International Journal of Innovative Computing, Information and Control
Volume **9**, Number **9**, September **2013**
ICIC International ©2013 ISSN 1349-4198
pp. 3667–3676

APPLICATION OF Q CHARTS FOR SHORT-RUN AUTOCORRELATED DATA

Hironobu Kawamura[1], Ken Nishina[2], Masanobu Higashide[3] and Tomomichi Suzuki[4]

[1]Information and Systems
University of Tsukuba
1-1-1, Tennodai, Tsukuba-shi 305-8577, Japan
kawamura@sk.tsukuba.ac.jp

[2]Faculty of Engineering
Nagoya Institute of Technology
Gokiso-Cho, Showa-Ku, Nagoya City 466-8555, Japan
nishina@nitech.ac.jp

[3]Quality Assurance Division
Renesas Electronics Corporation
1753, Shimonumabe, Nakahara-Ku, Kawasaki, Kanagawa 211-8668, Japan
masanobu.higashide.wt@renesas.com

[4]Faculty of Science and Technology
Tokyo University of Science
2641, Yamazaki, Noda City 278-8510, Japan
szk@rs.tus.ac.jp

Received July 2012; revised February 2013

Abstract. *Traditional control charts are designed for processes where data are independent and identically distributed and where large historical data sets are available before a production run for estimating process parameters and computing control limits. Many processes, particularly in semiconductor manufacturing, often involve limited data sets, which result from high-mix low-volume production; in many cases, data are correlated. Therefore, charting techniques for treating short-run autocorrelated data are required. In this paper, we assess the effectiveness of control charts based on Q statistics by applying them to actual data obtained from a horizontal low-pressure chemical vapor deposition process used in semiconductor manufacturing. Our results show that Q charts enable the plotting of different types of data on the same chart, and that Q charts can detect real anomalies in data. Furthermore, we show that both Q statistics and Q statistics applied to the residuals of a time series model are practical, useful methods for the processes employed in semiconductor manufacturing.*
Keywords: Control chart, Time series analysis, Autocorrelation, Semiconductor manufacturing process, High-mix low-volume production, Statistical process control, Q statistics

1. **Introduction.** Traditional control charts are designed for processes where data are independent and identically distributed (*i.i.d.*) and where large amounts of historical data are available before a production run for estimating process parameters and computing control limits. However, two drawbacks in traditional statistical process control (SPC) techniques have been noted recently.

First, because traditional SPC approaches are intended for use in high-volume manufacturing, they may not be effective for a short time series [1]. Montgomery [2] and Quesenberry [3] recommended collecting at least 25 or 30 samples, each consisting of four or five data, for establishing meaningful control limits before starting a production run. However, in many situations, a process does not yield sufficient observations for traditional SPC tools to be used effectively. For example, process start-ups, major tool changes, different raw materials, new equipment, and production processes do not allow frequent measurement [4].

Several control chart methods designed for short runs have been introduced in the SPC literature. The method suggested by Yang and Hillier [5] involves adjusting standard control limits in a control chart to achieve the desired type I error probability, irrespective of the number of initial subgroups. Farnun [6] proposed using X-bar and R charts to plot the quality characteristic scaled by a predetermined value. Quesenberry [7] developed Q charts, which allow a user to plot different statistics on the same chart. Del Castillo et al. [8] reviewed short-run SPC techniques.

Second, traditional SPC tools (e.g., control charts and process capability indices) assume that the process data are uncorrelated. However, the violation of this assumption is prevalent in chemical and continuous industries [9]. Autocorrelations decrease the effectiveness of control charts that assume independent process measurements [10].

Different methods for dealing with autocorrelated data have been discussed. One approach involves modifying existing SPC charts by adjusting the control limits [11,12]. A second approach assumes that the residuals of a time series model are statistically uncorrelated. This method fits an appropriate time series model to process observations and then applies traditional control charts to the residuals [10,13,14]. Mastrangelo and Montgomery [15] used the residuals from an EWMA. SPC procedures for dealing with autocorrelated processes were reviewed by Psarakis and Papaleonida [16].

The problem of short-run autocorrelated data has received marginal attention [17]. Prasad [18] recommended using the joint estimation (JE) procedure developed by Chen and Liu [19] for controlling chemical production processes. Wright et al. [20] investigated the JE procedure as an SPC tool for short-run autocorrelated data. Crowder and Halblieb [21] proposed an adaptive filtering approach for process monitoring with short-term autocorrelated data. Snoussi et al. [22] suggested using the residuals of a time series model in conjunction with Q statistics when the process parameters are unknown. Snoussi and Limam [17] also proposed the unknown parameters change-point formulation in conjunction with the residuals of various time series models.

High-mix low-volume production in semiconductor manufacturing allows only a few observations, and the data are often correlated. Kawamura et al. [23] showed that polysilicon film thicknesses in vertical low-pressure chemical vapor deposition (LPCVD) process are correlated. Therefore, charting techniques for short-run autocorrelated data are required in such processes. Del Castillo [24] considered a multivariate short-run case for autocorrelated data in the wet etching process used in semiconductor manufacturing, but did not focus on the methods for the univariate case, such as Q charts and the JE procedure.

However, the short-run and short-run autocorrelated methods have not been sufficiently investigated from a practical perspective. In this paper, we show that both Q statistics and Q statistics applied to the residuals of a time series model are practically useful methods for the horizontal LPCVD process used in semiconductor manufacturing. We have three reasons for focusing on a Q chart. First, among the other methods, only the use of a Q chart enables the plotting of data sets for each different type on the same chart, although most researchers have not strengthened the advantage. If a Q chart is not used, it will be necessary to plot a separate control chart for each type. As the number of types increases, many control charts are needed and must be interpreted, resulting in highly inefficient use of time. Second, users can easily understand Q charts, because the transformation that yields the Q statistics resembles the well-known standardization formula. Finally, it is not clear whether Q charts are practically useful because no known study considers Q charts with an actual production data set.

2. **Case Study of Horizontal LPCVD.** The horizontal LPCVD process is outlined in Figure 1. A wafer is introduced into a silicon carbide reaction chamber and heated to several hundred degrees Celsius under reduced pressure, after which polysilicon raw material gas is introduced. A polysilicon film is then formed by an endothermic chemical reaction. This process allows the production of a batch of multiple sets comprising up to 25 silicon wafers. The thickness of the resulting polysilicon film is a quality characteristic.

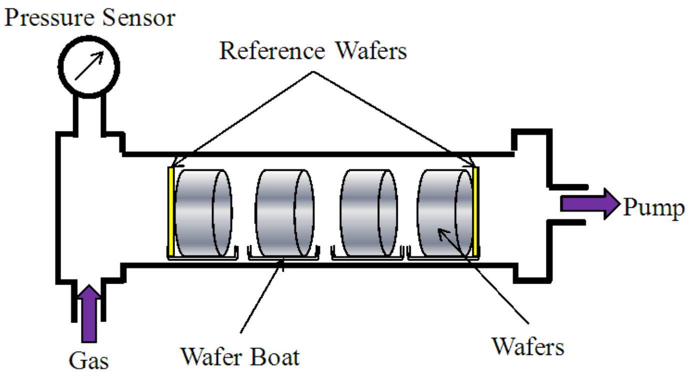

FIGURE 1. Horizontal LPCVD process

In this process, the film thickness data are obtained by measuring reference wafers, and the thickness of the films in the batches varies. Therefore, we use the average film thickness of the two reference wafers as data for assessing the Q statistics. We define ten types of wafer based on different target thickness values and obtain a data set for each type. Generally, a control chart should be designed for each type, because the mean and variance values of each type are different. As a result, many control charts are generated, and significant time is wasted in interpreting them. These are serious problems. Therefore, it is necessary to introduce Q charts that enable engineers to plot results for all different types on the same chart in this process.

3. Control Charts Using Q Statistics.

3.1. **Q charts for short-run processes.** Quesenberry [7] developed Q charts for the mean and variance values of measurements sampled from a normal distribution for cases in which both, either, or neither of the mean and variance are known. Advantages of the Q charts are as follows.

(1) They can be obtained in real time, starting at the onset of production, i.e., when the process parameters cannot be assumed before the beginning of a run and must be estimated from the available data sequence.
(2) They can be easily managed, e.g., by plotting different process variables on the same chart. This may help to identify assignable causes.
(3) Rules for point patterns to increase the power of the chart to detect assignable causes, such as the Western Electric rules [25], can be applied to all these charts.

This paper will focus on applying Q statistics to the process mean of individual measurements and concentrate on the case in which both the process mean and variance are unknown (Case UU). In many cases, both the process mean and variance are unknown because it is difficult to collect many data due to the shortening of a product life cycle.

Let X_1, X_2, $\cdots$ represent successive measurements of a sequence of types as they are produced over time. Assume that these values are *i.i.d.*, because they are collected from a normal distribution $N(\mu, \sigma^2)$ with a mean μ and variance σ^2. Q charts are generated for the cases in which both, either, or neither of these parameters is assumed prior to the production run. In this section, we focus on obtaining a Q chart for Case UU, because our objective is to apply it practically. We obtain a Q chart by calculating the following statistics:

$$Q_r\left(X_r\right) = \Phi^{-1}\left\{G_{r-2}\left[\left(\frac{r-1}{r}\right)^{\frac{1}{2}}\left(\frac{X_r - \bar{X}_{r-1}}{S_{r-1}}\right)\right]\right\}, \quad r = 3, 4, \cdots, \tag{1}$$

where Φ^{-1} denotes the inverse of the standard normal distribution function, $G_v(\bullet)$ is the Student's t distribution function with v degrees of freedom, and $\bar{X}_r$ and S_r are respectively defined as follows:

$$\bar{X}_r = \frac{1}{r}\sum_{j=1}^{r} X_j \tag{2}$$

$$S_r^2 = \frac{1}{r-1}\sum_{j=1}^{r}\left(X_j - \bar{X}_r\right)^2. \tag{3}$$

These values can be computed after each new measurement X_r according to [26]:

$$\bar{X}_r = \frac{1}{r}\left[(r-1)\,\bar{X}_{r-1} + X_r\right], \quad r = 2, 3, \cdots \tag{4}$$

$$S_r^2 = \frac{r-2}{r-1}S_{r-1}^2 + \frac{1}{r}\left(X_r - \bar{X}_{r-1}\right)^2, \quad r = 3, 4, \cdots \tag{5}$$

The corresponding Q statistics can then be plotted on Q charts centered on zero and bound by control limits at ± 3. This procedure is applicable to short-run situations where the parameters are not known in advance.

We illustrate the use of a Q chart for Case UU using simulated data with known properties. We sampled 50 values generated from $N(500, 10^2)$. Using the set parameters ($\mu = 500$ and $\sigma = 10$), we computed the 50 standardized values by the following equation:

$$Q_r(X_r) = \frac{X_r - \mu}{\sigma}, \quad r = 1, 2, \cdots \tag{6}$$

This is the formula for calculating Q statistics when both μ and σ^2 are known (Case KK). The results are plotted in Figure 2.

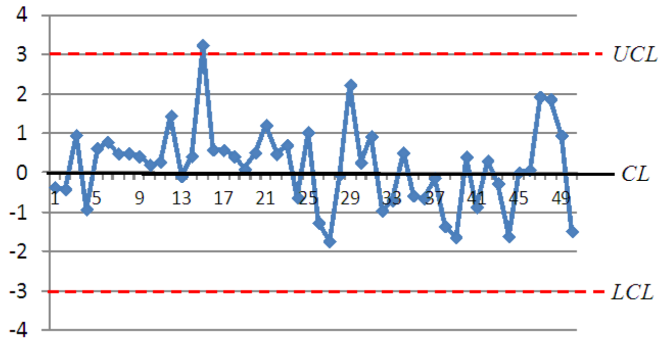

Figure 2. Q chart for Case KK: both μ and σ are known

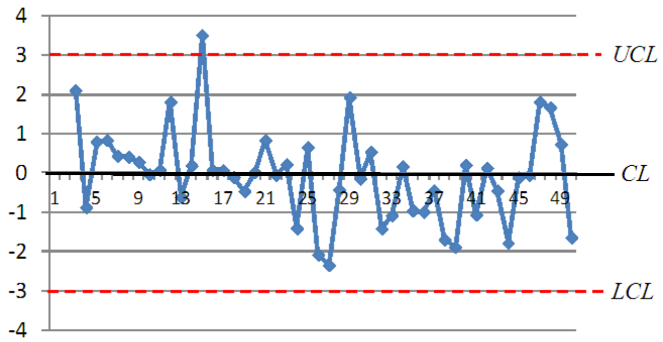

FIGURE 3. Q chart for Case UU: both μ and σ are unknown

Similarly, for Case UU, which is important for our purposes, we computed 48 Q statistics using Equation (1). These values are plotted in Figure 3. After the first few observations, the charts for Cases KK and UU become almost identical, with a correlation coefficient of 0.969.

3.2. **SCC Q charts for short-run autocorrelated data.** Q charts assume that the process data are independent; therefore, they cannot be applied to autocorrelated data. However, Snoussi et al. [22] suggested using the residuals of a time series model in conjunction with Q statistics when the process parameters are unknown.

Assume that the successive measurements X_1, X_2, $\cdots$ taken at different times are described by an ARIMA (p, d, q) time series model

$$\phi(B)\,\nabla^d X_t = \theta(B)\,a_t, \tag{7}$$

where $\phi(B)$ and $\theta(B)$ are polynomial functions (of orders p and q, respectively) of the backward difference operator (of order d) and a_t are the sequences of $i.i.d.$ normal random variables with a zero mean and a constant variance.

Let $\hat{X}_t$ be the predicted value for the observation at an instant t (i.e., calculated at the end of period $t-1$), obtained from an appropriately identified and fitted ARIMA model. Then, the residuals

$$\hat{e}_t = X_t - \hat{X}_t, \quad t = 2, 3, \cdots, \tag{8}$$

are $i.i.d.$ random variables. Applying the Q statistic with unknown process parameters to these residuals gives independent standard normal variables. The end result is a Shewhart control chart (SCC Q chart) with constant control limits [22]. The control limits are set to ± 3, i.e., UCL $= 3$ and LCL $= -3$.

4. **Application of Q Chart to Case UU.** In this section, we illustrate the use of the Q and SCC Q charts with real data sets obtained from a horizontal LPCVD process.

To illustrate and assess the behavior of the Q charts, each one is plotted after stratifying by type. Figure 4 shows the average film thickness for type 2, and Figure 5 shows the corresponding Q chart.

The high values of the correlation coefficients in Table 1 confirm that this high correlation also exists for other types. Even the lowest correlation coefficient value is as high as 0.774. Moreover, as shown in Table 1, it is also found that the horizontal LPCVD process is a short-run process, because each type has a small sample size, especially, a sample size for type 4 is very small. Therefore, if a Q chart is not used, we cannot control the process because many data are needed before designing traditional control charts.

Both charts look similar, and the correlation coefficient for an average film thickness (Figure 4) and the corresponding Q statistics (Figure 5) is 0.992. One of the advantages

TABLE 1. Correlation coefficients of Q statistics for each type

Types	Sample size	Correlation coefficient
Type 1	202	0.982
Type 2	322	0.992
Type 3	55	0.956
Type 4	9	0.993
Type 5	203	0.966
Type 6	38	0.961
Type 7	61	0.774
Type 8	47	0.882
Type 9	451	0.975

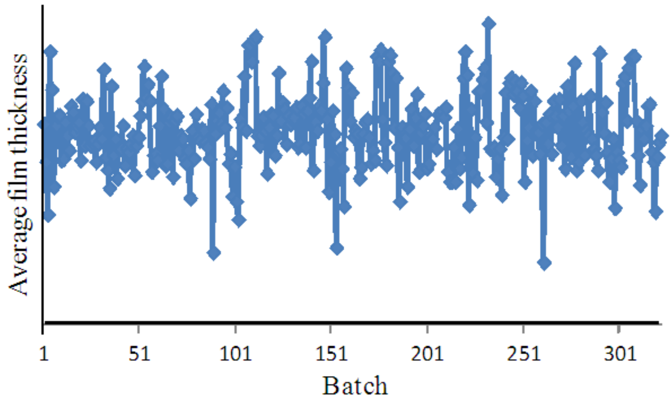

FIGURE 4. Average film thickness for type 2

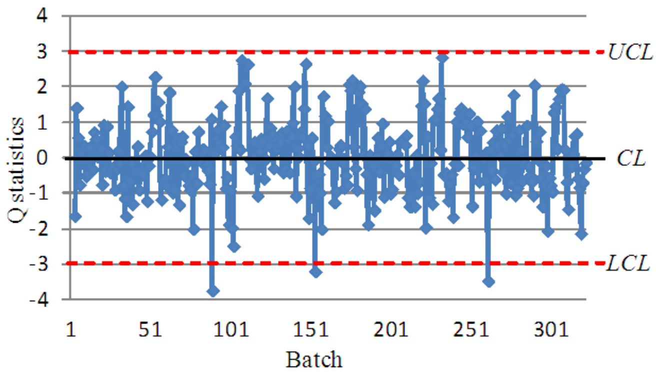

FIGURE 5. Q chart for type 2

of a Q chart is that results for all types can be plotted on the same chart, as shown in Figure 6.

The low values of the correlation coefficients for types 7 and 8 may be due to autocorrelation. Only the film thickness data for types 7 and 8 are correlated. Figure 7 shows the sample autocorrelation function for type 7. An SCC Q chart is better suited to correlated data.

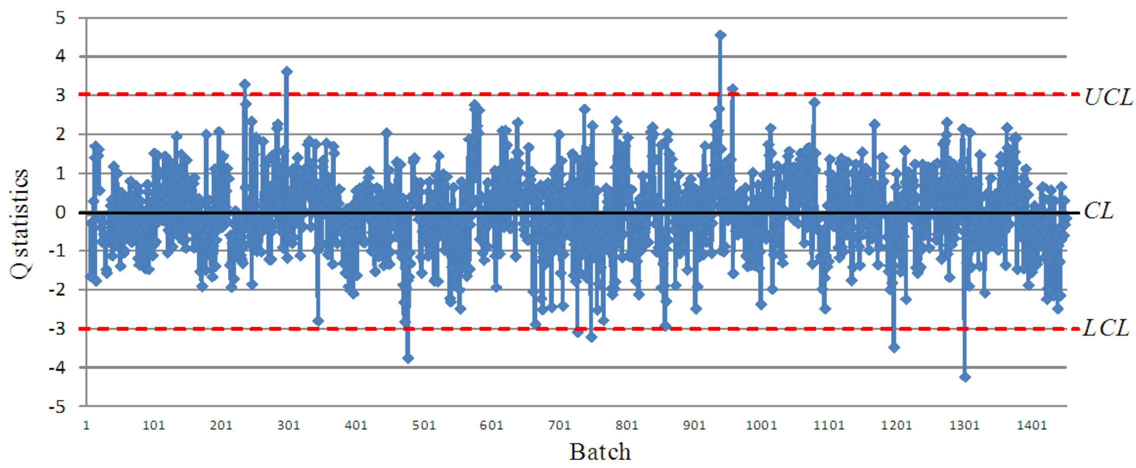

Figure 6. Q chart of all types

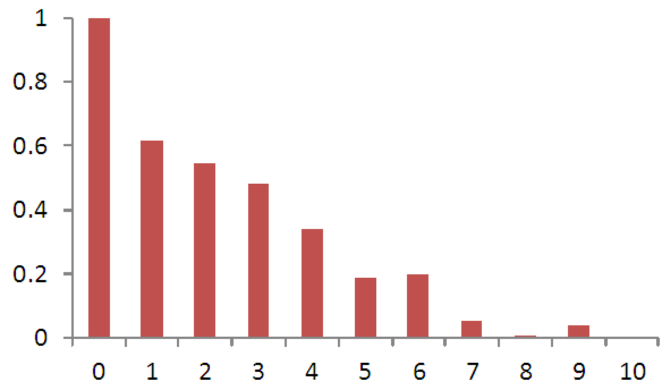

Figure 7. Sample autocorrelation function for type 7

Table 2. Assessment results of AIC and BIC

Model	AIC	BIC
ARIMA $(1, 0, 0)$	585.47	589.69
ARIMA $(2, 0, 0)$	582.11	588.44
ARIMA $(1, 1, 0)$	576.24	580.43
ARIMA $(2, 1, 0)$	574.61	580.89
ARIMA $(1, 0, 1)$	581.62	587.95
ARIMA $(2, 0, 1)$	583.25	591.69
ARIMA $(0, 1, 1)$	572.49	576.68

We select an appropriate ARIMA model based on the Akaike information criterion (AIC) and Bayesian information criterion (BIC) values, respectively defined as

$$AIC = -2\ln(L) + 2k \tag{9}$$

$$BIC = -2\ln(L) + \ln(T)k, \tag{10}$$

where k is the number of parameters in the time series model, L is the maximum likelihood function of the estimated model, and T is the sample size. We assessed several low-order ARIMA models in terms of AIC and BIC and found ARIMA $(0, 1, 1)$ to be the most efficient for types 7 and 8. Table 2 shows the assessment results of AIC and BIC.

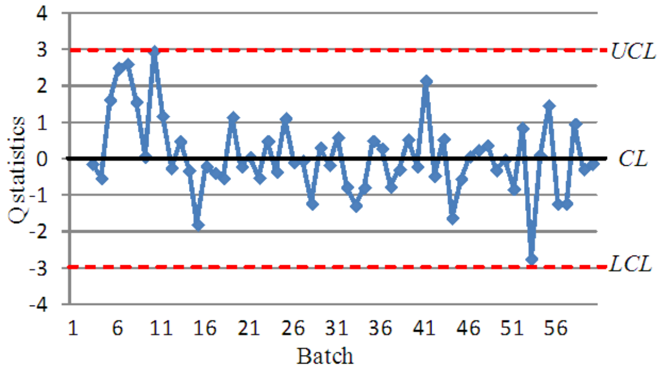

Figure 8. SCC Q chart for type 7

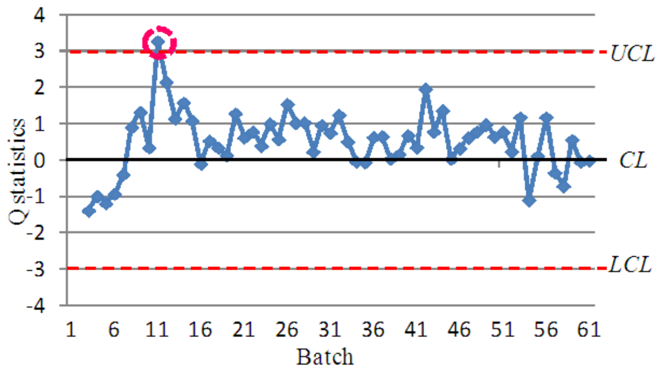

Figure 9. Q chart for type 7

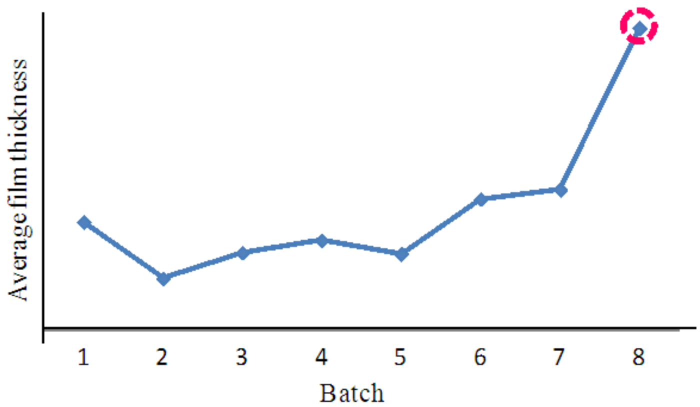

Figure 10. Data set with an anomalous data point

The SCC Q chart derived from ARIMA $(0, 1, 1)$ in Figure 8 shows no points outside the control limits at ± 3, whereas one point is outside the control limits in the Q chart for type 7 in Figure 9.

The number of points outside the control limits indicates a type I error, because the original data set was obtained from a stable process that did not include anomalous data. Therefore, the SCC Q chart is adequate for treating short-run autocorrelated data, because it decreases the type I error.

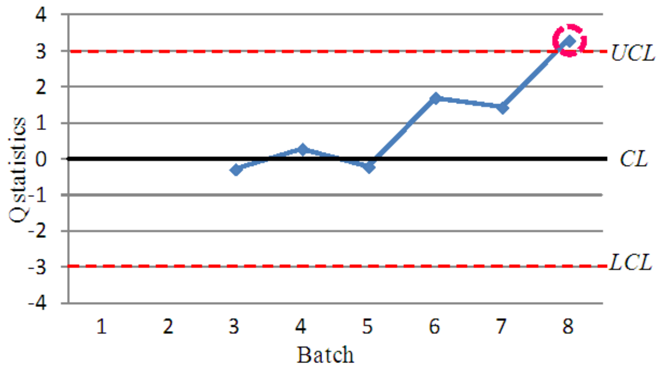

FIGURE 11. Q chart with an anomalous data point

5. **Q Chart Performance.** Many authors have argued that Q charts are inefficient at detecting shifts, e.g., [27]. In this section, we evaluate outlier detection capabilities of Q charts by using the data set of type 9, which includes real anomalous data. Figure 10 shows a detectable anomalous point corresponding to batch 8.

The corresponding Q chart, shown in Figure 11, identifies the anomaly, because the eighth point lies beyond UCL. Therefore, although Q charts are theoretically inefficient at detecting shifts, they may not be so poor for practical use. Moreover, an SCC Q chart was not drawn in this case, because the data for type 9 are uncorrelated.

6. **Conclusions.** Although Q charts for short-run data and SCC Q charts for short-run autocorrelated data are well known, the effectiveness of these charts has not yet been sufficiently investigated from a practical perspective using real data, because all previous studies theoretically consider the univariate methods for short-run and short-run autocorrelated data.

Our study focused on both Q charts and SCC Q charts and practically assessed their effectiveness by using film thickness data obtained from a horizontal LPCVD process. The results are as follows: (1) Q charts enable data of different types to be plotted on the same chart; (2) SCC Q charts are more suitable for correlated data, because Q charts can contain type I errors; and (3) Q charts can detect real anomalies in data, although it is known that their outlier detection capability is poor in theory. On the basis of these results, we conclude that both methods are useful tools for controlling semiconductor manufacturing processes.

We have not compared Q charts and SCC Q charts with other methods of treating short-run or short-run autocorrelated data. The relative complexity of these other methods can make them inconvenient for practical use.

Acknowledgment. The authors greatly acknowledge the Renesas Electronics Corporation Group for their constructive comments and provision of data. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation. This work was supported by KAKENHI (25750120).

REFERENCES

[1] S. V. Crowder, An SPC model for short production runs: Minimizing expected cost, *Technometrics*, vol.34, no.1, pp.64-73, 1992.
[2] D. D. Montgomery, *Introduction to Statistical Quality Control*, John Wiley and Sons, New York, 2009.
[3] C. P. Quesenberry, The effect of sample size on estimated limits for X-bar and X control charts, *J. Qual. Technol.*, vol.25, no.4, pp.237-247, 1993.

[4] C. M. Wright, M. Y. Hu and D. E. Booth, Effectiveness of joint estimation when the outlier is the last observation in an autocorrelated short time series, *Decision Sci.*, vol.30, no.3, pp.825-847, 1999.
[5] C. H. Yang and F. S. Hillier, Mean and variance control chart limits based on small number of subgroups, *J. Qual. Technol.*, vol.2, no.1, pp.9-16, 1970.
[6] N. R. Farnun, Control charts for short runs: Nonconstant process and measurement error, *J. Qual. Technol.*, vol.24, no.3, pp.138-144, 1992.
[7] C. P. Quesenberry, SPC Q chart for start-up processes and short or long runs, *J. Qual. Technol.*, vol.23, no.3, pp.223-224, 1991.
[8] E. Del Castillo, J. M. Grayson, D. C. Montgomery and G. C. Runger, A review of statistical process control for short run manufacturing systems, *Comm. Stat. Theor. Meth.*, vol.25, no.11, pp.2723-2737, 1996.
[9] T. H. Harris and W. H. Ross, Statistical process control procedures for correlated observations, *Can. J. Chem. Eng.*, vol.69, no.48-57, 1991.
[10] L. C. Alwan and H. V. Roberts, The problem of misplaced control limits, *J. Roy Stat. Soc.*, vol.44, pp.269-306, 1995.
[11] L. VanBrackle and M. R. Reynolds, EWMA and CUSUM control charts in the presence of correlation, *Comm. Stat. Simulat. Comput.*, vol.26, no.4, pp.979-1008, 1997.
[12] A. V. Vasipoulos and A. P. Stamboulis, Modification of control chart limits in the presence of data correlation, *J. Qual. Technol.*, vol.10, no.1, pp.20-30, 1978.
[13] D. G. Wardell, H. Moskowitz and R. D. Plante, Control charts in the presence of data correlation, *Manage. Sci.*, vol.38, no.8, pp.1084-1105, 1992.
[14] C. W. Lu and M. R. Reynolds, CUSUM chart for monitoring an autocorrelated process, *J. Qual. Technol.*, vol.33, no.3, pp.316-334, 2001.
[15] C. M. Mastrangelo and D. C. Montgomery, SPC with correlated observations for the chemical and process industries, *Qual. Reliab. Eng. Int.*, vol.11, no.2, pp.79-89, 1995.
[16] S. Psarakis and G. Papaleonida, SPC procedures for monitoring autocorrelated processes, *Qual. Technol. Quant. Manag.*, vol.4, no.4, pp.501-540, 2007.
[17] A. Snoussi and M. Limam, The change point model: SPC method for short run autocorrelated data, *Qual. Technol. Quant. Manag.*, vol.4, no.3, pp.313-329, 2007.
[18] S. Prasad, Monitoring the quality of a chemical production process using the joint estimation method, *J. Chem. Inf. Comp. Sci.*, vol.35, no.1, pp.53-58, 1995.
[19] C. Chen and L. Liu, Joint estimation of model parameters and outlier effects in time series, *J. Am. Stat. Assoc.*, vol.88, no.421, pp.284-297, 1993.
[20] G. M. Wright, D. E. Booth and M. Y. Hu, Joint estimation: SPC method for short run autocorrelated data, *J. Qual. Technol.*, vol.33, no.3, pp.265-278, 2001.
[21] S. V. Crowder and L. L. Halblieb, Small sample properties of an adaptive filter applied to low volume SPC, *J. Qual. Technol.*, vol.33, no.1, pp.29-45, 2001.
[22] A. Snoussi, M. E. Ghourabi and M. Limam, On SPC for short run autocorrelated data, *Comm. Stat. Simulat. Comput.*, vol.34, pp.219-234, 2005.
[23] H. Kawamura, K. Nishina and T. Suzuki, Process adjustment control chart for simultaneous monitoring of process capability and state of statistical control, *Frontiers in Statistical Quality Control*, vol.10, pp.121-130, 2012.
[24] E. A. Del Castillo, A multivariate self-tuning controller for run-to-run process control under shift and trend disturbances, *IIE Trans.*, vol.28, no.12, pp.1011-1021, 1996.
[25] *Statistical Quality Control Handbook*, Western Electric Co., 1956.
[26] E. A. Youngs and E. M. Cramer, Some results relevant to choice of sum and sum-of-product algorithms, *Technometrics*, vol.13, no.3, pp.657-665, 1971.
[27] G. S. Wasserman and A. Sudjianto, Short run SPC based upon the second order dynamic linear model for trend detection, *Comm. Stat. Simulat. Comput.*, vol.22, pp.1011-1036, 1993.