

## Mean Estimate for Shewhart-Chart-Monitored Processes Subject to Random Shifts

Argon Chen

Graduate Institute of Industrial Engineering  
National Taiwan University  
1, Sec. 4, Roosevelt Rd., Taipei, Taiwan, 106

E. A. Elsayed

Department of Industrial Engineering  
Rutgers, The State University of New Jersey  
96 Frelinghuysen Rd., Piscataway, NJ 08854, USA

### ABSTRACT

The Shewhart control chart is a widely used statistical control tool that helps detect a possible process shift (out-of-control process). Though the Shewhart chart is in theory a monitoring tool that reveals only the result of a hypothesis testing, in practice the chart's signaling is often used as the evidence for making process adjustment. In this paper, we develop a process mean estimator for processes monitored by Shewhart charts. This mean estimate can serve as an important reference for investigating assignable causes and taking appropriate corrective actions. A semiconductor fabrication process will be used as an example to illustrate the methodology.

### INTRODUCTION

In practice, the Shewhart control chart remains the most used, if not the most effective, statistical process control tool. Though the statistical theories and assumptions behind the chart have been fully discussed, it is the employment of the chart that draws continuous debates and discussions. Different types of industries have different needs in monitoring or controlling production processes. Consequently, distinct employment strategies are developed to accommodate the Shewhart charts. Some use the chart as a straightforward monitoring device that alarms workers of a possible out-of-control process. Others use it as a control device that not only alerts workers but also provides basis for process adjustment.

This Shewhart chart scheme [21] is in effect a statistical hypothesis testing that reveals only whether the process is still in-control. This is certainly not enough for practical applications where the causes of the problem should be found and the production should be restored to its satisfactory operational state. It is the search for the out-of-control causes and reinstatement of the process that complete the statistical process control loop. To reinstate the process, necessary corrective actions or process adjustment are performed following the chart's signaling. When corrective actions are taken, a critical step is to search for the cause of the out-of-control occurrence. During this search process, a robust process mean estimate can serve as a valuable reference that helps to determine the size of the process deviation and thus limit the search to a selective set of possible causes. Sometimes, process adjustment is made following the out-of-control signal to bring back the departed process. The size of the adjustment is usually based on the runs pattern of the data points and most frequently on the judgment of experienced process engineers. An accurate estimate of the mean will certainly provide a more precise adjustment of the process. In this paper, a mean estimate for Shewhart-chart-monitored processes is developed to fulfill these needs.

Conventional process mean estimates usually assume a

process that is governed by a linear stochastic system and is described using state-space models or autoregressive integrated moving average (ARIMA) models. Harrison and Stevens [11][12], apply state-space models to situations where shifts are exhibited both in the process level and the slope. As the authors point out, their model for shifts in the mean is, in a way, analogous to ARIMA (0, 1, 1) model in [4][5]. This integrated moving average (IMA or ARIMA (0, 1, 1)) model can be transformed into an exponentially weighted moving average (EWMA) estimator, which in turn can be realized as the popular PID controller. Recently, the EWMA estimator attracts attention from both researchers [6][14][16][26] and practitioners [3][20]. Though our proposed mean estimator is in essence a weighted moving average, the idea behind the estimator is entirely different from the conventional ones.

Process adjustment based on process mean estimation is a feedback control scheme [1][4][5][10] that is generally referred to as Automatic Process Control (APC) approach. Researchers in the APC area focus most on identifying the process's dynamic nature and constructing a linear system to describe the process. Whereas, the Shewhart chart is intended to identify the process departure from its normal state. This is referred to as SPC approach. Extensive discussions about APC and SPC can be found in [6][16][17][20][22][23]. While we agree that these two approaches have their own roles and should not be obscured with their seemingly related names, we also think that a plausible mean estimate should be developed to provide basis for further actions when a process is signaled out-of-control by a statistical control chart.

Barnard [2] is the first to point out the possibility of using control chart as an element in a feedback control loop. He suggests that the control chart should not only be used as a monitoring device but also as an estimator. The work, however, receives limited attention from industrial practitioners. This is possibly because the cumulative sum (CUSUM) control chart [8][18][19] used by Barnard is relatively complex and is less popular than the Shewhart control chart. The Shewhart chart remains the dominant control chart in practice because of its simplicity and the development of supplementary runs rules [13]. Nevertheless, Barnard's original idea provides a new perspective for applications of control charts. In this paper, we follow Barnard's philosophy but focus on the more popular Shewhart control chart.

This paper is organized into four sections. First, the process model is described and the basic assumptions of the Shewhart control chart are discussed. In the second section, we develop the mean estimate for the Shewhart chart with a single  $3\text{-}\sigma$  rule and a constant shift occurrence probability followed by an example and some discussions. We also investigate the situation where supplementary runs rules are

used and how the methodology can be applied.

### PROCESS MODEL

Throughout this paper, we will use a semiconductor fabrication process as an example to illustrate the SPC ideas and to demonstrate the use of the proposed mean estimate. Among various semiconductor processes, a very important process is to grow a layer of silicon dioxide (SiO<sub>2</sub>). To grow an SiO<sub>2</sub> layer, oxidation in a furnace is a widely used technique especially for growing a gate oxide layer on the Metal-Oxide-Semiconductor (MOS) device. The critical output characteristic of this oxidation process is the thickness of the SiO<sub>2</sub> layer. After oxidation, the thickness (Å) is measured and plotted on a Shewhart control chart to track the process condition. Process engineers often use the control chart not only as a monitoring tool but also a trigger for making process adjustment.

Let the random variables  $X_1, X_2, X_3, \dots$  denote the 1st, 2nd, 3rd, ... successive sample statistics of the quality characteristic observations, say the thickness of the SiO<sub>2</sub> layer. The sample statistic  $X_i$  can be a single observed value or a sample mean of several observed values. It is assumed that  $X_i$  is normally distributed with mean  $\mu_i$  and variance  $\sigma^2$ . In the Shewhart control chart scheme, the process mean  $\mu$  is initially set to coincide with the target ( $T$ ). For instance, let the target of the oxide thickness be 250 Å. When the process is in-control,  $X_i$ 's are said to share the same mean  $\mu (=250)$  unless the following condition is violated:

$$\mu - 3\sigma \leq X_i \leq \mu + 3\sigma$$

where  $\mu - 3\sigma$  and  $\mu + 3\sigma$  are called, respectively, lower and upper control limits. Let the standard deviation ( $\sigma$ ) of the thickness be 1.5. When one of the control limits is exceeded at observation  $i$ , the chance that the mean of  $X_i$  remains 250 is

$$\alpha = P(X_i > 254.5; \mu_i = 250) = P(X_i < 245.5; \mu_i = 250) \\ = 1 - \Phi(3) = \Phi(-3) = 0.0013$$

where  $\Phi(\bullet)$  is the cumulative distribution function of a standard normal random variable. When any of the control limits is exceeded, the process is said to be likely out-of-control; i.e., the process mean has probably departed from the target 250. However, there is still a probability of 0.0013 that the process is actually in control. This probability is referred to as Type I error probability and is denoted by  $\alpha$ .

When one of the control limits is exceeded at the  $i$ th observation, this does not mean that the process goes out-of-control exactly at this sample point. The process could go out-of-control long before  $i$  but the control limit is not exceeded until observation  $i$ . When the process is actually out-of-control but control limits are not exceeded, this misinterpretation of the process state is called Type II error. The probability of this Type II error's occurrence can be calculated given the size of the shift. For instance, if the mean changes from 250 to 253; that is, the mean of the thickness deviates from the target by  $2\sigma$ , then the Type II error probability, denoted by  $\beta$ , using the Shewhart chart scheme is:

$$\beta = P(X_i \in [245.5, 254.5] | \mu_i = 253) \\ = \Phi\left(3 - \frac{253 - 250}{1.5}\right) - \Phi\left(-3 - \frac{253 - 250}{1.5}\right) = 0.841$$

As mentioned earlier, this Shewhart scheme reveals only a hypothesis testing result. In the following sections, we will go one step further to estimate the process mean when the Shewhart chart signals an out-of-control.

### BASIC ESTIMATION PROCEDURE

Before estimating the process mean based on the Shewhart chart's signaling, we first examine the case where the point of the mean change is known. Given this change point information, two conditional mean estimators can be easily derived. Suppose the current observation is  $j$  and the most recent mean change occurred at observation  $k$ . A straightforward estimate of the process mean is the sample mean; i.e., the arithmetic average of observations  $k$  to  $j$ :

$$\bar{x}_{j,k} = \frac{\sum_{i=k}^j x_i}{j - k + 1} \quad (1)$$

where  $x_i$  is the observed value of  $X_i$ . Since the shift occurred at  $k$ , all observations from observation  $k$  forward share the same mean. This justifies the estimate of Equation (1). However, this estimate suffers from high instability especially when the sample size  $j - k + 1$  is small. After enough experience is cumulated and information about the amount of mean changes is gathered, the past experience can actually serve as the prior knowledge on how the process mean is likely to shift. Bayes's estimation technique can then be applied to stabilize the mean estimate. Since the Shewhart chart's assumption about the process is normal distribution, a proper conjugate prior distribution is, therefore, also a normal distribution. Let the prior distribution's mean be  $\xi$  and variance be  $\tau^2$ . Bayes estimator under a quadratic loss function can be obtained (Lehmann (1991), pp 236-249) as follows:

$$\hat{\mu}_{j,k} = \frac{\tau^2 \bar{x}_{j,k} + \frac{\sigma^2}{j - k + 1} \xi}{\frac{\sigma^2}{j - k + 1} + \tau^2} \quad (2)$$

Examination of Equation (2), one can find that when the number of observations ( $j - k + 1$ ) between the current point and the change point is large, then more weight is assigned to  $\bar{x}_{j,k}$ . On the other hand, when  $j - k + 1$  is small,  $\xi$  is assigned a higher weight; that is, the past experience becomes more important in the estimate. Also, the mean estimate approaches  $\xi$  when  $\tau^2$  is relatively small. This is because a small  $\tau^2$  means that we learn from experience that the process mean is almost a deterministic value  $\xi$  whenever a shift occurs.

We have estimated the process mean given that the change point is known. The change point is, however, unlikely to be known. The Shewhart chart is thus devised to detect the process change. The control chart does not reveal the actual location of the change point either. In other words, with the control chart detection, we only know that a mean change may have occurred at or before this detection point. To more

accurately estimate the process mean, we need to know where the change point is likely to reside. A lag-window of observation points prior to the control chart's detection is defined to include all possible points at which the shift may occur.

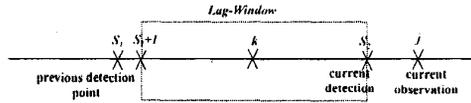


Figure 1: Shift Occurrence Lag-Window

Suppose that the current control chart's detection point is at  $S_2$  and the one preceding it is at  $S_1$ . In practice, corrective actions are taken after  $S_1$  and when the process is in operation again the process mean is immediately subject to possible new shifts. As shown in Figure 1, the lag-window is defined as the time period containing all observation points between  $S_1$  and  $S_2$  including  $S_2$ . The lag-window's definition depends greatly on the out-of-control practice. In any case, the lag-window's starting point should lie at the point where the process is deemed back to its mean condition.

Given the current detection point at  $S_2$ , it is plausible to say that there is a probability of  $1-\alpha$  that a shift has occurred in the lag-window. Let the random variable  $K$  represent the location at which the shift occurs. We would like to know now the probability distribution of  $K$ :

$$P_{k,S_2} = Pr(K = k, \text{detection at } S_2, K \in (S_1, S_2))$$

Applying Bayes' theorem, the probability that a shift occurs at an observation point, say  $k$ , in the lag-window given the control chart detection at  $S_2$  can be expressed as:

$$P_{k,S_2} = \frac{Pr(K = k, \text{detection at } S_2, K \in (S_1, S_2))}{Pr(\text{detection at } S_2, K \in (S_1, S_2))} = \frac{Pr(\text{detection at } S_2, K = k, K \in (S_1, S_2)) Pr(K = k, K \in (S_1, S_2))}{\sum_{l=S_1+1}^{S_2} Pr(\text{detection at } S_2, K = l, K \in (S_1, S_2)) Pr(K = l, K \in (S_1, S_2))} \quad (3)$$

To compute  $P_{k,S_2}$ , two probabilities  $Pr(\text{detection at } S_2, K = k, K \in (S_1, S_2))$  and  $Pr(K = k, K \in (S_1, S_2))$  should be first calculated.

$Pr(\text{detection at } S_2, K = k, K \in (S_1, S_2))$  is the probability for the shift to be discovered at  $S_2$  given that the shift is known to occur at  $k$  and  $k \in (S_1, S_2)$ . This probability can be seen as the probability of taking  $S_2 - k + 1$  observations to detect the mean change.  $S_2 - k + 1$  is also called the run length, i.e. the number of observations elapsed since the change point, before the control chart signals. Given an estimate of the shift size one can calculate the Type II error probability ( $\beta_{S_2,k}$ ), and the run length probability distribution is simply the geometric distribution with the parameter  $\beta_{S_2,k}$ .

$$Pr(\text{detection at } S_2, K = k, K \in (S_1, S_2)) = \beta_{S_2,k}^{S_2-k} (1 - \beta_{S_2,k}) \quad (4)$$

Equation (4) can be explained as follows. When the mean is shifted at  $k$ , the shift is not detected (with Type II error probability  $\beta_{S_2,k}$ ) until observation  $S_2$  (with probability  $1 - \beta_{S_2,k}$ ). To calculate  $\beta_{S_2,k}$ , one should know first how much the mean is shifted. Since the shift occurred at  $k$ , the shifted mean can be estimated by Equation (1) or (2). And

$$\hat{\mu}_{S_2,k} = \Phi\left(3 - \frac{\hat{\mu}_{S_2,k} - T}{\sigma}\right) - \Phi\left(-3 - \frac{\hat{\mu}_{S_2,k} - T}{\sigma}\right) \quad (5)$$

Again, we assume here that when the process is in-control the process mean is consistent with the target  $T$  and when the process is shifted the process mean is deviated from the target by  $\hat{\mu}_{S_2,k} - T$ .

$Pr(K = k | K \in (S_1, S_2))$  is the probability for the shift to occur at  $k$  given that the shift is known to be in the lag-window. To calculate this probability, we need to consider the probability for a shift to occur at any time. Here, we first assume that the occurrence probability at any time is a constant  $p$ . This implies that the probability for a shift to occur is always constant. Under this assumption, we can immediately obtain:

$$Pr(K = k | K \in (S_1, S_2)) = \frac{p(1-p)^{S_2-S_1-1}}{\sum_{l=S_1}^{S_2} p(1-p)^{S_2-S_l-1}} = \frac{1}{S_2 - S_1} \quad (6)$$

Substituting Equations (4) and (6) into Equation (3), we have

$$P_{k,S_2} = \frac{\beta_{S_2,k}^{S_2-k} (1 - \beta_{S_2,k})}{\sum_{l=S_1+1}^{S_2} \beta_{S_2,l}^{S_2-l} (1 - \beta_{S_2,l})} \quad (7)$$

We are now ready to estimate the process mean at the detection point. There is a probability of  $1-\alpha$  that the process mean is indeed shifted and only a probability of  $\alpha$  that the control chart gives a false alarm and the process mean is actually unchanged. The mean of  $X_{S_2}$  can be then estimated as:

$$\hat{\mu}_{S_2} = \alpha T + (1 - \alpha) \sum_{K=S_1+1}^{S_2} \hat{\mu}_{S_2,K} P_{K,S_2} \quad (8)$$

where  $\hat{\mu}_{S_2}$  is the estimated mean of  $X_{S_2}$  given a known change point at  $K$  and  $P_{K,S_2}$  is the probability for the shift to occur at  $K$  given a control chart detection at  $S_2$ .

When the process is not shut down after the control chart's signaling and observation samples taken from the process are continued, then the process mean can be estimated for later observation points, say  $j$  and  $j > S_2$ , in the same manner:

$$\hat{\mu}_j = \alpha T + (1 - \alpha) \sum_{K=S_1+1}^{S_2} \hat{\mu}_{j,K} P_{K,S_2} \quad (9)$$

where

$$P_{kS_j} = \frac{\beta_{jk}^{S_j-k} (1 - \beta_{jk})}{\sum_{l=1}^{S_j} \beta_{jl}^{S_j-l} (1 - \beta_{jl})} \quad (10)$$

and

$$\beta_{jk} = \Phi\left(3 - \frac{\hat{\mu}_{jk} - T}{\sigma}\right) - \Phi\left(-3 - \frac{\hat{\mu}_{jk} - T}{\sigma}\right)$$

As mentioned earlier, the mean estimate by Equation (9) is in essence a weighted average of the observations. When  $\hat{\mu}_{jk}$  is estimated by  $\bar{X}_{jk}$ , we can summarize the weights for each observation  $X_i$  as follows:

$$\hat{\mu}_j = \alpha T + (1 - \alpha) \sum_{i=N_j+1}^{S_j} \sum_{j=K+1}^i \frac{P_{K S_j}}{j - K + 1} X_i + \sum_{i=N_j}^j \sum_{K=N_j+1}^i \frac{P_{K S_j}}{j - K + 1} X_i \quad (11)$$

### EXAMPLE

In the oxidation process, the target SiO<sub>2</sub> thickness is 250Å. After preventive maintenance (PM), the process is in operation with process mean set to the target. Also, the thickness S.D. is known to be 1.5 Å. A Shewhart control chart is then constructed to monitor the oxidation condition. We have here 70 observations. The thickness mean is shifted at the 50th observation with a magnitude 3.0Å to become 253Å. Figure 2 shows the Shewhart control chart.

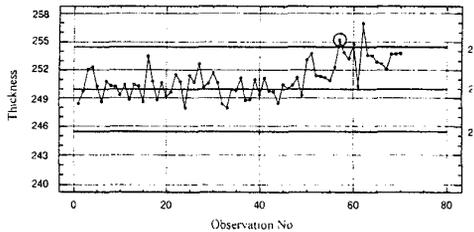


Figure 2: The Shewhart Control Chart for Oxidation Thickness

As can be seen, at observation 57 the control chart gives an alarm. We now apply the mean estimate to this detection observation and the observations that follow. The lag-window here includes Observations 1 to 57. That is, the shift could have actually occurred at any of these observations. To compute the mean estimate for the detection observation 57, we first calculate sample means for  $\bar{X}_{j^*j^*}, \bar{X}_{j^*-1j^*}, \dots, \bar{X}_{1j^*}$ . Corresponding to these sample means, we can estimate the Type II error probabilities:  $\beta_{j^*j^*}, \beta_{j^*-1j^*}, \dots, \beta_{1j^*}$  using equation (5):

$$\beta_{j^*k} = \Phi\left(3 - \frac{\hat{\mu}_{j^*k} - 250}{1.5}\right) - \Phi\left(-3 - \frac{\hat{\mu}_{j^*k} - 250}{1.5}\right)$$

for  $K=1, \dots, 57$

The shift occurrence probability given signaling at 57 can be then estimated using Equation (7):

$$P_{kS_j} = \frac{\beta_{j^*k}^{S_j-k} (1 - \beta_{j^*k})}{\sum_{l=1}^{S_j} \beta_{j^*l}^{S_j-l} (1 - \beta_{j^*l})} \quad \text{for } K=1, \dots, 57$$

The mean estimate for observation 57 is calculated as

$$\hat{\mu}_{57} = (0.00135)(250) + 0.99865 \sum_{k=1}^{57} \bar{X}_{j^*k} P_{K S_j}$$

The same procedure is repeated for estimating means for observations 58 and beyond. Figure 3 shows the estimated mean together with the observations.

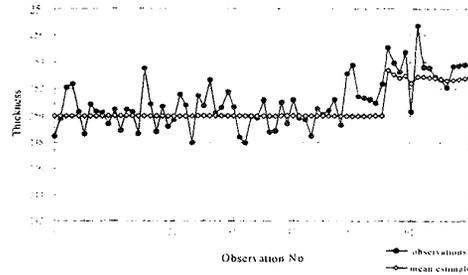


Figure 3: Mean Estimate for Shewhart-Chart-Monitored Process

One can observe from Figure 3 that the mean estimate at the signal observation overshoots a little given only one observation available after detection. The mean estimate is also seen to oscillate before gradually stabilize. Using Bayes estimate given by Equation (2), when enough experience is gained, can significantly reduce the overshooting and oscillation at the beginning of the estimation. Suppose the shifted mean is learned to vary around 252.5Å with S.D. 2.0Å. Then,

$$\hat{\mu}_{jk} = \frac{4.0 \bar{X}_{jk} + \frac{2.25}{j-k+1} 252.5}{\frac{2.25}{j-k+1} + 4.0} \quad \text{for } k=1, \dots, 57$$

Figure 4 shows a comparison between mean estimates with and without the Bayes stabilization.

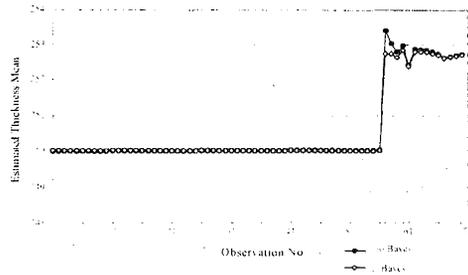


Figure 4: Mean Estimates with and without Bayes Stabilization

The Shewhart-based mean estimate is a weighted average of

available observations. It is interesting to see how the weights vary over the age of the observation. Figure 5 shows the weight assigned by the estimate on each observation.

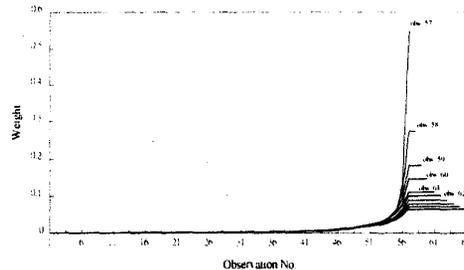


Figure 5: Weights for Observations taken by Shewhart-based Mean Estimate

From Figure 5, the proposed mean estimate looks very much like a dynamic weighted average of available observations. For example, for the detection observation's mean estimate, the weight taken on observation 57 itself is about 0.55 and then exponentially decreases for older observations. When the mean estimate continues for observations after signaling, the weight distribution is dynamically adjusted. While the weights exponentially decrease for those observations sampled before detection, the observations after signaling are equally weighted. Take Observation 62 as an example, the mean estimate for this data point takes observations 57 to 61 with an equal weight about 0.1. But weights for observations 56 and earlier decrease drastically. This kind of weight distribution seems quite plausible for the situation where the out-of-control occurrence takes only the form of mean shift and its occurrence is solely tested by the Shewhart chart.

#### MEAN ESTIMATE WITH RUNS RULES

It is well known that the original Shewhart chart is not effective in detecting smaller process shifts. The popularity of the Shewhart chart lies not only on its simplicity but also on the development of customized runs rules. In addition to improving the control chart's ability in detecting shifts, runs rules can be installed to identify patterns that are rather difficult to detect by the single out-of- $3\sigma$  rule. Typical runs rules to detect small shifts are:

1. A single observation falls outside of the  $3\sigma$  limits;
2. Two out of three consecutive observations fall between  $2\sigma$  and beyond from the average or between  $-2\sigma$  and beyond from the average;
3. Four out of five consecutive observations fall between  $+1\sigma$  and beyond from the average or between  $-1\sigma$  and beyond from the average; and
4. Eight consecutive observations fall on a given side of the central line.

These four runs rules were first introduced in *Statistical Quality Control Handbook* (1956) and has been intensively studied since.

The effectiveness of a control chart is usually evaluated by Type I and Type II error probabilities. When runs rules are applied, the testing is not based on a single test statistic but is based on runs of several observations. Here, Average Run

Length's (ARL's) are used instead to evaluate the control chart's ability. While the in-control ARL ( $ARL_0$ ) assesses how frequently the control chart is expected to give false alarms, the out-of-control ARL ( $ARL_1$ ) measures how quickly the control chart is expected to signal a process shift. To evaluate the ARL's, one can apply Markov chain approach to calculate the Run Length (RL) probability distribution. Champ and Woodall (1987) have defined the Markov chain states for the runs rules that yield the minimal number of states.

As explained in the previous section, the out-of-control run length ( $RL_j$ ) probability given an approximated shift size is needed in order to calculate  $P_{k,S_j}$ . That is, we need to calculate

$$P_{k,S_j} = P(\text{detection at } S_2, K = k, K \in (S_j, S_2, j)) \\ = P(RL_j = S_2 - k + 1 \mid \text{shift} = \hat{\mu}_{1,k} - T), j \geq S_2 \quad (12)$$

Depending on the runs rules used, the calculation of the run length probability can become extensive. For example, it takes 216 states in the Markov chain to calculate the run length distribution when the above four runs rules are implemented simultaneously. Computer programs will be required in this case to implement the mean estimate for on-line production use. The calculation procedure using the Markov chain approach will not be repeated in this paper. Interested readers should refer to [7].

In Equation (9), we also need to know the probability of false alarm at  $S_2$ . Since runs rules are used, the Type I error probability  $\alpha$  is not clearly defined. Instead, the false alarm probability here can be calculated using in-control run length ( $RL_0$ ) probabilities. It is equivalent to the probability for the control chart to falsely signal after  $S_2-S_1$  observations, that is, the probability that  $RL_0$  equals  $S_2-S_1$  given  $RL_0 \geq S_2 - S_1$ :

$$P_f = P(\text{false alarm at } S_2, j) = \frac{P(RL_0 = S_2 - S_1)}{P(RL_0 \geq S_2 - S_1)} \quad (13)$$

When the single  $3-\sigma$  rule is used, the false alarm probability can be calculated using Equation (13) as:

$$P_f = \frac{(1-\alpha)^{S_2-S_1} + \alpha}{(1-\alpha)^{S_2-S_1} + 1} = \alpha$$

which is equal to the Type I error probability.

#### CONCLUSIONS

In this paper, we have developed methodologies to estimate the process mean with a Shewhart chart installed to test possible process shifts. The proposed mean estimate was actually a dynamic weighted average of historic data. It estimated the Shewhart chart's test sensitivity and used it to dynamically adjust the weights for the historic data. The mean estimate was further refined by utilizing the past experience as the prior knowledge for shift occurrence rate and shift size. The use of the mean estimate was then demonstrated through an example of the semiconductor oxidation process.

#### REFERENCES

- [1] Astrom, K. J. (1970). *Stochastic Control Theory*. Academic Press, New York.

- [2] Barnard, G. A. (1959), "Control Charts and Stochastic Processes", *Journal of Royal Statistics Society, Series B*, 21, 239-257.
- [3] . Boning, D., A., Hurwitz, J. Moyne, W. Moyne, T. Smith, J. Taylor and R. Telfeyan, "Run by Run Control of Chemical Mechanical Polishing." *IEEE Trans. on Components, Packaging and Manufacturing Technology - Part C*, 19, 1, pp. 307-314.
- [4] Box, G. E. and G. M. Jenkins (1968), "Some Recent Advances in Forecasting and Control", *Applied Statistics*, 17, 2, 91-109.
- [5] . Box, G. E. and G. M. Jenkins (1970), *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.
- [6] Box, G. E. and T. Kramer (1992), *Statistical Process Monitoring and Feedback Adjustment - A Discussion*", *Technometrics*, 34, 3, 251-285.
- [7] . Champ and Woodall (1987), "Exact Results for Shewhart Control Charts with Supplementary Runs Rules", *Technometrics*, 29, 4, 393-399.
- [8] Chernoff, H. and S. Zacks (1964), "Estimating the current mean of a normal distribution which is subject to change in time", *Ann. Math. Stat.*, 35, 999-1018.
- [9] Cox, D. R. (1962), *Renewal Theory*. Chapman and Hall, New York.
- [10] Davis, M. H. A. and R. B. Vinter (1985), *Stochastic Modeling and Control*. Chapman and Hall, New York.
- [11] Harrison, P. J. and C. F. Stevens (1971), "A Bayesian Approach to Short-TERM Forecasting", *Operational Research quarterly*, 22, 4, 341-362.
- [12] . Harrison, P. J. and C. F. Stevens (1976), "Bayesian Forecasting", *Journal of Royal Statistics Society, Series B*, 38, 205-247.
- [13] Hoerl, R. W. and A. Palm (1992), "Discussion: Integrating SPC and APC", *Technometrics*, 34, 3, 268-272.
- [14] . Ingolfsson, A. and E. Sachs (1993), "Stability and Sensitivity of an EWMA Controller," *Journal of Quality Technology*, 25, 4, pp. 271-287.
- [15] . Lehmann, E. L. (1991), *Theory of Point Estimation*, 236-249. Wadsworth, Pacific Grove, California.
- [16] MacGregor, J. F. (1987), "Interfaces Between Process Control and Online Statistical Process Control", *Computing and Systems Technology Division Communications*, 10, 9-20.
- [17] MacGregor, J. F. (1988), "On Line Statistical Process Control", *Chemical Engineering Progress*, Oct, 21-31.
- [18] Page, E. S. (1954), "Continuous inspection schemes", *Biometrika*, 41, 100-116.
- [19] . Page, E. S. (1955), "A test for a change in a parameter occurring at an unknown point", *Biometrika*, 42, 523-526.
- [20] . Page, E. S. (1955), "A test for a change in a parameter occurring at an unknown point", *Biometrika*, 42, 523-526.
- [21] Sachs, E., A. Hu, and A. Ingolfsson (1995), "Run by Run Process Control: Combining SPC and Feedback Control," *IEEE Trans. Semiconductor Manufacturing*, 8, 1, pp. 26-43.
- [22] Shewhart, W. A. (1931), *Economic Control of Quality Improvement*. Van Nostrand Reinhold Company, Princeton.
- [23] Tucker, W. T., F. W. Faltin and S. A. Vander Wiel (1993), "Algorithmic Statistical Process Control: An Elaboration", *Technometrics*, 35, 4, 363-375.
- [24] . Vander Wiel, S. A., W. T. Tucker, F. W. Faltin, and N. Doganaksoy (1992), "Algorithmic Statistical Process Control: Concepts and an Application", *Technometrics*, 34, 3, 286-297.
- [25] Western Electric Company (1958), *Statistical Quality Control Handbook*. Western Electric Company, New York.
- [26] . Woodall, W.H. and M. R. R. Reynolds, Jr. (1983), "A Discrete Markov Chain Representation of the Sequential Probability Ratio Test", *Communications in Statistics - Sequential Analysis*, 2, 1, 27- 44.
- [27] Yashchin, E. (1995), "Estimating the Current Mean of a Process Subject to Abrupt Changes." *Technometrics*, 37, 3, pp. 311-323.