Simultaneous Monitoring of Bias, Linearity, and Precision of Multiple Measurement Gauges

TEE-CHIN CHANG

University of Macau, Macao, China

In this paper, we consider a Shewhart charting scheme for simultaneous monitoring of multiple measurement gauges used for measuring a product or process attribute. The proposed chart is useful in detecting any malfunctioning gauges that cause a shift of bias, linearity, or precision. An in-control condition of the charting scheme provides evidence that the gauges are functioning properly and thus consistent and accurate measurements from the gauges are ensured. The run-length performance of the proposed charting scheme is compared with that of a Shewhart charting scheme for monitoring linearity between two measurement gauges. Applications of the proposed charting scheme are demonstrated with real manufacturing data sets.

Key Words: Average Run Length; Chi-Squared Distribution; *F* Distribution; Joint Probability Distribution; Noncentral Chi-Squared Distribution; Statistical Process Control.

In the manufacturing industry, process and product characterization, process control and improvement, and quality assurance are the core activities to ensure that high-quality products are consistently being produced. These activities rely greatly on the measurements of product quality or process parameters for statistical inferences. The quality of measured data depends on the performance of the measurement gauge, which is characterized by a measure of bias, accuracy, and precision. A malfunctioning gauge will cause a shift of either bias, linearity, or precision and thus result in inconsistent and inaccurate data being generated. Inconsistent and inaccurate data will lead to a biased statistical inference. The stability of bias, precision, and linearity is critical for obtaining consistent and accurate data all the time.

Precision of a gauge will determine the level of consistency of data and can be evaluated from a gauge repeatability and reproducibility (GR&R) study. Various statistical methods, such as those by Montgomery (2001), Montgomery and Runger (1993a, 1993b) and Burdick, Borror, and Montgomery (2003), have been proposed for conducting a GR&R study. Bias is often evaluated by comparing the average measurement of a reference standard with the known measurement value of the corresponding reference standard. Linearity measures the slope of a linear regression obtained based on the paired data of the measured and known measurement values of reference standards. The definition of linearity can also be found in the reference manual of Measurement System Analysis (1995). Both the bias and linearity of a gauge will determine the level of accuracy of data. In the literature, there are various statistical tests for comparing the bias, precision, and linearity of two or more gauges. Pitman (1939) and Maloney and Rastogi (1970) considered a statistical test for comparing the precision of two gauges. Blackwood and Bradley (1991) compared two gauges using paired data to test equality of bias and precision of measurements. Grubbs (1973, 1983) and Christensen and Blackwood (1993) have developed statistical tests for bias and precision of multiple gauges. Tan and Iglewicz (1999) developed the confidence intervals and equivalence test for the linearity of two different measurement methods.

Dr. Chang is a Senior Statistician with GlaxoSmithKline, Singapore. His email address is tee-chin.t.chang@gsk.com.

Due to the fact that extensive process-control and quality-assurance activities arising from high-volume production, multiple gauges are often used in the manufacturing operation for measuring the same type of product quality or process parameters. These gauges must be statistically comparable to ensure that measured data with similar consistency and accuracy are being produced. In integrated circuits (ICs) manufacturing, multiple gauges are used in almost all the processes, for example, in the IC electrical testing and wire bonding. In the IC electrical testing, all the ICs are tested with respect to various electrical measurements using the automatic test equipments (testers). The testing process is to ensure that the ICs are able to function electrically before they are shipped to the customers. Many testers are used in this process in order to cope with the high-volume of production. In order to ensure consistency of measurements from the testers, statistical comparison in terms of bias, linearity, and precision is regularly performed among the testers.

Another process that uses multiple measurement gauges is the wire-bonding process. The function of this process is to bond an internal semiconductor die to the external leads using gold wires, which is critical for the electrical functionality of an IC. There are two tests involved in this process, which are known as pull test and shear test. A pull test involves pulling off the bonded wires and the respective pull strengths are measured using a pull gauge. Whereas a shear test is performed by shearing the ball-shaped connections at the bond pads located on the internal die and the respective shear strengths are measured using a shear gauge. The main purpose of both tests is to measure the bonding quality. A high pull- or shear-test measurement indicates a good bonding quality. The bonding quality depends very much on the capability of the wire-bonding machine. Due to the fact that many wire-bonding machines need to be monitored, multiple pull and shear gauges are used.

In the statistical process control (SPC) literature, control charts are developed mainly for monitoring the product quality or process parameters. Regular monitoring of a measure of bias, linearity, and precision of measurement gauges is critical for ensuring gauges are functioning properly over time and thus providing consistent and accurate measurements at all times. Consistent and accurate data are critical for obtaining unbiased statistical inferences: unbiased inferences on the process stability, process capability, and product quality, for example. Chang and Gan (2006) have proposed a Shewhart chart for monitoring the linearity between two gauges based on the slope estimator of a linear statistical model of two gauges. The chart proposed by Chang and Gan (2006) will issue an out-of-control signal if either one or both gauges linearity and precision parameters have shifted. In this paper, we consider a Shewhart chart for simultaneous monitoring of bias, linearity, and precision of multiple measurement gauges. This charting procedure is useful in detecting malfunctioning gauges that cause a shift of either bias, precision, or linearity. The second section contains a description of a linear statistical model with bias, linearity, and precision parameters for multiple gauges. In the third section, the proposed Shewhart charts and the associated monitoring statistics are discussed. The Shewhart chart for monitoring linearity proposed by Chang and Gan (2006) is briefly described as well. The formulas for determining the control limits of the proposed Shewhart charts is also provided in the third section. The design of the proposed Shewhart chart based on the real data sets from the IC manufacturing is demonstrated in the fourth section. This is followed by a study on the run-length performances of the proposed charts in the fifth section. In addition, comparison of the average run length (ARL) performances of the proposed charts and the charts considered by Chang and Gan (2006) is also performed in the fifth section. Finally, conclusions are given in the sixth section.

Linear Statistical Model for Multiple Gauges

A set of "standard devices" (or reference standards) with n different but known measurement values, $\{u_1, u_2, \ldots, u_n\}$, are considered for monitoring multiple measurement gauges. For example, a set of standard devices associated with a weighing machine would be a set of different weights of which the exact weights are known. Let x_{ij} be the observed measurement on the jth standard device with known measurement value u_j obtained from the ith gauge. For q number of independent gauges, the observed measurements on n standard devices obtained from each gauge can be expressed as

$$x_{ij} = \gamma_i + \theta_i u_j + \varepsilon_{ij}, \qquad i = 1, 2, \ldots, q \text{ and } j = 1, 2, \ldots, n, \tag{1}$$

where γ_i and θ_i are the bias and linearity parameters of the ith gauge, respectively. Both the parameters γ and θ jointly constitute the accuracy profile of a gauge. The quantity ε_{ij} is the measurement er-

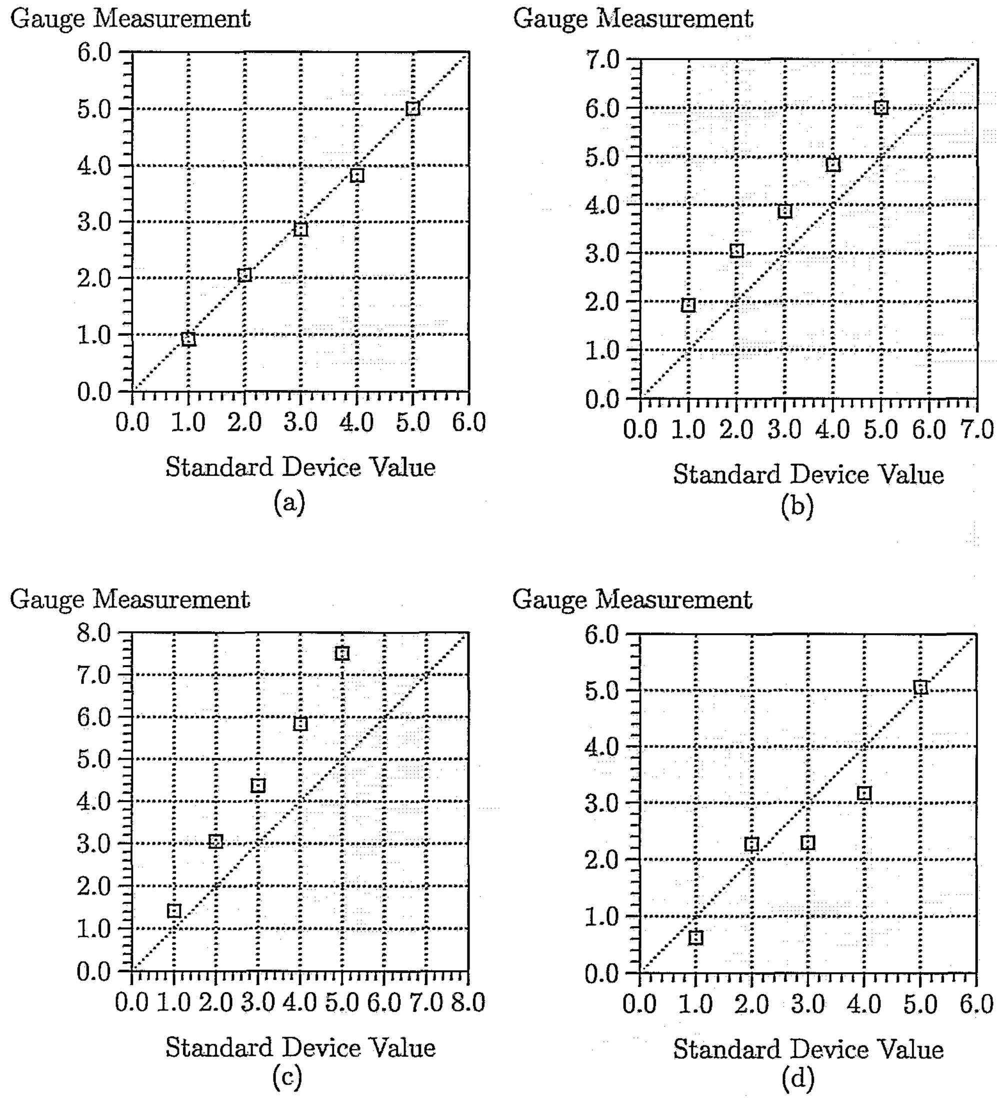

FIGURE 1. Plots of Linear Relationship of Gauge Measurements and Standard Devices Values with (a) $\gamma = 0$, $\theta = 1.0$, $\sigma = 0.1$ (the Measurement Gauge Is In Control with Respect to Bias, Linearity, and Precision), (b) $\gamma = 1.0$, $\theta = 1.0$, $\sigma = 0.1$ (the Measurement Gauge is Out of Control with Respect to Bias), (c) $\gamma = 0$, $\theta = 1.5$, $\sigma = 0.1$ (the Measurement Gauge Is Out of Control with Respect to Linearity), and (d) $\gamma = 0$, $\theta = 1.0$, $\sigma = 0.5$ (the Measurement Gauge Is Out of Control with Respect to Precision). The linear dotted line is the plot of ideal linear relationship of gauge measurements and standard devices values.

ror of the ith gauge on the jth standard device and $\varepsilon_{ij} \sim N(0, \sigma_i^2)$, where σ_i is the precision parameter of the ith gauge. The quantities ε_{ij}'s for $j = 1, 2, \ldots, n$ are assumed to be independent. When $\gamma_i = 0$ and $\theta_i = 1$ for $i = 1, 2, \ldots, q$, this indicates that the q gauges are functioning properly and are providing acceptable and identical measurements on average or they are simply having similar accuracy profiles. On the other hand, if $\sigma_1 = \sigma_2 = \cdots = \sigma_q$, then the q gauges are having similar precisions.

Figure 1 shows simulated examples of a linear relationship between the measurements obtained based on a set of standard devices with different values of γ, θ, and σ, and the corresponding known measurement values of the standard devices. Plot (a) of Figure 1 illustrates the in-control linear relationship or linear relationship when the gauge is functioning properly with $\gamma = 0$, $\theta = 1$, and $\sigma = 0.1$. The sample points are very close to the linear dotted line, which is the plot of the ideal linear relationship between the gauge measurements and the standard devices' known measurement values. Plots (b), (c), and (d) of Figure 1 show the out-of-control linear relationship or the linear relationship when there is a shift of γ, θ, or σ from its in-control value due to malfunction of the gauge. Sample points of plots (b) and (c) are consistently deviated from the ideal linear relationship plot. Although the sample points of plot (d) hover around the ideal linear relationship plot, they are not as close to the ideal linear relationship plot as those of plot (a). A detection of a shift of γ, θ, or σ is critical for detecting malfunction of a gauge.

Charting Schemes for Monitoring Multiple Gauges

The Proposed Charting Schemes

Measurement Precision σ Is Known

Let (x_{ij}, u_j) be the observed measurement on the jth standard device obtained using the ith gauge and the corresponding known measurement value of the jth standard device, respectively. A sample of n pairs of measurements, $\{(x_{i1}, u_1), (x_{i2}, u_2), \ldots, (x_{in}, u_n)\}$, based on n different standard devices is obtained from the ith gauge at time t, assuming that the precision of the ith gauge when the gauge is in control (or functioning properly) is known and is given as σ_i^0. The proposed statistic, which is the sum of squares of standardized measurement errors, for monitoring the bias, linearity, and precision parameters of the ith gauge at time t can be computed from

$$G_{it} = \sum_{j=1}^{n} \left(\frac{x_{ij} - u_j}{\sigma_i^0} \right)^2. \tag{2}$$

When the ith gauge is in control, we would expect that bias $\gamma_i = 0$ and linearity $\theta_i = 1$ and thus the distribution of the G_{it} statistic are a chi-square distribution with n degrees of freedom or $G_{it} \sim \chi_n^2$. The G statistics at time t for other gauges can also be obtained using Equation (2). If there is a shift of the ith gauge parameters such that $\gamma_i = \gamma_i^1$, $\theta_i = \theta_i^1$, and $\sigma_i = k_i \sigma_i^0$, where $\gamma_i^1 \neq 0$, $\theta_i^1 \neq 1$, and $k_i > 1$, due to malfunction of the gauge, then we would expect the G_{it} statistic value to be larger than the G_{it} statistic when the gauge is in control. The distribution of G_{it}/k_i^2 statistics, when the gauge is out of control, is a noncentral chi-square distribution with n degrees of freedom and noncentral parameter η_i, where

$$\eta_i = \sum_{j=1}^{n} \left(\frac{\gamma_i^1 + (\theta_i^1 - 1) u_j}{k_i \sigma_i^0} \right)^2 \tag{3}$$

or $G_{it}/k_i^2 \sim \chi_n^2(\eta_i)$ (see Appendix).

For simultaneous monitoring of q independent gauges, a Shewhart chart based on M_t statistic is given as

$$M_t = \text{Max}\{G_{it}, i = 1, 2, \ldots, q\}, \qquad t = 1, 2, \ldots; \tag{4}$$

the maximum value of the G statistics of the q gauges is considered. For ease of reference, the above chart is denoted as a Shewhart M chart. A signal is issued at time t if $M_t > UCL$, where UCL is the upper control limit. This indicates that at least one of the q gauges is out of control with respect to a shift of bias, linearity, or precision due to malfunction. In order to determine which gauge is malfunctioning and which parameters of the malfunctioning gauge have shifted, an exploratory plot of measurement errors and the corresponding known measurement values of standard devices can be considered. A measurement error is the deviation of the gauge measurement from the known measurement value of a standard device. If the measurement errors corresponding to all the standard devices are similarly high or low, then this indicates a possible shift of the bias parameter. However, if the measurement errors display a linear trend, then this indicates a possible shift of the linearity parameter. A fit of a linear regression on the measurement errors and performing significance tests on the intercept and slope will further verify whether a shift of bias or linearity or both has occurred. If the measurement errors vary randomly at zero level with high absolute values, then this might indicate a decrease of precision level. A significance test on the precision parameter will further verify any shift of precision. With the identification of the out-of-control parameter, a corresponding corrective action can then be carried out on the malfunctioning gauge.

The probability of false alarm, α, issued by a Shewhart M chart can be determined based on the following:

$$\begin{aligned} \alpha &= \text{Pb}(M_t > UCL \mid \text{the gauges are in-control}) \\ &= 1 - \text{Pb}(M_t < UCL \mid \text{the gauges are in-control}) \\ &= 1 - \{\text{Pb}(G_{it} < UCL)\}^q. \end{aligned} \tag{5}$$

Given a fixed value of α, the UCL can be determined based on the following:

$$UCL = \chi_{\zeta,n}^2, \tag{6}$$

where ζ is the level of significance, which is given as $(1-\alpha)^{q^{-1}}$.

Assuming that v gauges out of q gauges are out of control with respect to a shift of gauge parameters, the power of a Shewhart M chart π can be computed as

$$\begin{aligned}\pi &= \mathrm{Pb}(M_t > UCL \mid v \text{ gauges are out-of-control}) \\ &= 1 - \mathrm{Pb}(M_t < UCL \mid v \text{ gauges are out-of-control}) \\ &= 1 - \left\{ \{\mathrm{Pb}(G_{it} < UCL)\}^{q-v} \right. \\ &\qquad \left. \times \prod_{l=1}^{v} \mathrm{Pb}(G_{lt}/k_l^2 < UCL/k_l^2) \right\}, \qquad (7)\end{aligned}$$

where $G_{it} \sim \chi_n^2$ and $G_{lt}/k_l^2 \sim \chi_n^2(\eta_l)$. Run length denotes the number of samples taken until a signal is issued by a control chart. When the parameters are known, the run-length distribution of a Shewhart chart with the monitoring statistics are independent is a geometric distribution. The ARL and standard deviation of run length (SDRL) of a Shewhart M chart can be obtained based on the mean and standard deviation of a geometric distribution. The in-control ARL, out-of-control ARL, and SDRL are given as $1/\alpha$, $1/\pi$, and $\sqrt{\mathrm{ARL}(\mathrm{ARL}-1)}$, respectively.

Measurement Precision σ Is Unknown

Assume that the precision of the ith gauge when the gauge is in control, σ_i^0, is unknown. The in-control precision can be estimated from the m samples of n pairs of measurements $\{(x_{i1}, u_1), (x_{i2}, u_2), \ldots, (x_{in}, u_n)\}$, which are obtained from the ith gauge when the gauge is deemed in control. Assume $d_{jr} = x_{ij} - u_j$ is the difference between the observed measurement on the jth standard device obtained from the ith gauge and the jth standard device known measurement value of the rth sample. The precision of the ith gauge can be estimated based on

$$\hat{\sigma}_i = \sqrt{\frac{(S_1^2 + S_2^2 + \cdots + S_n^2)}{n}}, \qquad (8)$$

where

$$S_j^2 = \sqrt{\frac{1}{m-1}\sum_{r=1}^{m}(d_{jr} - \bar{d}_j)^2}$$

and

$$\bar{d}_j = \sum_{r=1}^{m} d_{jr}/m,$$

which is a pooled estimate obtained from the precision estimates associated with n different standard devices. The precision of other gauges can be estimated in a similar manner. The proposed statistics for monitoring bias, linearity, and precision parameters of the ith gauge at time t can be computed as

$$H_{it} = \sum_{j=1}^{n} \left(\frac{x_{ij} - u_j}{\hat{\sigma}_i} \right)^2. \qquad (9)$$

The distribution of H_{it}/n statistics is an F distribution with degrees of freedom n and $n(m-1)$ or $H_{it}/n \sim F_{n,n(m-1)}$ when the gauge is in control. If A and B are independent chi-squared variables with a and b degrees of freedom, respectively, then the variable $C = (A/a)/(B/b)$ is an F variable with a and b degrees of freedom (see Mood et al. (1974), p. 247). The distribution of H_{it}/n is derived by applying the above theorem on the variables G_{it} and $n(m-1)\hat{\sigma}_i^2/\sigma_i^{02}$, which are independent chi-squared variables with n and $n(m-1)$ degrees of freedom, respectively.

For simultaneous monitoring of q independent gauges, a Shewhart chart is based on the N_t statistic,

$$N_t = \mathrm{Max}\{H_{it}, i = 1, 2, \ldots, q\}, \qquad t = 1, 2, \ldots. \qquad (10)$$

The maximum value of the H statistics of the q gauges is considered. For ease of reference, the above chart is denoted as a Shewhart N chart. A signal is issued at time t if $N_t > UCL$, which indicates that at least one of the q gauges is out of control with respect to a shift of bias, linearity, or precision due to malfunction. The probability of false alarm, α, issued by a Shewhart N chart can be determined based on the following:

$$\begin{aligned}\alpha &= \mathrm{Pb}(N_t > UCL \mid \text{the gauges are in-control}) \\ &= 1 - \mathrm{Pb}(N_t < UCL \mid \text{the gauges are in-control}) \\ &= 1 - \{\mathrm{Pb}(H_{it} < UCL)\}^q. \qquad (11)\end{aligned}$$

Given a fixed value of α, the UCL can be determined based on the following:

$$UCL = nF_{\zeta,n,n(m-1)}, \qquad (12)$$

where ζ is the level of significance, which is given as $(1-\alpha)^{q^{-1}}$.

Shewhart $\hat{\beta}$ Charting Schemes

Chang and Gan (2006) have proposed a Shewhart chart for monitoring linearity between two gauges. Let (x_{1j}, x_{2j}) be the observed measurements on the jth standard device obtained using two gauges, respectively. Thus, a sample of n pairs of observed measurements $\{(x_{11}, x_{21}), (x_{12}, x_{22}), \ldots, (x_{1n}, x_{2n})\}$ based on n standard devices with different measurement values $\{u_1, u_2, \ldots, u_n\}$ can be obtained at a given time. The linear relationship between the explanatory variable, x_{1j}, and the dependent variable, x_{2j}, which is an error-in-variables model with slope β, measures the linearity between the two gauges and is considered by Chang and Gan (2006). A shift of β such that $\beta \neq 1$ indicates a shift of linearity between two gauges. Based on the models in Equation (1) for x_{1j} and x_{2j}, $\beta = \theta_2/\theta_1$ can be derived. Except for a simultaneous equal shift of θ_1 and θ_2, a shift of θ_1 or/and θ_2 could lead to a shift of β. Maximum-likelihood estimator $\hat{\beta}$ based on (x_{1j}, x_{2j}) observations and the precision ratio of two gauges, $\delta = \sigma_2^2/\sigma_1^2$, is used to estimate β. The $\hat{\beta}$ is an unbiased estimator with its variance proportional to the precision of the two gauges, σ_1^2 and σ_2^2. The plot of $\hat{\beta}$'s together with the lower and upper control limits (UCL) can be used to detect a shift of linearity parameters, θ's, or precision parameters, σ's, of two gauges. However, the chart is not able to detect a shift of the bias parameters, γ's, of the two gauges. The design procedure for determining the Shewhart chart limits based on the approximate distribution of $\hat{\beta}$ can be found in Chang and Gan (2006).

Monitoring Multiple Gauges in IC Manufacturing

The consistency and accuracy of measurements given by two pull gauges, X_1 and X_2, are regularly compared based on the measurements obtained from four "standard weights" with values 10, 25, 50, and 100 g. A real sample of measurements based on the four standard weights, which are 9.9691, 24.946, 50.122, and 100.01 g, is obtained from the pull gauge X_1. Similarly, a real sample of measurements based on the four standard weights, which are 9.9472, 24.896, 49.948, and 99.995 g, is obtained from the pull gauge X_2. In a similar manner, 30 real-samples data are obtained from each of the pull gauges X_1 and X_2. The in-control precision of pull gauges X_1 and X_2 are unknown, and using Equation (8) and the 30-samples data, the estimated values 0.03126 and 0.04908 are obtained, respectively. With the estimated in-control precision, the H statistics based on the above observed measurements are determined as 19.29 and 6.78 for pull gauges X_1 and X_2, respectively, using Equation (9). Thus, the N statistic obtained from Equation (10) is 19.29. Similarly, the N statistics based on the 30-samples data of gauges X_1 and X_2 can be obtained and are displayed in Figure 2(a).

Assuming that a Shewhart chart based on N statistics for simultaneous monitoring of bias, linearity, and precision of two pull gauges is needed, the false-alarm rate $\alpha = 0.002$, which corresponds to an in-control ARL of 500, is desired. With $n = 4$, $m = 30$, $q = 2$, and $\alpha = 0.002$, the UCL of the Shewhart chart can be obtained as 19.835 using Equation (12). The plot of N statistics given in Figure 2(a) shows that the 25th sample is above the UCL, which indicates that at least one pull gauge is out of control. The precision of the pull gauges X_1 and X_2 are re-estimated without the 25th sample data, and the estimates are 0.03157 and 0.0402, respectively. The N statistics based on the re-estimated precision are computed and are plotted in Figure 2(b) together with the new UCL. The plot in Figure 2(b) shows that, in addition to the 25th sample, the 27th sample also issues an out-of-control signal. Following similar steps, Figure 2(c) is obtained where the N statistics are computed based on the precision of pull gauges X_1 and X_2, estimated as 0.03206 and 0.03761, respectively. The plot in Figure 2(c) shows that, in addition to the 25th and 27th sample, the 26th sample also issues an out-of-control signal. The precision of the pull gauges X_1 and X_2 are again re-estimated without the 25th, 26th, and 27th sample data, and the estimates are 0.03229 and 0.03459, respectively. The N statistics based on the re-estimated precision are computed and are plotted in Figure 2(d). The plot in Figure 2(d) shows that there is no additional out-of-control sample besides the 25th, 26th, and 27th samples.

An investigation of which pull gauge is accountable for the out-of-control samples and which parameters have shifted has been carried out. Plots of the measurement errors correspond to different standard devices based on the out-of-control samples (the 25th, 26th, and 27th samples) for both the pull gauges are given in Figures 3(a)–3(c) respectively. For comparison purposes, the plot of measurement errors for an in-control sample, the 28th sample, has been included in Figure 3(d). From the plots in Figures 3(a)–3(c), it seems that gauge X_2 has higher variation of measurement errors as compared with gauge X_1. The measurement errors based on the 28th sample for both gauges, as shown in Figure 3(d), are hovering around zero with similar variation. A statistical comparison between the precision of the out-of-control samples and the precision of the in-control samples at a 5% level of significance has further verified the deterioration of the precision of gauge X_2. No statistical evidence indicates a shift of precision of gauge X_1. Based on a linear regression fit and signif-

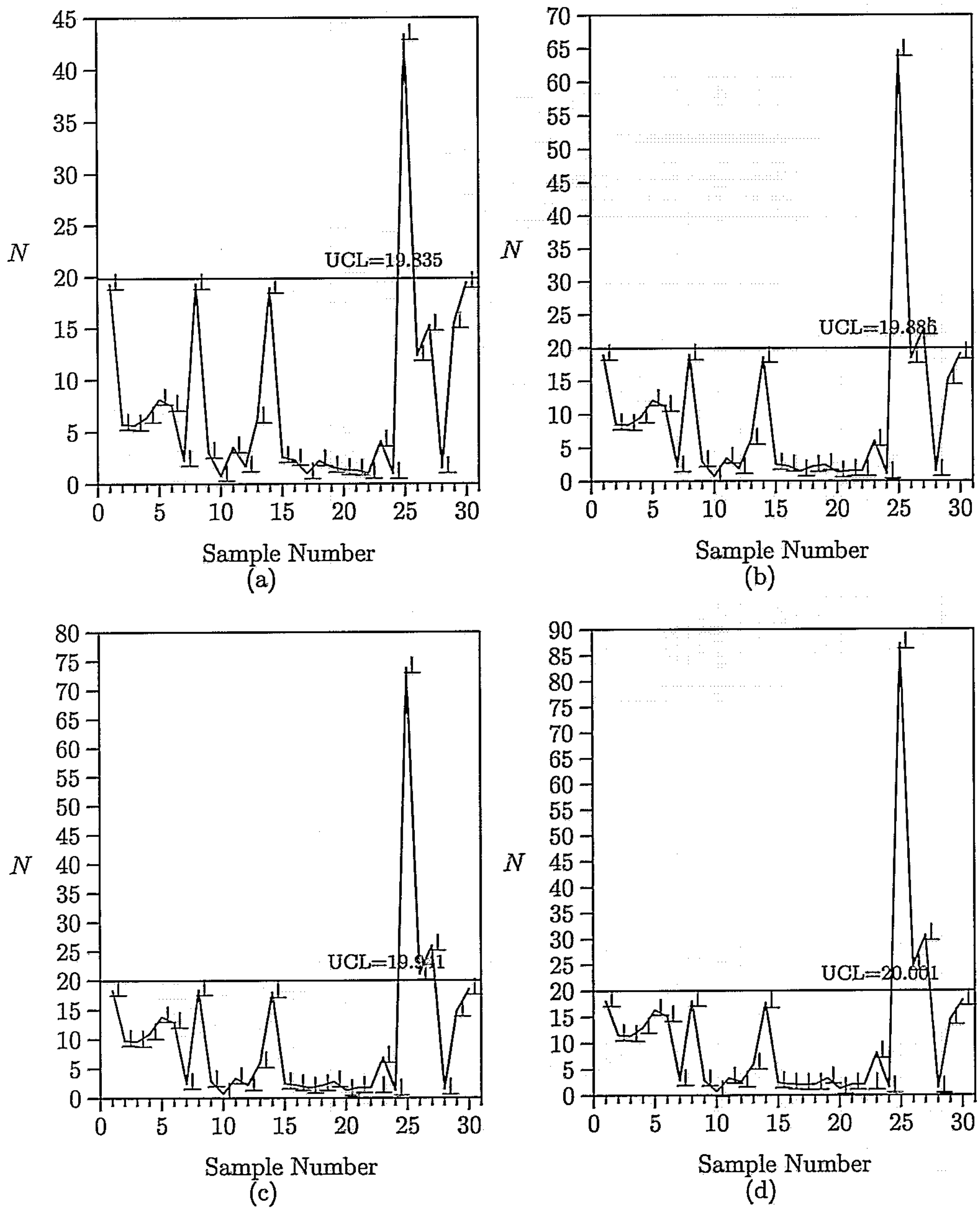

FIGURE 2. Shewhart Charts for Monitoring Bias, Linearity, and Precision of Two Pull Gauges. Upper control limit and N statistics are determined with precision estimated based on (a) all the samples, (b) all the samples except the 25th sample, (c) all the samples except the 25th and 27th samples, and (d) all the samples except the 25th, 26th, and 27th samples.

icance tests on the intercept and slope at a 5% level of significance, there is no statistical evidence to indicate any shift of the bias or the linearity of either gauge X_1 or gauge X_2. A scheduled maintenance was in fact performed on both the pull gauges between the time when the 27th and 28th samples were collected. This explains why both the pull gauges are within control after the 27th sample.

The Shewhart chart based on $\hat{\beta}$ proposed by Chang and Gan (2006) is applied on the same data set of pull gauges X_1 and X_2. The $\hat{\beta}$'s for the 30-sample data are obtained using the equation given in Chang and Gan (2006), with the estimated precision ratio δ. Figure 4(a) shows the plot of $\hat{\beta}$'s with δ estimated based on all the sample data. The control limits correspond to a false alarm rate $\alpha = 0.002$

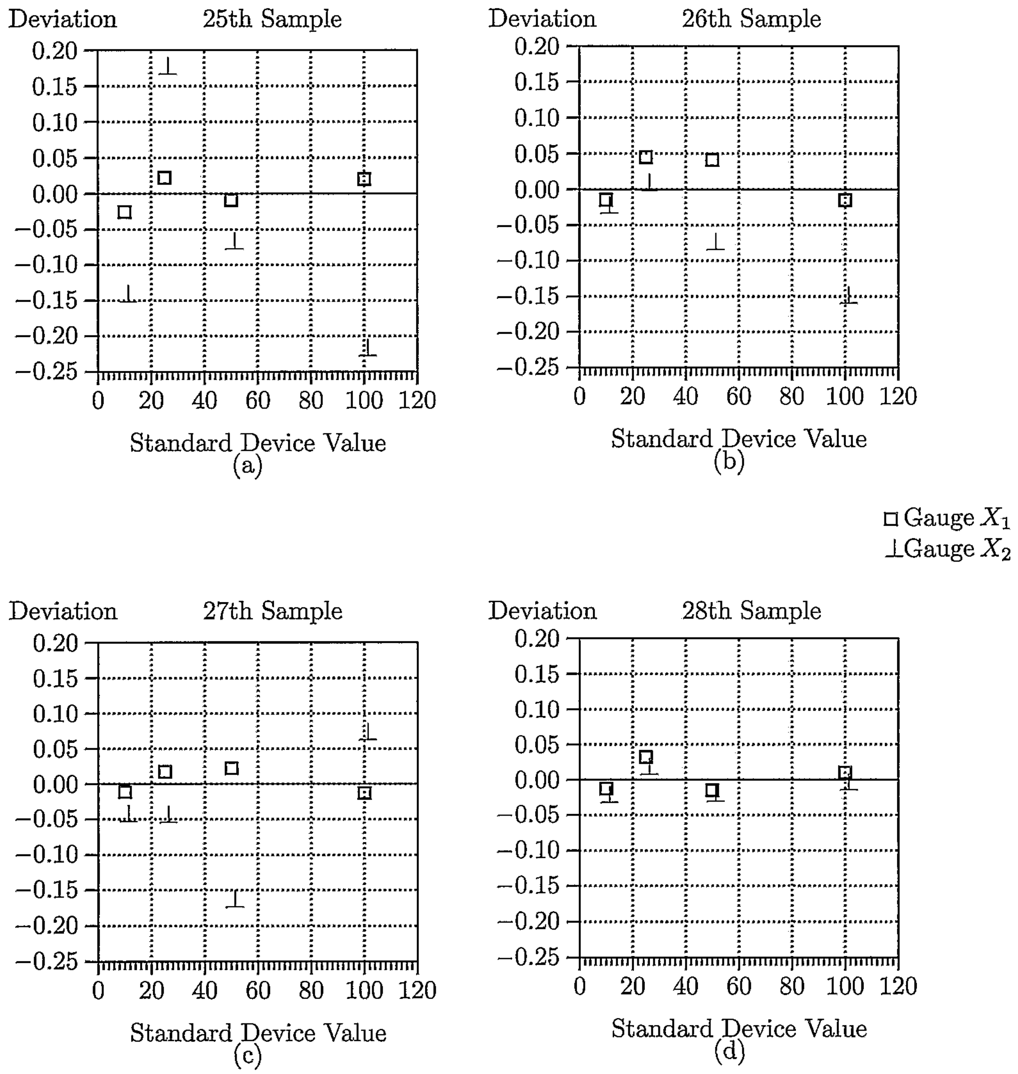

FIGURE 3. Plots of Deviations of the Gauge Measurements from the Standard Devices Values for the (a) 25th, (b) 26th, (c) 27th Samples, Which Are the Out-of-Control Samples in Figure 2, and the (d) 28th Sample, Which Is the In-Control Sample in Figure 2.

and are obtained using the design procedure given in Chang and Gan (2006). From Figure 4(a), a signal is issued at the 25th sample that indicates a possible shift of linearity or precision parameter of pull gauges X_1 and X_2 due to malfunction of at least one pull gauge. The δ is re-estimated with the precision of pull gauges X_1 and X_2 estimated without the 25th sample data. The $\hat{\beta}$'s are then estimated with the revised δ and are plotted in Figure 4(b). There is no additional out-of-control sample after the 25th sample. The Shewhart $\hat{\beta}$ chart only indicates that the 25th sample is out of control, in contrast with the proposed Shewhart N chart, which indicates that the 25th, 26th, and 27th samples are out of control. Although the 26th and 27th samples are within the control limits of Shewhart $\hat{\beta}$ chart, $\hat{\beta}$'s of both samples are relatively low and high, respectively, as compared with other $\hat{\beta}$'s of in-control samples. The run-length performances of both the Shewhart charts will be further evaluated and compared in the next section.

Another example of application of the proposed Shewhart chart is based on the real-data set obtained from two shear gauges, X_1 and X_2. The consistency and accuracy of the measurements given by the two shear gauges are regularly compared based on the measurements obtained from four standard weights, with values 25, 50, 100, and 250 g. Using an ap-

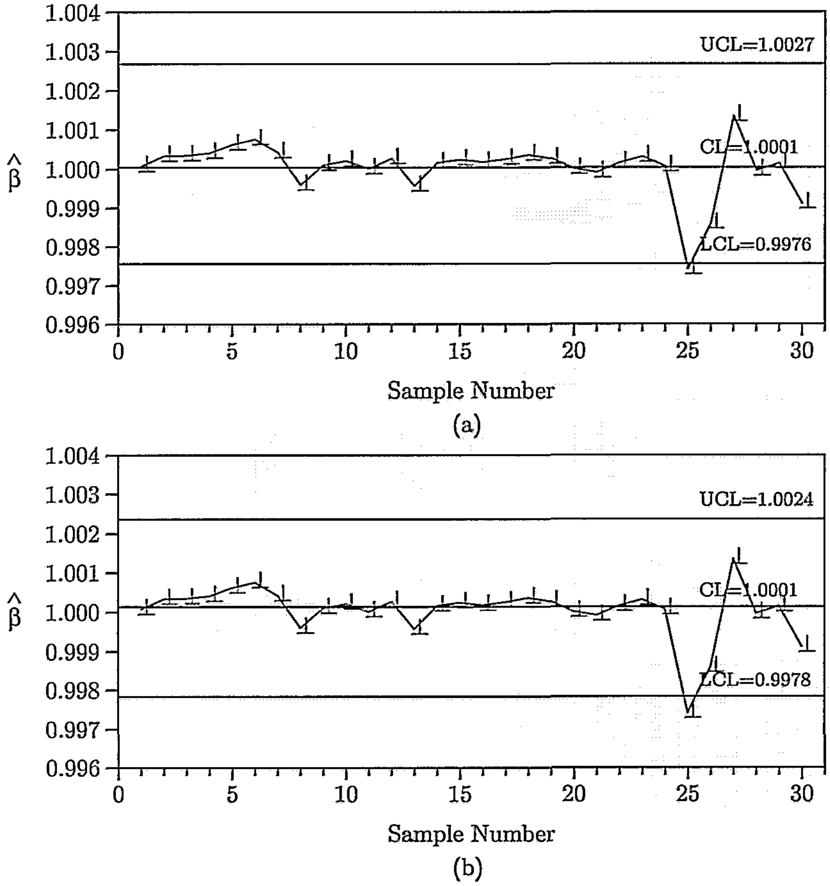

FIGURE 4. Shewhart $\widehat{\beta}$ Charts for Monitoring Linearity and Precision of Two Pull Gauges. Control limits and $\widehat{\beta}$ statistics are determined with precision estimated based on (a) all the samples and (b) all the samples except the 25th sample.

proach similar to that of the pull gauges example, a Shewhart chart based on N statistics for monitoring shear gauges is obtained and is shown in Figure 5. All the samples are in control, and this indicates that both shear gauges, X_1 and X_2, are in control in terms of bias, linearity, and precision or both the shear gauges are functioning properly within that period. The Shewhart $\hat{\beta}$ chart is applied on the same shear gauges data set and is given in Figure 6. All the $\hat{\beta}$'s are also within the control limits, which indicates that the shear gauges X_1 and X_2 are in control in terms of linearity and precision. This result is consistent with that concluded based on the proposed Shewhart N chart.

Run-Length Performances of Shewhart N Charts

Average run length (ARL) is often used to evaluate the sensitivity of a control chart in issuing an out-of-control signal. Besides, ARL is also used as a means of comparison for various competing charts. The precision σ of a gauge when the gauge is in control is usually unknown and needs to be estimated from the in-control samples. Thus, the use of Shewhart N charts would be more common in practice and therefore it would be more useful to know the run-length performance of a Shewhart N chart. The run-length performances of Shewhart M charts, however, can be referenced from the technical report generated by the author. Here, we study the run-length performances of the proposed Shewhart N charts for monitoring two measurement gauges based on various in-control samples. The exact ARLs of a Shewhart N chart with $m < \infty$ cannot be computed because the run length of a Shewhart N chart is not geometrically distributed due to the N statistics being dependent on the estimated in-control σ's. In addition to ARL, the estimated standard deviation of run length (SDRL) is also included in the comparison.

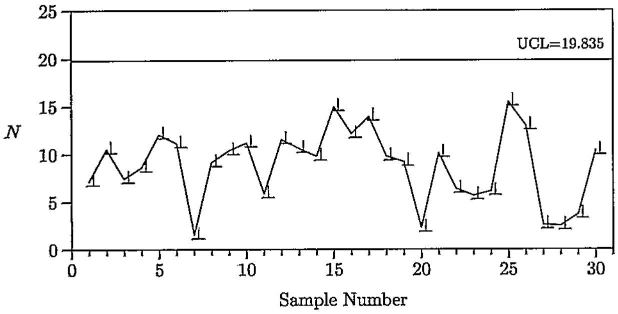

FIGURE 5. Shewhart Chart for Monitoring Bias, Linearity, and Precision of Two Shear Gauges. The upper control limit and N statistics are determined with the precision estimated based on all the samples.

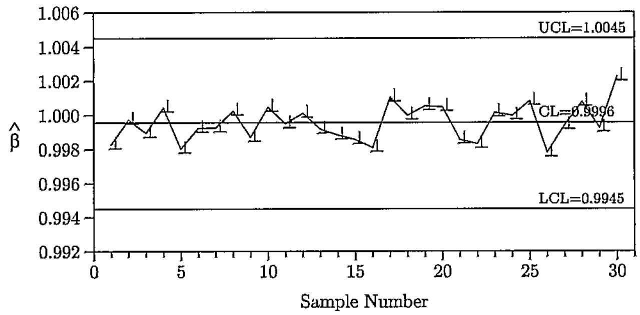

FIGURE 6. Shewhart $\widehat{\beta}$ Chart for Monitoring Linearity and Precision of Two Shear Gauges. The control limits and $\widehat{\beta}$ statistics are determined with the precision estimated based on all the samples.

The ARLs and SDRLs of Shewhart N charts correspond to various shifts of bias, linearity, and precision parameters of two gauges for number of samples $m = 30, 100, 200, 300$, and ∞ are provided in Tables 1–3, respectively. Four standard devices with known measurement values $u = 10, 25, 50, 100$ are considered. Except for the ARL and SDRL associated with $m = \infty$, each ARL and SDRL is estimated based on 30,000 simulation runs. In general, the ARL and SDRL of a Shewhart N chart approach the ARL and SDRL of a Shewhart M chart when m increases. It can be observed that the use of $m = 30$, a number that is commonly recommended in determining the Shewhart $\bar{X}$ chart limits for sample size $n = 4$ or 5, may not be adequate in terms of sensitivity in detecting out-of-control situations. For a Shewhart $\bar{X}$ chart used for monitoring process mean, Quesenberry (1993) has recommended about $m = 400/(n - 1)$ samples of size n in order for the $\bar{X}$ chart with estimated parameters to have performance similar to $\bar{X}$ chart with known parameters. This recommendation is found to be appropriate for the Shewhart N chart based on the run-length comparison results obtained from Tables 1–3.

The ARL performance of the Shewhart N chart with $m = \infty$, which is equivalent to a Shewhart M chart, will be compared with that of the Shewhart $\hat{\beta}$ chart. The ARL comparison, however, only considers the performance of both charts when there is a shift of either linearity θ or precision σ, as the Shewhart $\hat{\beta}$ chart is developed only to detect such a shift. Table 2 gives the ARLs of Shewhart M and $\hat{\beta}$ charts at various shifts of linearity parameters θ_1 and θ_2 with no change of bias and precision parameters. The ARL is 100 when θ_1 and θ_2 are in control at 1. In general, the Shewhart M charts are consistently more sensitive than the Shewhart $\hat{\beta}$ charts in detecting most of the out-of-control shifts of θ's. Both the Shewhart charts are more sensitive in issuing an out-of-control signal when there is a simultaneous shift of θ_1 and θ_2 as compared with a single shift of either θ_1 or θ_2, except for the Shewhart $\hat{\beta}$ chart in detecting a simultaneous equal shift of θ's. The Shewhart $\hat{\beta}$ chart is found to be comparatively very insensitive in issuing an out-of-control signal when there is a simultaneous equal shift of θ's. This is due to the fact that $\hat{\beta}$'s are close to the in-control value $\beta = 1$ when there is a simultaneous equal shift of θ's. For a fixed precision, the larger the simultaneous equal shift of θ's, the less sensitive is the Shewhart $\hat{\beta}$ chart. This is due to the variation of $\hat{\beta}$, of which the formula can be referenced to Chang and Gan (2006), being smaller when the simultaneous equal shift of θ's is larger.

Table 3 gives the ARLs of both the Shewhart M and $\hat{\beta}$ charts at various shifts of precision parameters σ_1 and σ_2 in terms of different values of k_1 and k_2, respectively, with no change of bias and linearity parameters. The ARL is 100 when k_1 and k_2 are in control at 1. It can be observed from the ARL values that both the Shewhart charts are more sensitive in issuing an out-of-control signal when there is a simultaneous shift of k_1 and k_2 as compared with a single shift of either k_1 or k_2. The Shewhart M chart is found to be substantially more sensitive in issuing an out-of-control signal than the Shewhart $\hat{\beta}$ chart for all the shifts of precision. Similar conclusions can be drawn from the run-length comparisons between the Shewhart N and Shewhart $\hat{\beta}$ charts based on standard devices with different set of known measurement values.

TABLE 1. ARLs* and SDRLs* (in Brackets) of Shewhart N Charts Based on Various Number of Samples m's with Respect to Various Shifts of Bias of Two Measurement Gauges γ_1 and γ_2. The chart limits are obtained based on in-control ARL of 100 with bias $\gamma_1 = \gamma_2 = 0$, linearity $\theta_1 = \theta_2 = 1$, and precision $\sigma_1 = \sigma_2 = 5$

	UCL	15.706	15.098	14.975	14.935	14.855
$\|\gamma_1\|$	$\|\gamma_2\|$	$m = 30$	$m = 100$	$m = 200$	$m = 300$	$m = \infty$
0.00	0.00	153.84	112.54	105.59	104.89	100.00
		(236.00)	(125.03)	(111.08)	(108.63)	(99.50)
	0.50	147.35	107.78	101.91	100.99	96.79
		(222.96)	(119.40)	(107.21)	(105.24)	(96.29)
	1.00	134.21	98.13	93.53	90.48	87.84
		(202.22)	(110.74)	(96.93)	(92.40)	(87.34)
	2.00	90.48	67.87	63.93	62.39	60.57
		(134.76)	(75.21)	(67.31)	(64.37)	(60.06)
	3.00	52.41	38.90	37.14	36.09	35.02
		(76.61)	(41.90)	(38.66)	(36.75)	(34.52)
	5.00	14.41	11.13	10.69	10.34	10.13
		(18.39)	(11.52)	(10.54)	(10.06)	(9.61)
0.50	0.50	142.77	105.63	98.86	96.81	93.79
		(211.43)	(118.24)	(103.69)	(99.62)	(93.28)
	1.00	129.34	95.77	89.86	89.13	85.36
		(195.04)	(107.73)	(94.14)	(91.89)	(84.86)
	2.00	88.83	65.21	62.47	61.85	59.38
		(126.72)	(70.83)	(65.45)	(63.59)	(58.88)
	3.00	50.65	38.60	36.69	35.82	34.63
		(69.93)	(42.09)	(38.06)	(36.17)	(34.13)
	5.00	14.06	11.16	10.58	10.48	10.09
		(17.66)	(11.28)	(10.42)	(10.17)	(9.58)
1.00	1.00	116.84	86.90	82.91	80.80	78.33
		(170.63)	(95.96)	(86.35)	(83.03)	(77.83)
	2.00	83.34	62.64	58.64	57.67	55.91
		(120.48)	69.49)	(60.54)	(59.07)	(55.41)
	3.00	49.07	37.30	35.17	34.54	33.43
		(68.34)	(40.76)	(35.99)	(35.17)	(32.93)
	5.00	13.87	10.91	10.35	10.33	10.00
		(17.09)	(11.22)	(10.17)	(9.98)	(9.48)
2.00	2.00	62.53	48.03	45.54	44.87	43.52
		(85.12)	(51.89)	(47.45)	(46.25)	(43.02)
	3.00	40.91	31.45	30.33	29.80	28.62
		(54.50)	(33.34)	(31.33)	(30.21)	(28.12)
	5.00	13.11	10.38	9.98	9.71	9.55
		(16.18)	(10.60)	(9.76)	(9.44)	(9.04)
3.00	3.00	29.60	23.32	22.36	22.10	21.39
		(37.88)	(24.11)	(22.66)	(21.95)	(20.88)
	5.00	11.29	9.29	8.93	8.90	8.65
		(13.29)	(9.25)	(8.59)	(8.58)	(8.13)
5.00	5.00	6.89	5.95	5.75	5.74	5.57
		(7.55)	(5.64)	(5.31)	(5.28)	(5.04)

*Except for the ARL and SDRL of the Shewhart N chart with $m = \infty$, all the ARLs and SDRLs are estimated using simulations with each value estimated based on 30,000 simulation runs.

TABLE 2. ARLs* and SDRLs* (in brackets) of Shewhart N and Shewhart $\hat{\beta}$ Charts Based on Various Number of Samples m's with Respect to Various Shifts of Linearity of Two Measurement Gauges θ_1 and θ_2. The chart limits are obtained based on in-control ARL of 100 with bias $\gamma_1 = \gamma_2 = 0$, linearity $\theta_1 = \theta_2 = 1$, and precision $\sigma_1 = \sigma_2 = 5$

	Chart	N	N	N	N	N	$\hat{\beta}$
	UCL	15.706	15.098	14.975	14.935	14.855	1.312
	LCL						0.762
θ_1	θ_2	$m = 30$	$m = 100$	$m = 200$	$m = 300$	$m = \infty$	
0.90	0.90	4.47	3.88	3.76	3.72	3.67	48.65
		(4.41)	(3.45)	(3.27)	(3.22)	(3.13)	(48.15)
	0.95	7.80	6.50	6.25	6.18	6.04	32.81
		(8.81)	(6.29)	(5.89)	(5.73)	(5.52)	(32.31)
	1.00	8.88	7.19	6.87	6.76	6.60	14.72
		(10.70)	(7.16)	(6.54)	(6.35)	(6.08)	(14.21)
	1.05	7.85	6.54	6.27	6.18	6.04	7.16
		(8.94)	(6.37)	(5.86)	(5.81)	(5.52)	(6.64)
	1.10	4.40	3.84	3.78	3.70	3.67	3.97
		(4.40)	(3.43)	(3.29)	(3.22)	(3.13)	(3.44)
0.95	0.95	32.14	25.52	24.43	23.85	23.41	69.22
		(39.99)	(26.45)	(24.53)	(23.85)	(22.90)	(68.72)
	1.00	57.23	42.28	40.23	39.60	37.74	44.83
		(80.20)	(45.83)	(41.61)	(40.32)	(37.24)	(44.33)
	1.05	31.79	25.80	24.51	24.00	23.41	19.06
		(39.49)	(26.87)	(25.08)	(23.96)	(22.90)	(18.55)
	1.10	7.88	6.48	6.28	6.20	6.04	8.85
		(9.11)	(6.23)	(5.93)	(5.72)	(5.52)	(8.33)
1.00	1.00	152.19	111.84	106.63	104.09	100.00	100.00
		(229.59)	(125.08)	(113.01)	(107.37)	(99.50)	(99.50)
	1.05	56.67	42.19	40.07	39.30	37.74	62.13
		(78.02)	(45.87)	(41.37)	(40.16)	(37.24)	(61.63)
	1.10	8.89	7.19	6.88	6.78	6.60	25.03
		(10.57)	(7.04)	(6.55)	(6.34)	(6.08)	(24.53)
1.05	1.05	32.59	25.57	24.41	24.05	23.41	146.73
		(41.70)	(27.29)	(24.77)	(24.22)	(22.90)	(146.23)
	1.10	7.80	6.46	6.23	6.23	6.04	87.33
		(8.66)	(6.27)	(5.89)	(5.85)	(5.52)	(86.83)
1.10	1.10	4.39	3.87	3.75	3.73	3.67	218.72
		(4.38)	(3.43)	(3.26)	(3.20)	(3.13)	(218.22)

*Except for the ARL and SDRL of the Shewhart N with $m = \infty$ and Shewhart $\hat{\beta}$ charts, all the ARLs and SDRLs are estimated using simulations with each value estimated based on 30,000 simulation runs.

Conclusions

Multiple measurement gauges are often used for measuring a type of product quality or process parameter due to extensive quality-assurance and process-control activities. Consistent and accurate measured data obtained from functioning gauges are critical for effective quality assurance and process control. A regular monitoring of the performance of measurement gauges in terms of bias, linearity, and precision will enable one to detect malfunctioning gauges, if any. In this paper, we have developed a Shewhart chart for simultaneous monitoring of multiple measurement gauges with respect to bias, linearity, and precision. Charting schemes for both cases where the in-control precision is known and unknown are considered. For the case of monitoring two

TABLE 3. ARLs* and SDRLs* (in Brackets) of Shewhart N and Shewhart $\hat{\beta}$ Charts Based on Various Number of Samples m's with Respect to Various Shifts of Precision of Two Measurement Gauges σ_1 and σ_2 in Terms of Different Values of k_1 and k_2. The chart limits are obtained based on in-control ARL of 100 with bias $\gamma_1 = \gamma_2 = 0$, linearity $\theta_1 = \theta_2 = 1$, and precision $\sigma_1 = \sigma_2 = 1$

	Chart	N	N	N	N	N	$\hat{\beta}$
	UCL	15.706	15.098	14.975	14.935	14.855	1.055
	LCL						0.948
k_1	k_2	$m = 30$	$m = 100$	$m = 200$	$m = 300$	$m = \infty$	
1.00	1.00	154.64	111.87	106.67	104.58	100.00	100.00
		(238.59)	(124.90)	(113.46)	(108.59)	(99.50)	(99.50)
	1.05	107.20	79.97	74.04	73.67	70.69	83.35
		(155.16)	(88.07)	(76.94)	(75.16)	(70.19)	(82.85)
	1.10	72.07	54.26	52.20	51.11	49.15	70.07
		(98.60)	(59.45)	(54.22)	(53.00)	(48.65)	(69.56)
	1.30	18.85	15.26	14.73	14.41	14.03	37.93
		(23.09)	(15.71)	(14.80)	(14.25)	(13.52)	(37.43)
	1.50	7.50	6.54	6.39	6.27	6.15	23.07
		(8.07)	(6.25)	(6.03)	(5.80)	(5.62)	(22.57)
	2.00	2.42	2.29	2.23	2.23	2.23	9.68
		(1.93)	(1.75)	(1.66)	(1.68)	(1.65)	(9.17)
1.05	1.05	78.50	60.70	57.10	56.26	54.72	70.61
		(106.32)	(65.51)	(60.05)	(57.09)	(54.22)	(70.11)
	1.10	57.66	45.12	43.44	42.15	40.90	60.23
		(77.24)	(48.46)	(45.16)	(42.42)	(40.40)	(59.73)
	1.30	17.47	14.48	13.84	13.72	13.31	34.15
		(20.80)	(14.68)	(13.72)	(13.47)	(12.80)	(33.65)
	1.50	7.28	6.36	6.16	6.15	6.02	21.43
		(7.66)	(6.01)	(5.78)	(5.78)	(5.49)	(20.92)
	2.00	2.42	2.27	2.23	2.24	2.22	9.36
		(1.95)	(1.74)	(1.67)	(1.69)	(1.64)	(8.85)
1.10	1.10	44.88	35.63	33.99	33.45	32.70	52.08
		(57.86)	(37.60)	(34.70)	(34.14)	(32.20)	(51.58)
	1.30	15.88	13.33	12.94	12.70	12.35	30.83
		(18.24)	(13.56)	(12.72)	(12.31)	(11.84)	(30.32)
	1.50	7.07	6.16	6.00	5.92	5.83	19.91
		(7.57)	(5.79)	(5.54)	(5.46)	(5.31)	(19.41)
	2.00	2.40	2.25	2.24	2.22	2.20	9.05
		(1.93)	(1.71)	(1.69)	(1.64)	(1.62)	(8.53)
1.30	1.30	9.57	8.23	8.02	7.89	7.77	21.02
		(10.30)	(8.00)	(7.62)	(7.50)	(7.25)	(20.52)
	1.50	5.50	4.86	4.75	4.71	4.66	15.03
		(5.43)	(4.43)	(4.30)	(4.23)	(4.13)	(14.53)
	2.00	2.23	2.10	2.10	2.08	2.07	7.89
		(1.73)	(1.56)	(1.53)	(1.49)	(1.49)	(7.37)
1.50	1.50	3.94	3.53	3.49	3.47	3.43	11.63
		(3.69)	(3.03)	(3.00)	(2.97)	(2.88)	(11.12)
	2.00	2.03	1.92	1.90	1.89	1.87	6.89
		(1.49)	(1.33)	(1.29)	(1.31)	(1.28)	(6.37)
2.00	2.00	1.53	1.46	1.45	1.45	1.44	5.05
		(0.93)	(0.83)	(0.81)	(0.81)	(0.80)	(4.52)

* Except for the ARL and SDRL of the Shewhart N chart with $m = \infty$ and Shewhart $\hat{\beta}$ charts, all the ARLs and SDRLs are estimated using simulations with each value estimated based on 30,000 simulation runs

gauges, the proposed Shewhart chart outperforms the Shewhart $\hat{\beta}$ chart in terms of the run-length performances. The proposed Shewhart chart is found to be more sensitive in detecting out-of-control shifts of linearity and precision parameters, which the Shewhart $\hat{\beta}$ chart is developed to detect. For a Shewhart chart with the monitoring statistics obtained based on the estimated in-control precision σ, the number of samples required for estimating the in-control precision σ is at least $400/(n-1)$. The applications of the proposed Shewhart charts are illustrated with real semiconductor data sets. The new application of control charts considered here has further exemplified the diversity and usefulness of control charting in quality improvement.

Appendix
The Probability Distribution of G_{it}/k_i^2 When the Gauge Is Out of Control

If $Z_1, Z_2, \ldots, Z_\nu$ are ν independent standard normal variables and if $a_1, a_2, \ldots, a_\nu$ are constants with

$$V = \sum_{i=1}^{\nu}(Z_i + a_i)^2$$

and

$$\lambda = \sum_{i=1}^{\nu} a_i^2,$$

then V has a noncentral chi-squared distribution with ν degrees of freedom and noncentral parameter λ (see Johnson and Pearson (1969)). From Equation (1), given that the ith measurement gauge is out of control with $\gamma_i = \gamma_i^1$, $\theta_i = \theta_i^1$, and $\sigma_i = k_i\sigma_i^0$, then $x_{ij}, j = 1, 2, \ldots, n$ are normally and independently distributed with means $\gamma_i^1 + \theta_i^1 u_j$ and standard deviations $k_i\sigma_i^0$. The standard normal variable $z_{ij} = (x_{ij} - (\gamma_i^1 + \theta_i^1 u_j))/k_i\sigma_i^0$ can be obtained. The G_{it} in Equation (2) can be expressed as

$$G_{it} = \sum_{j=1}^{n} k_i^2 \left(\frac{x_{ij} - (\gamma_i^1 + \theta_i^1 u_j) + (\gamma_i^1 + \theta_i^1 u_j) - u_j}{k_i\sigma_i^0}\right)^2$$

$$= \sum_{j=1}^{n} k_i^2 \left(z_{ij} + \frac{\gamma_i^1 + (\theta_i^1 - 1)u_j}{k_i\sigma_i^0}\right)^2.$$

Thus,

$$\frac{G_{it}}{k_i^2} = \sum_{j=1}^{n}\left(z_{ij} + \frac{\gamma_i^1 + (\theta_i^1 - 1)u_j}{k_i\sigma_i^0}\right)^2$$

has a noncentral chi-squared distribution with degrees of freedom n and noncentral parameter

$$\eta = \sum_{j=1}^{n}\left(\frac{\gamma_i^1 + (\theta_i^1 - 1)u_j}{k_i\sigma_i^0}\right)^2.$$

References

BLACKWOOD, L. G. and BRADLEY, E. L. (1991). "An Omnibus Test for Comparing Two Measuring Devices". *Journal of Quality Technology* 23, pp. 12–16.

BURDICK, R. K.; BORROR, C. M.; and MONTGOMERY, D. C. (2003). "A Review of Methods for Measurement Systems Capability Analysis". *Journal of Quality Technology* 35, pp. 342–354.

CHANG, T. C. and GAN, F. F. (2006). "Monitoring Linearity of Measurement Gauges". *Journal of Statistical Computation and Simulation* 76, pp. 889–911.

CHRISTENSEN, R. and BLACKWOOD, L. G. (1993). "Test for Precision and Accuracy of Multiple Measuring Devices". *Technometrics* 35, pp. 411–420.

GRUBBS, F. E. (1973). "Errors of Measurements, Precision, Accuracy and the Statistical Comparison of Measuring Instruments". *Technometrics* 15, pp. 53–66.

GRUBBS, F. E. (1983). "Grubbs' Estimators (Precision and Accuracy of Measurements)". In *Encyclopedia of Statistical Sciences*, Vol. 3, eds. S. Kotz and N. L. Johnson. John Wiley, New York, NY, pp. 542–549.

JOHNSON, N. L. and PEARSON, E. S. (1969). "Tables of Percentage Points of Non-Central χ". *Biometrika* 56, pp. 255–272.

MALONEY, C. J. and RASTOGI, S. C. (1970). "Significance Test for Grubbs' Estimators". *Biometrics* 26, pp. 671–676.

(1995). *Measurement System Analysis Reference Manual*, 2nd ed. Chrysler Corp., Ford Motor Corp., General Motors Corp.

MOOD, A. M.; GRAYBILL, F. A.; and BOES, D. C. (1974). *Introduction to Theory of Statistics*, 3rd ed. McGraw-Hill.

MONTGOMERY, D. C. (2001). *Introduction to Statistical Quality Control*, 4th ed. John Wiley & Sons, New York, NY.

MONTGOMERY, D. C. and RUNGER, G. C. (1993a). "Gauge Capability Analysis and Designed Experiments. Part I: Basic Methods". *Quality Engineering* 6, pp. 115–135.

MONTGOMERY, D. C. and RUNGER, G. C. (1993b). "Gauge Capability Analysis and Designed Experiments. Part II: Experimental Design Models and Variance Component Estimation". *Quality Engineering* 6, pp. 289–305.

PITMAN, E. J. G. (1939). "A Note of Normal Correlation". *Biometrika* 31, pp. 9–12.

QUESENBERRY, C. P. (1993). "The Effects of Sample Size on Estimated Limits for $\bar{X}$ and X Control Charts". *Journal of Quality Technology* 25, pp. 237–247.

TAN, C. Y. and IGLEWICZ, B. (1999). "Measurement-Methods Comparisons and Linear Statistical Relationship". *Technometrics* 41, pp. 192–201.