated, it is used for the EWMA/SD control chart to perform trend detection. If any EWMA/SD value is out of control limits, an out-of-control signal is generated and integration engineers must find out the root causes as quickly as possible to prevent further yield crisis.

5. Data Validation

The simulation data generated in Section 3 is used for evaluating the effectiveness of EWMA/SD. Process characteristics in this case are $M_p = 1$, $R_p^* = 50$, $\delta = 1\sigma$, and the minimum $ARL1$ requirement is set as 300, which corresponds to an optimal vector $(\lambda, h) = (0.1, 2.75)$. As shown in figure 3, EWMA/SD lines not only show more significant trend patterns than the original end-of-line data but also avoid the false alarm by the traditional Shewhart SPC at the 63th data point. We can see that the use of optimal weighting factor ($\lambda = 0.1$) results in a faster detection speed than a lower value ($\lambda = 0.01$). However, if λ becomes close to one, EWMA/SD values will be the same as the original end-of-line data which resembles the Shewhart chart ($(\lambda, h) = (1,3)$) and shows no abnormal trend.

Next, real fab data is used for method validation. In this case, there is a shift in the standard deviation of end-of-line measurement Vt_P3. The nature log of standard deviation is plotted in figure 6, which seems to drift slowly over time due to the sequence-disordered effect. If we trace the end-of-line sequence back to the in-line step (a well-implant step, $M_p = 3, R_p^* = 75$), we can see a significant shift pattern in one of the three machines. The amount of shift is determined as $1\sigma_Y < \delta < 2\sigma_Y$. Then the optimal EWMA/SD parameter vector is derived as $(\lambda, h) = (0.08, 2.75)$, which results in an earliest detection of all. When the weighting factor is too small (e.g., $\lambda = 0.01$), the EWMA/SD cannot detect any abnormal trend.

6. Conclusions

In this paper, we focus on the trend detection of sequence-disordered end-of-line data. An EWMA/SD method is proposed and analyzed. It can generate the optimal parameter vector based on the process characteristics. The results of simulation and fab data validation show that EWMA/SD method smoothes out the sequence-disordered data, is sensitive to process change, and is robust to nature noise.

Acknowledgment

This work was supported in part by the National Science Council of the Republic of China under Grant NSC86-2416-H-002-012 and Taiwan Semiconductor manufacturing Corporation under Contract 84-S-119 and 85-S-047.

Reference

[1] R.E. Devor, T. Chang and J.W. Sutherland, *Statistical Quality Design and Control: Contemporary Concepts and Methods,* New York : Macmillan, 1992.

[2] P.J. Brockwell and Richard A. Davis, *Time Series: Theory and Methods*, New York: Springer-Verlag, 1991.

[3] J.S. Hunter, "The Exponentially Weighted Moving Average," *Journal of Quality Technology*, vol. 18, no. 4, 1986, pp. 203-210.

[4] J.M. Fan, R.S. Guo, S.C. Chang and J.H. Lee, "Abnormal Trend Detection of Sequence-disordered Data Using EWMA Method," *Proceedings of 1996 Advanced Semiconductor Manufacturing Conference and Workshop*, 1996, pp. 169-174.

[5] J.M. Lucus and M.S. Saccucci, "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements," *Technometrics*, vol. 32, no. 1, 1990, pp. 1-12.

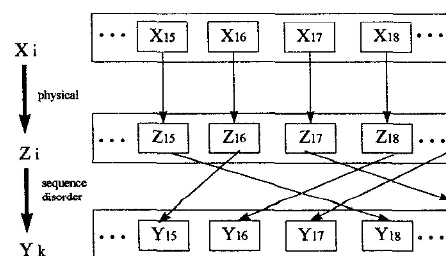

Figure 1 End-of-line data generation process

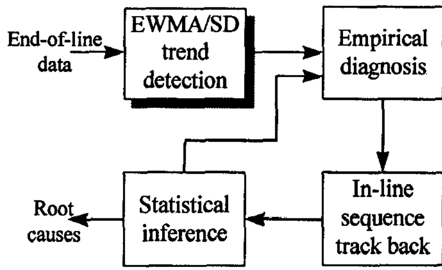

Figure 2 A framework for end-of-line trend detection and diagnosis

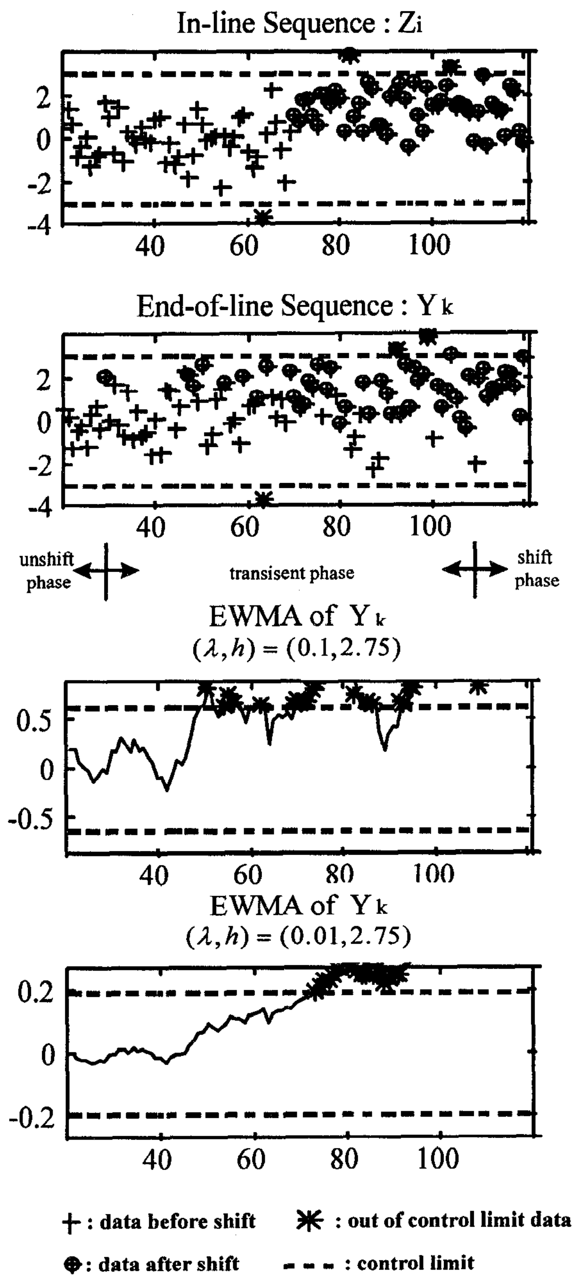

Figure 3 A simulation to demonstrate the sequence-disordered effect and the EWMA/SD trend detection capability

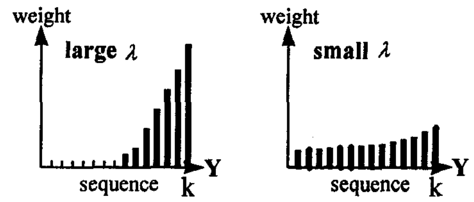

Figure 4 The decaying behavior of EWMA weighting coefficient

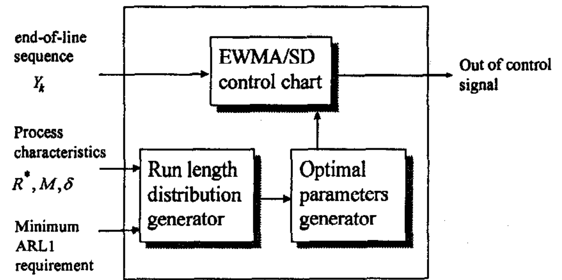

Figure 5 An EWMA/SD trend detection system

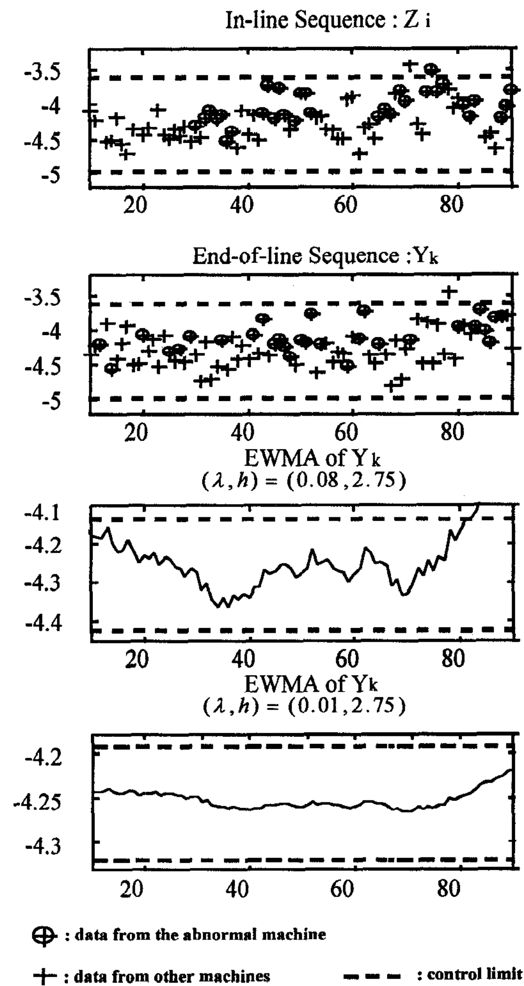

Figure 6 Fab data validation for the EWMA/SD trend detection capability