Quality Engineering
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597292

Monitoring Multiple Stream Processes

Xuyuan Liu [a]; R. Jock MacKay [b];Stefan H. Steiner [b]
[a] School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia [b] Business and Industrial Statistics Research Group, Department of Statistics and Actuarial Sciences, University of Waterloo, Waterloo, N2L 3G1, Canada

To cite this Article Liu, Xuyuan , MacKay, R. Jock andSteiner, Stefan H.(2008) 'Monitoring Multiple Stream Processes', Quality Engineering, 20: 3, 296 — 308
To link to this Article: DOI: 10.1080/08982110802035404
URL: http://dx.doi.org/10.1080/08982110802035404

Quality Engineering, 20:296–308, 2008

ISSN: 0898-2112 print/1532-4222 online
DOI: 10.1080/08982110802035404

Monitoring Multiple Stream Processes

Xuyuan Liu[1],
R. Jock MacKay[2],
Stefan H. Steiner[2]
[1]School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia
[2]Business and Industrial Statistics Research Group, Department of Statistics and Actuarial Sciences, University of Waterloo, Waterloo, N2L 3G1, Canada

ABSTRACT In industry, processes with multiple streams or gauges in parallel are common. We discuss monitoring such processes to detect changes in both the overall process mean and changes in the individual stream or gauge means. We propose two new control chart statistics based on an F test and a likelihood ratio test. One appealing aspect of these approaches is that they can be implemented either with or without process parameter estimates obtained from previous data (i.e., from phase 1 implementation of the control chart). These proposals are shown to compare favorably to available methods. The article is motivated by a truck assembly process in which wheel alignment characteristics are measured on every truck by one of four alignment machines, arranged in parallel within the overall process.

KEYWORDS control charts, measurement system monitoring, parallel gauges, parallel streams, statistical process control

INTRODUCTION

In the assembly of light trucks, front wheel alignment is a key characteristic that affects the handling of the vehicle and the life of its tires. The alignment characteristics include toe, caster, and camber angles on both the left and right front wheels. In the assembly process, there is a single production line. Due to high volumes, four gauges are used to measure alignment characteristics. Each truck is measured by one gauge (100% inspection) and adjustments are made as necessary. Trucks are assigned to a gauge based on gauge availability; accordingly, we assume that the distributions of the true values of the alignment characteristics are the same for each gauge. Checking a truck's alignment takes three to four minutes and all adjustment/rework is done offline. A process map is given in Figure 1.

The need for formal process monitoring became apparent during a project to reduce caster angle variation. At one point, the project team noticed that there were substantial differences in the average right caster angles among the four gauges.

The truck alignment process is an example of a multiple stream process, where the multiple streams correspond to the four gauges. In general, we may have multiple streams at any process step, a common feature of high-volume manufacturing.

To monitor a multiple stream process, it is helpful if we distinguish between assignable causes that result in changes to the overall process output (i.e., across all streams/gauges) and changes that only affect one

Address correspondence to Dr. Stefan H. Steiner, Department of Statistics and Actuarial Sciences, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada. E-mail: shsteiner@uwaterloo.ca

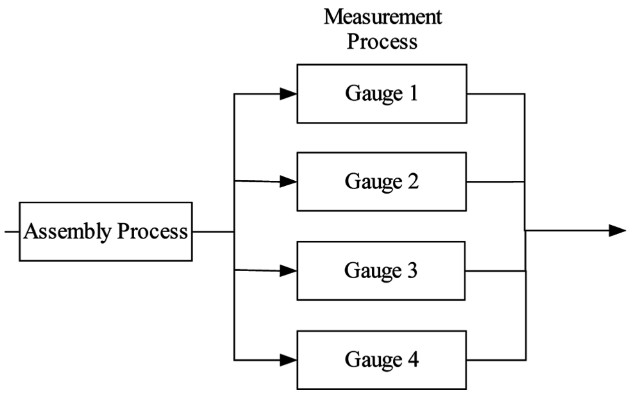

FIGURE 1 **Truck assembly and alignment measurement process with four parallel gauges.**

(or some) of the streams. Figure 2 illustrates some possible process changes using a hypothetical process with three parallel streams. The top left plot shows the process output stratified by stream when the process is in-control. We see that there is little difference in the distribution of the output values across the streams. The top right plot shows the process after a process mean shift upwards that affects all streams, while the plot in the bottom left shows the process after an increase in the variation of stream three. Finally, the bottom right plot shows the process after the mean of the third stream has decreased. Our goal is to distinguish among these types of process changes. Making this distinction is helpful in diagnosing the problem and eliminating the assignable cause. For example, a signal suggesting a shift in the mean of a single stream would require different corrective action than one suggesting a shift in the overall process mean.

If we ignore the streams, we can monitor the process using a single $\bar{X}$ and s chart (Montgomery, 2004). This is effective for detecting step changes in the overall process mean or variation. However, note that changes in the stream means (bottom right in Figure 2) will increase the overall process variation, and detecting changes in the output variation in an individual stream (bottom left) will be difficult. One immediate solution to this problem is to maintain separate $\bar{X}$ and s charts for each stream of the process. This, however, results in a large number of charts and is only feasible if the number of streams is small. In addition, using multiple $\bar{X}$ and s charts is less effective at detecting changes to the process that affect all streams at the same time. As well, with

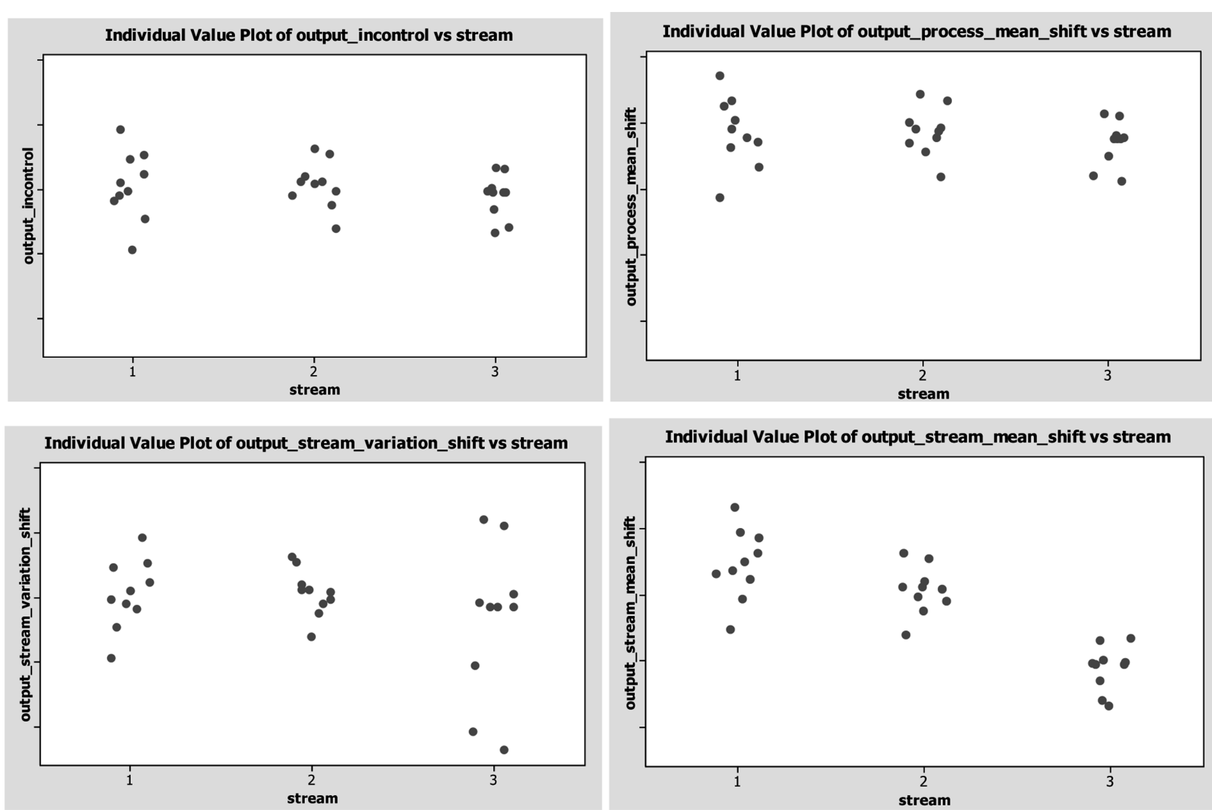

FIGURE 2 **Illustrative example plots of output by stream.**

multiple charts, we should widen the control limits of the stream-specific charts to maintain a reasonable overall false alarm rate.

We propose a number of new control charts, to be used in conjunction with the overall $\bar{X}$ and s charts, that are specifically designed to look for changes in the stream means. While we focus on detecting changes in the stream means, the ideas can be extended to look for changes in the within-stream variation. However, generally for multiple stream processes (Mortell and Runger, 1995), and especially when the streams represent multiple gauges, the process variation (e.g., assembly variation) is large relative to the within-stream variation (i.e., assembly plus measurement variation). In this case, changes in the variation of a single stream are difficult to detect and have little impact on the overall variation.

The next section provides some background and discusses existing approaches. The subsequent section gives a detailed description of the proposed new control charts and the derivation of appropriate control limits. Next, we compare the performance of the proposed control charts to some existing charts using a simulation study. Finally, we illustrate the use of the new control charts in the truck alignment example.

BACKGROUND

To monitor a multiple stream process, Boyd (1950; see also Nelson, 1986; Montgomery, 2004, section 9.3) proposed a group control chart. In this approach, observations are collected at regular time intervals from the process in subgroups consisting of n observations from each of the m streams. All the control charts considered here use the same sampling plan. We define the observed data for each subgroup as

$$y_{ij},\ i = 1, \ldots, m(\text{streams}),\ j = 1, \ldots, n(\text{observations})$$

For a group control chart the two statistics charted for each subgroup are the maximum and minimum of the stream averages; that is, $\max_i \bar{y}_{i.} = \text{maximum}(\bar{y}_{1.}, \bar{y}_{2.}, \ldots, \bar{y}_{m.})$ and $\min_i \bar{y}_{i.}$, where $\bar{y}_{i.} = \sum_{j=1}^{n} y_{ij}/n$. The control limits are derived using the standard procedure for an $\bar{X}$ control chart using the data from each stream as a subgroup (Montgomery, 2004). With a group control chart, there are two decision rules. First, the control chart signals a shift in the overall process mean whenever the maximum or minimum stream average falls outside the control limits. Second, to detect a shift in the mean of an individual stream, the stream numbers associated with the maximum and minimum average are used as plotting symbols on the chart. A signal occurs whenever the average of a particular stream is an extreme in c consecutive samples. That is, we see the same plotting symbol for c consecutive subgroups.

Mortell and Runger (1995) discuss the group control chart, concluding that the methodology is not effective at detecting changes in the overall process mean when the overall process variation is large relative to the stream variation. This would be the case in most multiple gauge applications where the measurement variability and relative bias of the measurement systems tend to be small compared to the variation in the "true" values of the characteristic. They also conclude that the discreteness in the runs rule (since c has to be integer) is a major drawback of the group control chart, severely limiting the choice of in-control average run length. Also, they note that the runs rule works poorly if more than one stream mean shifts by a similar amount at the same time.

Using the same sampling scheme, Mortell and Runger (1995) propose new control charts to detect changes in the overall or stream means. They recommend a standard $\bar{X}$ control chart based on the overall subgroup average, that is, $\bar{y}_{..} = \sum_i \sum_j y_{ij}/nm$, to look for changes in the overall process average. This provides an improvement over the group control chart that uses the min and max of the stream averages because we increase the sample size in the averaging. Note, while not considered by Mortell and Runger, we could similarly monitor for changes in the overall process variation using an s chart based on the overall subgroup sample standard deviation; i.e., $\sqrt{\sum_{i,j}(y_{ij} - \bar{y}_{..})^2/nm - 1)}$. To detect a shift in the mean of an individual stream, they suggest plotting the range of the stream averages, that is, $r = \max_i \bar{y}_{i.} - \min_i \bar{y}_{i.}$, the difference of the two statistics plotted on the group control chart. To construct the control limits for a Shewhart chart based on r, they use standard methods for range charts applied to the ranges of the stream averages. The upper control limit (UCL) for the chart is set at $D_{.999}(\overline{R}/d_2)$, where $\overline{R}$ represents the average of the ranges across all subgroups in phase I. The constant $D_{.999}$ can be obtained from the table given by Harter (1960) and d_2 from the control chart constant tables in Montgomery (2004). We refer to the method proposed by Mortell and Runger as the R chart or the range method.

Chang and Gan (2006) considered the special case of monitoring the linearity of two parallel measurement systems. Their approach requires a different sampling plan, involving the repeated measurement of parts with known dimensions. It is not clear how to extend their approach when there are more than two parallel gauges or when there are multiple streams other than measurement systems. As a result, we do not consider their approach further here.

PROPOSED METHODOLOGY

We compare the Mortell and Runger (1995) range method to four new test statistics, two based on the well known F test and two based on a likelihood ratio test. For ease of presentation, we use terminology from the multiple gauge context though the methodology is applicable to any situation with multiple parallel process streams. We focus on the goal of detecting changes in the stream means; that is, in the measurement bias of one or more of the gauges.

To explain the new approaches, for each subgroup we assume the following model:

$$Y_{ij} = T_{ij} + R_{ij}, \quad i = 1, 2, \ldots, m, \ j = 1, 2, \ldots, n \quad (1)$$

In model (1), Y_{ij} represents the possible values for the jth measured value from the ith gauge, T_{ij} the true values of the characteristic, and R_{ij} the measurement errors. We assume the T_{ij} are independent and distributed $N(\mu, \sigma_p^2)$ and the R_{ij} are independent of the true values and each other and distributed $N(\beta_i, \sigma_m^2)$. In this model, μ and σ_p are properties of the process upstream of the gauges, while β_i denotes the bias of the ith gauge and σ_m represents the measurement variation. From the model, we have $Y_{ij} \sim N(\mu + \beta_i, \sigma^2)$, where $\sigma^2 = \sigma_p^2 + \sigma_m^2$. We call σ^2 the overall variation as it represents the combined effect of the process and measurement variation.

The assembly process is in control if μ and σ_p do not change over time. The two objectives of the process monitoring are to detect changes in the (production) process parameters μ and σ_p from their stable values and to detect differences among the gauge biases; i.e., the β_i s. For the first goal we propose using $\overline{X}$ and s charts based on the overall sample mean and standard deviation. We focus on the second goal. When the measurement system is in control, $\beta_i = \beta$ for all $i = 1, 2, \ldots, m$, i.e., for all gauges, thus β is the stable (in-control) value of the measurement bias, not necessarily equal to zero. Note that using the proposed sampling scheme, a simultaneous consistent change in the bias of all measurement gauges is indistinguishable from changes in the assembly process mean. Next, we describe the four proposed chart statistics for detecting changes in the within-gauge biases.

F Statistic

Monitoring the consistency of the gauge bias is equivalent to repeatedly testing the hypothesis

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_m = \beta \text{ versus}$$
$$H_a : \beta_i \text{ not all equal, } \ i = 1, \ldots, m.$$

This general alternative hypothesis suggests the well-known F-test for comparing means in a one-way analysis of variance (ANOVA) context. Specifically, for each subgroup we calculate the F ratio

$$f = \frac{n \sum_{i=1}^{m} (\bar{y}_{i.} - \bar{y}_{..})^2}{m - 1} \Bigg/ \frac{\sum_{i,j} (y_{ij} - \bar{y}_{i.})^2}{nm - m} \quad (2)$$

Assuming model (1), $f \sim F(m-1, nm-m)$ under the null hypothesis; i.e., when the measurement system is stable. If we observe a large value of f at time t, we conclude that the measurement system is out of control. That is, one or more of the gauge biases has shifted. We assign an upper control limit, using the probability method, at $UCL = F_{.999}(m-1, mn-m)$. Note that this control limit does not depend on any unknowns. Thus, for the F-ratio control chart, phase 1, where we collect data from the in-control process to set an appropriate control limit, is not necessary. We can immediately begin phase 2, where we monitor the process on an ongoing basis.

Note that with the assumed model (1), the denominator of the F ratio (2) is an internal estimate of σ^2 under both the null and alternate hypotheses. In the usual measurement system studies, where parts are repeatedly measured, the denominator gives an estimate for σ_m^2 under both the null and alternative hypotheses. Also, the F test is well known to be robust to the distributional assumptions.

Gauge Average S Chart

With the F ratio, we compare the numerator, a measure of variability in the gauge averages, to the denominator, a measure of the variability within gauges. The denominator provides a benchmark that allows us to determine whether the observed differences in gauge averages, as measured by the numerator, should be considered large or not.

However, in situations where we have observed the in-control process for some time, i.e., we have a substantial amount of phase 1 data, we can determine a prior estimate of σ^2. In that case, we could increase the sensitivity of the chart by no longer reestimating the overall variation from the data in each subgroup. This suggests charting the numerator of the F-ratio, namely $\sum_{i=1}^{m}(\bar{y}_{i.} - \bar{y}_{..})^2$. Under rescaling, this statistic is equivalent to the sample standard deviation of the gauge averages. For this reason we propose a chart based on s, the standard deviation of the gauge averages in each subgroup; i.e.,

$$s = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}(\bar{y}_{i.} - \bar{y}_{..})^2} \tag{3}$$

We can also motivate the S chart more directly by noticing that when monitoring a process for changes in variation, an S chart using the subgroup standard deviation is an alternative to an R chart based on the subgroup range. In our context, Mortell and Runger's (1995) R chart looks for changes in the gauge means, which will look like increased variation in the output. We expect the R chart to work well when a single gauge bias shifts. However, a chart based on s may be better when multiple gauges shift. Also, we may expect the S chart to be somewhat more sensitive than the R chart when the number of parallel gauges is large.

To set an upper control limit for the gauge average S chart, use the standard probability limit given by $UCL = \bar{s}/c_4\sqrt{\chi^2_{.999}/n-1}$, where $\bar{s}$ represents the average of the standard deviations across all subgroups in phase I, the constant $\chi^2_{.999}$ comes from the chi-square distribution, and c_4 from control chart constant tables in Montgomery (2004).

Likelihood Ratio Statistic

The F test that leads to the chart statistic f is designed to look for any differences among the gauge biases. In some situations, where we are able to give a more specific alternative hypothesis, a test based on a likelihood ratio will be more powerful. We can make the alternative hypothesis more specific since, in many production processes, it is reasonable to assume that the bias of only one gauge at a time changes. As a result, we propose a chart statistic that will signal quickly when there is a shift in the bias of only one gauge. This idea can extend to cases where we can make the alternative hypothesis even more specific; for instance, we may be willing to assume the direction of the expected change in bias. This is not done here because it did not make sense in the truck alignment example.

The chart statistic l arises from the likelihood ratio test of the null hypothesis against the specific alternative hypothesis corresponding to a bias shift in one unspecified gauge denoted k; i.e.,

$$H_a^*: \beta_i = \beta \text{ for } i = 1, 2, \ldots, m, i \neq k \text{ and } \beta_k \neq \beta.$$

For convenience, we derive the expression for the log-likelihood ratio in two steps. First, recall that the log-likelihood ratio statistic l is of the form $-2\log[L(\hat{\beta}_0, \hat{\sigma}_0)/L(\hat{\beta}_a, \hat{\sigma}_a)]$, where L is the likelihood function, $\hat{\beta}_0$ and $\hat{\sigma}_0$ are the maximum likelihood estimates under the null hypothesis, and $\hat{\beta}_a$, $\hat{\sigma}_a$ are the maximum likelihood estimates under the specified alternative H_a^*. For a specific gauge k the log-likelihood ratio is

$$l_k = nm\log\left\{\sum_{i,j}^{m,n}(y_{ij} - \bar{y}_{..})^2 \Big/ \left[\sum_{j=1}^{n}(y_{kj} - \bar{y}_{k.})^2 + \sum_{i\neq k,\, j=1}^{m,n}(y_{ij} - \bar{y}^*_{k.})^2\right]\right\}$$

where $\bar{y}^*_{k.}$ denotes the average of the observations from all gauges except the kth gauge. The numerator gives an estimate of the overall variation while the first term in the denominator estimates the variation within gauge k and the second term the variation within all the other gauges.

Since we do not know which gauge has shifted, we propose the following likelihood ratio test statistic for the hypothesis that the bias of exactly one gauge has shifted

$$l = \max_k l_k = \max_k nm \log\left\{\sum_{i,j}^{m,n} (y_{ij} - \bar{y}_{..})^2 \middle/ \left[\sum_{j=1}^{n} (y_{kj} - \bar{y}_{k.})^2 + \sum_{i \neq k,\, j=1}^{m,n} (y_{ij} - \bar{y}_{k.}^*)^2\right]\right\} \quad (4)$$

Large values of the statistic suggest that the process is out of control since there is evidence one gauge bias is different from the others. Since we use a specific alternative hypothesis, we expect this chart statistic to perform better than other methods for detecting a shift in the bias of a single gauge.

Likelihood Ratio Statistic with Known Variation

The likelihood ratio statistic with known variation is similar to the standard likelihood ratio statistic. The difference is that when computing the test statistic, we assume a prior estimate of σ^2 is available from phase 1. The resulting test statistic for testing against H_a^* for a specified k, is $q_k = \frac{nm}{m-1}(\bar{y}_{k.} - \bar{y}_{..})^2/\sigma^2$.

Then, the chart statistic q is defined by

$$q = \max_k \left[\frac{nm}{m-1}(\bar{y}_{k.} - \bar{y}_{..})^2/\sigma^2\right]. \quad (5)$$

When the process is in control, each of the q_ks follow a χ_1^2 distribution but they are not independent.

Setting the Control Limits for the Likelihood Ratio–Based Control Charts

Since it is difficult to determine the distributions of l and q, we use simulation to determine appropriate control limits for these likelihood ratio–based charts. In the simulation, we assume that the in-control measurement errors follow model (1) and, without loss of generality, have mean zero ($\beta = 0$) and standard deviation $\sigma_m = 1$. We have to choose the number of simulated subgroups to achieve sufficient precision for the control limits. As is well known, $Var(\hat{X}_p) \approx p(1-p)/nf^2(x_p)$, where $\hat{X}_p$ is the pth sample quantile, n is the sample size, and $f(x)$ is the density function of distribution. We select 10^7 subgroups, which gives a standard error of the estimated quantile less than 0.03.

TABLE 1 **Upper Control Limit for Likelihood Ratio Statistic (*l*) with Unknown Variance**

Number of gauges (*m*)	Number of observations per gauge (*n*)		
	6	12	20
2	13.55	12.08	11.55
4	14.95	14.10	13.83
12	16.05	15.75	15.63
24	17.09	16.91	16.87

False alarm rate = 0.001.

TABLE 2 **Critical Values for Likelihood Ratio Statistic (*q*) with Known Variance**

Number of gauges (*m*)	Number of observations per gauge (*n*)		
	6	12	20
2	0.90	0.45	0.27
4	1.67	0.84	0.50
12	2.36	1.18	0.71
24	2.68	1.34	0.80

Control limit = table entry*estimate for σ^2 from phase 1.
False alarm rate = 0.001.

We give the critical values for the likelihood ratio statistics l and q derived by simulation for a false alarm rate of 0.001 in Tables 1 and 2. The critical values from Table 1 can be used directly for setting the control limit for the likelihood ratio statistic l since the distribution does not depend on the (unknown) overall variation σ^2. The critical values for use with the likelihood ratio statistic q in Table 2 need to be multiplied by σ^2, the phase 1 estimate of the overall variation.

COMPARISONS

Table 3 summarizes the chart statistics for the various control charts we will compare. Note that changes from subgroup to subgroup in the overall process mean μ, either jumps or drifting, have no effect on any of the proposed control charts.

One major difference among the proposed charts is that the Mortell and Runger range chart R, the likelihood method based on q, and the gauge average S chart all require estimates of the in-control value of σ to construct control limits. To get an estimate of σ we need phase 1 as in the usual implementation of statistical process control charts. On the other hand, the charts using f and l re estimate the overall variability

TABLE 3 **Proposed Test Statistics to Detect Change in Bias**

Approach	Chart statistic
Mortell and Runger, *R* chart	$r = \max_i \bar{y}_{i.} - \min_i \bar{y}_{i.}$
F-test *f* chart	$f = \sum_{i=1}^{m} (\bar{y}_{i.} - \bar{y}_{..})^2 \Big/ \sum_{i,j} (y_{ij} - \bar{y}_{i.})^2$
Gauge average *S* chart (with known σ^2)	$s = \sum_{i=1}^{m} (\bar{y}_{i.} - \bar{y}_{..})^2$
Likelihood (*l*) ratio chart	$l = \max_k \left\{ \sum_{i,j}^{m,n} (y_{ij} - \bar{y}_{..})^2 \Big/ \left[\sum_{j=1}^{n} (y_{kj} - \bar{y}_{k.})^2 + \sum_{i \neq k, j=1}^{m,n} (y_{ij} - \bar{y}^*_{k.})^2 \right] \right\}$
Likelihood (*q*) ratio chart (known σ^2)	$q = \max_k (\bar{y}_{k.} - \bar{y}_{..})^2$

for each subgroup. A major advantage of the *f* and *l* charts is that they can be implemented immediately without first collecting phase 1 data. Assuming the phase 1 estimates are precise, we expect the *r*, *q*, and *s* charts to have a performance advantage over the *f* and *l* control charts. However, this advantage could easily become a disadvantage if the overall variation changes over time or is poorly estimated in phase 1. In cases where each subgroup is reasonably large, the *R*, *S*, and *q* charts based only on the gauge averages from each subgroup ignore potentially valuable information given by the within gauge differences. In the *f* and *l* charts we use this within-gauge variability to assess the magnitude of the differences in the gauge averages.

A second qualitative difference among the proposed charts is the nature of the underlying alternate hypothesis. For the *f*, *S* and *R* charts, we specify a very general alternate hypothesis, while the two likelihood ratio–based methods, *l* and *q* are based on a more specific alternate hypothesis. The control charts based on the more specific hypothesis are expected to yield improved performance when the process change matches that hypothesis but not work as well for other kinds of process changes.

We compare the charting procedures quantitatively using the power to detect various changes in the biases in the multiple gauge process. We choose control limits for each chart statistic to give a false alarm rate of roughly 0.001. The power of a chart statistic is the probability that a point on that control chart falls outside the control limit when the biases of the measurement system have changed. In this comparison, we estimate the power of each chart using a simulation with 100,000 trials. Figures 3 through 5 provide graphical comparisons of the power for the different chart statistics for 2, 4, and 24 (m) gauges and 6 and 20 (n) observations for each gauge. For each chart statistic, one gauge is shifted from a relative bias ($\beta_k - \beta$) of zero to a bias of 0.2, 0.4, 0.6,…, 3.0. We compare the unknown variance chart statistics (likelihood ratio statistic *l*, F statistic *f*) and the known variance chart statistics (likelihood ratio statistic *q*, *r* statistic, and *s* statistic) separately. Note that when there are only two gauges the Mortell and Runger *R* and Gauge Average *S* charts are equivalent.

Among the cases considered, and we believe more generally, if the process change matches the alternative hypothesis H_a^* for the unknown variance group, the likelihood ratio statistic performs consistently better than F statistic. In the known variance group, the chart based on the likelihood ratio statistic performs consistently better than R and S charts. These results are expected.

When the measurement system consists of only two parallel gauges, we see from Figure 3 that there is virtually no difference among the power of the chart statistics within each group. The known variance group performs noticeably better than the unknown variance group with six observations from each gauge. However, as the number of observations from each gauge increases, the difference between two groups becomes smaller. With 20 observations from each of the two gauges, the known variance methods are only marginally better than the unknown variance group.

As shown in Figures 4 and 5, the number of observations from each gauge plays a similar role when there are 4 and 24 gauges in the measurement system. Generally, as the number of observations increases, the difference between the known and unknown variance methods becomes smaller. This suggests that if we could obtain a large number of observations from the process in each subgroup, regardless of the number of gauges in the measurement system, the likelihood ratio statistic with unknown variance performs as well as all the methods in the known variance group.

Next, we consider the effect of the number of gauges. From Figures 3 through 5, we see that as

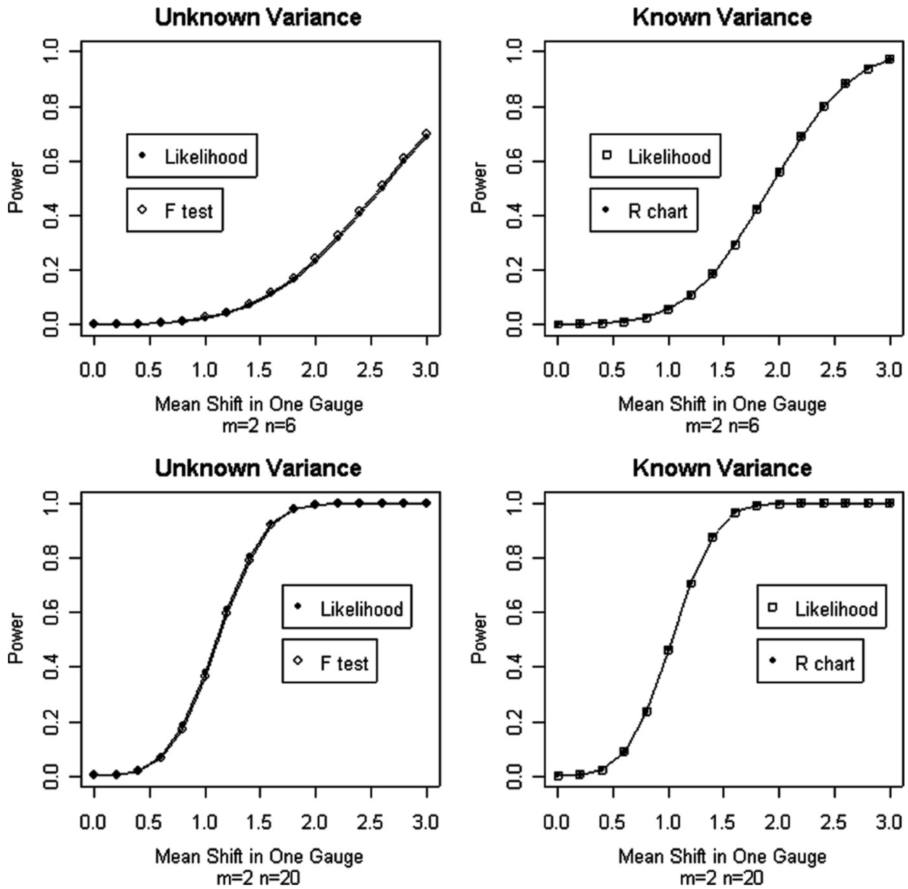

FIGURE 3 Comparison of the power with two gauges.

the number of gauges increases, the difference among chart statistics within each group increases. For example, clearly the difference between the likelihood ratio statistic with unknown variance and the F statistic becomes larger as the number of gauges increases. The pattern in the known variance group is similar. Also, increasing the number of gauges reduces the difference between the best chart within each group. Regardless of the sample size, with a large number of gauges in the measurement system, the likelihood ratio statistic with unknown variance performs (almost) as well as the likelihood ratio statistic with known variance.

The results in Figures 3 through 5 are summarized numerically in Table 4, which shows the in-control false alarm rates and the out-of-control power of the different methods under various conditions. We only give the power when the mean of one gauge is shifted from a relative bias of zero to a bias of 1, 2, and 3. Table 4 allows a direct comparison of the various methods in the two groups.

In the comparison, we have assumed the process variation is constant, and we assumed the constant process variation was known when we constructed control charts for three chart statistics in the known variance group. In application, as the variance is not known but estimated, the R, S, and q charts will not necessarily match the results given here. We may have a higher or lower false alarm rate and higher or lower power to detect bias changes than expected.

Another concern is the possibility that multiple gauge biases shift simultaneously. This is especially a concern for the likelihood-based approaches where the specific alternative hypothesis is that the bias of only one gauge shifts. It is not feasible to consider all possible combinations of bias shifts in the multiple gauges. To get an idea, we consider the case where two gauge biases shift simultaneously *in the same direction and by the same magnitude*. Figure 6 shows the comparisons of the power for 4 and 24 gauges and six observations from each gauge.

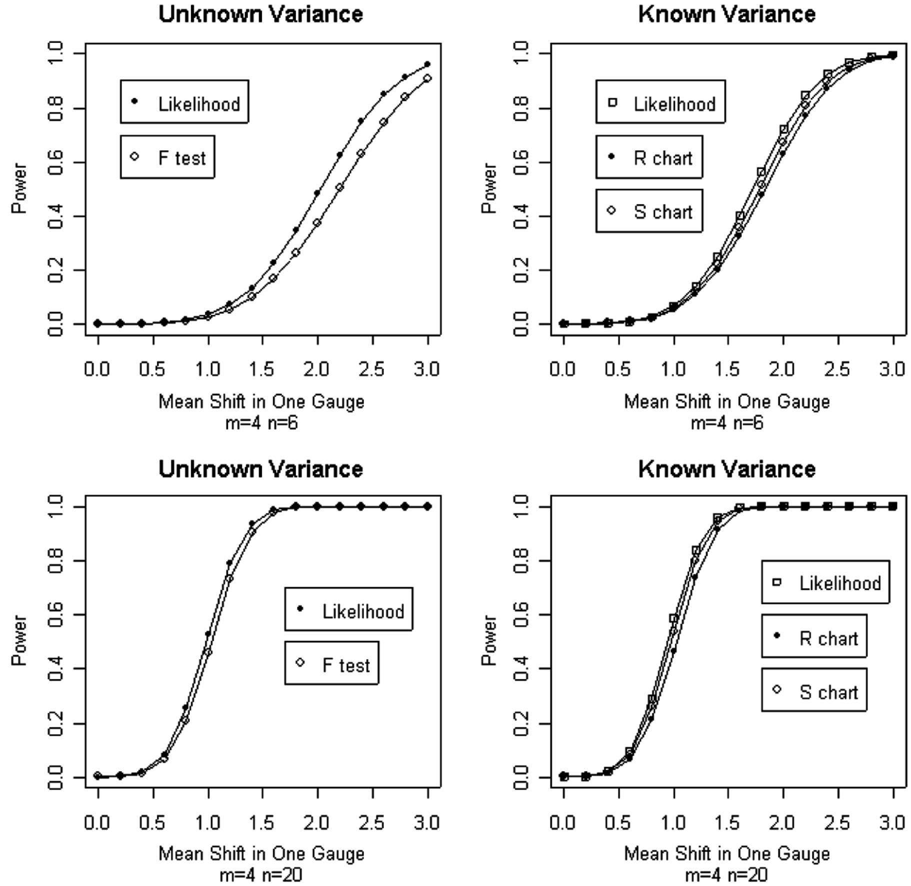

FIGURE 4 Comparison of the power with four gauges.

For two gauges a simultaneous shift in the same direction will look like a shift in the process mean so we skip this case in Figure 6. For a small number of gauges, the likelihood ratio chart with unknown variance shows extremely low power compared with other methods. Since the likelihood ratio chart statistic is designed to be sensitive to a single gauge shift, this is not surprising. However, this difference disappears as the number of gauges increases from 4 to 24.

EXAMPLE

Consider the truck assembly process discussed in the Introduction. As shown in Figure 1, there are four parallel alignment gauges. The goal is to detect changes in the gauge biases or changes in the overall process mean or standard deviation.

In this application, 100% inspection was used. We have data from a total of 2950 trucks over 10 days production. For monitoring purposes, we define a subgroup as 50 consecutive trucks. The choice of 50, as apposed to say 10 or 100, is arbitrary. Generally, we may only sample a well-defined subgroup from the production process at some predetermined frequency.

With subgroups of 50 trucks, since there are four gauges, on average we will have 12.5 trucks per gauge. In practice the number of trucks varies due to measurement times and production timing. The number of observations per gauge ranges from 5 to 22 across the 58 subgroups. This variability is not taken into account; i.e., we do not change the control limits based on the composition of the subgroup.

To look for changes in the overall process mean and variability we use an $\bar{X}$ chart of the overall subgroup averages and an s chart of the overall subgroup standard deviations. Note that these charts do not use the classification based on gauges. The resulting charts are given in Figure 7.

Figure 7 shows the result of applying the control charts retrospectively to the phase 1 data. We see evidence that the process is out of control in terms of its mean. This is perhaps not overly surprising

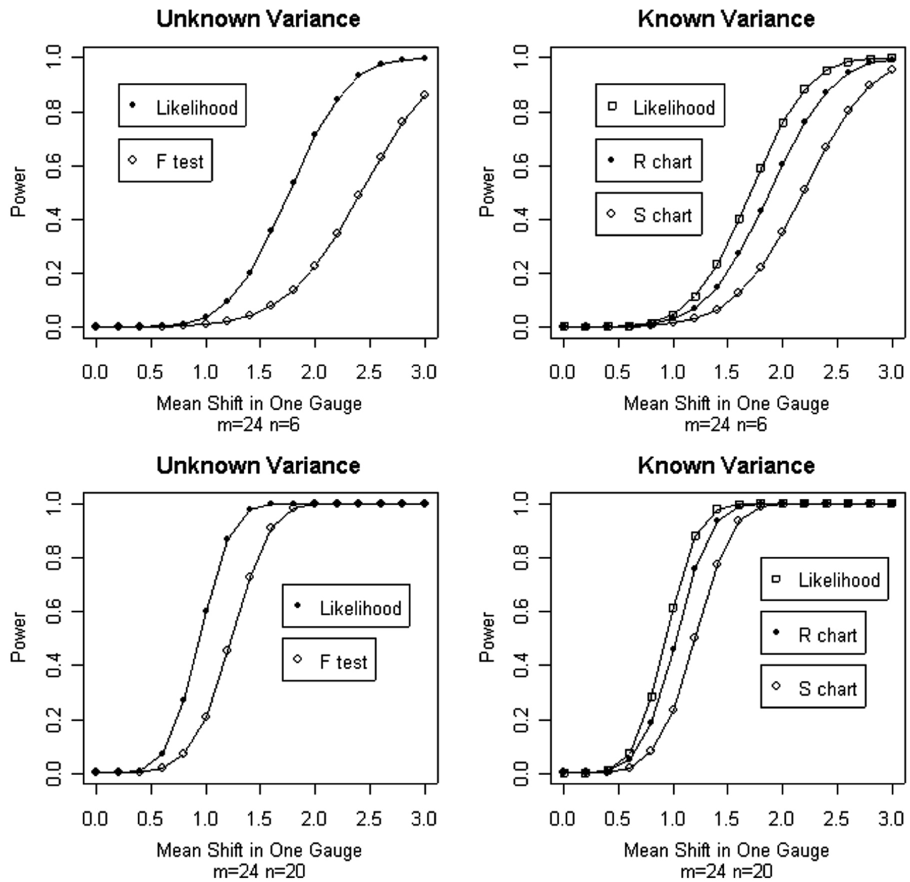

FIGURE 5 **Comparison of the power with 24 gauges.**

with the given definition of a subgroup. The variability within subgroups (50 consecutive vehicles) is smaller than the long-term variability in the right caster. This is due to some autocorrelation in the right caster. This will not affect the proposed "control charts" to detect changes in the gauges biases.

To look for changes in the gauge biases, we illustrate using the *R*, *S*, *l*, and *f* charts. The first two can only be applied retrospectively since to determine the control limits we need some phase 1 data. Redefining the available data as the subgroup gauge averages, i.e., for each subgroup, the data are $\bar{y}_{i.}$, $i = 1, 2, \ldots, m$ (m is the number of gauges), we can use standard software implementation of $\bar{X}$ and s charts to generate the *R* and *S* charts given in Figure 8. The results suggest that the process is in control.

The control charts based on both *l* and *f* can be applied without any phase 1 data. In other words, we can set appropriate control limits without any process data. The control limit for the *f* chart is set at $F_{.999}(4-1, 50-4) = 6.42$. From Table 1, with 4 gauges and 12 observations per gauge we get a control limit of 14.1 for the *l* chart. Figure 9 gives the two resulting control charts. In the *l* chart, we see an out-of-control signal at subgroup 57. However, the signal is not sustained.

For any of the charts, if there is evidence of a difference between the gauges we can compare the gauge means $\bar{y}_{i.}$ and/or the individual data stratified by gauge to determine which guage is likely the main cause of concern. With the likelihood-based methods we can also examine the l_k and q_k values for each gauge to see which is the largest. In the example, further investigation of the data in subgroup 57 gives Figure 10. We see that gauge 4 has a substantially lower average than the other three gauges. Looking further at subgroups 58 and 59, gauge 4 consistently gives the smallest right caster average. This suggests that gauge 4 may be experiencing problems.

CONCLUSIONS AND DISCUSSION

We have proposed a variety of different control charts for monitoring processes with multiple parallel gauges or streams. The most suitable method

TABLE 4 Comparison of Power

Chart type	Number of gauges (m)	Number of observations (n)	Power for a Mean Shift in One Gauge of b σ units			
			b = 0	b = 1	b = 2	b = 3
Likelihood (l)	2	6	0.00105	0.024	0.241	0.699
F	2	6	0.00112	0.005	0.073	0.362
R (Mortell et al.)	2	6	0.00116	0.062	0.580	0.972
Likelihood (l)	2	20	0.00118	0.395	0.901	0.998
F	2	20	0.00101	0.388	0.898	0.998
R (Mortell et al.)	2	20	0.00113	0.487	0.955	1
Likelihood (l)	4	6	0.00095	0.035	0.484	0.958
F	4	6	0.00118	0.028	0.375	0.906
R (Mortell et al.)	4	6	0.00121	0.055	0.630	0.990
S	4	6	0.00130	0.059	0.672	0.994
Likelihood (Known)	4	6	0.00133	0.067	0.721	0.997
Likelihood (l)	4	20	0.00091	0.254	0.999	1
F	4	20	0.00095	0.210	0.997	1
R (Mortell et al.)	4	20	0.00150	0.214	0.999	1
S	4	20	0.00205	0.258	1	1
Likelihood (Known)	4	20	0.00092	0.289	1	1
Likelihood (l)	24	6	0.00102	0.014	0.536	0.992
F	24	6	0.0011	0.005	0.139	0.761
R (Mortell et al.)	24	6	0.00101	0.012	0.432	0.979
S	24	6	0.00108	0.007	0.222	0.897
Likelihood (Known)	24	6	0.00109	0.016	0.590	0.995
Likelihood (l)	24	20	0.00124	0.273	1	1
F	24	20	0.00114	0.070	0.983	1
R (Mortell et al.)	24	20	0.00112	0.186	0.999	1
S	24	20	0.00108	0.080	0.990	1
Likelihood (Known)	24	20	0.00121	0.282	1	1

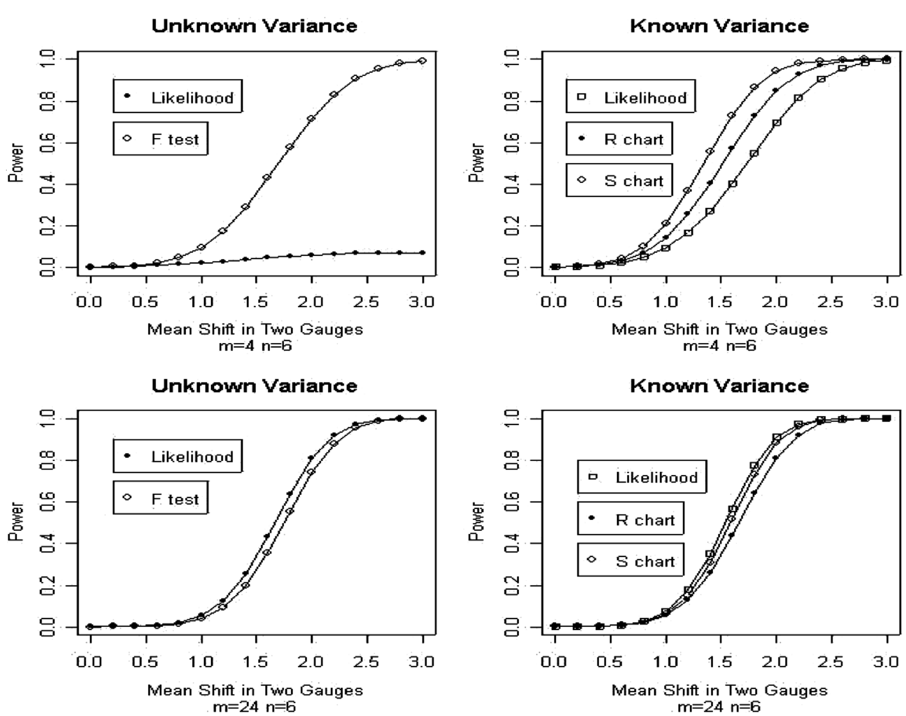

FIGURE 6 Comparison of the power with two gauge bias shifts, $m = \#$gauges, $n = \#$of observations per gauge.

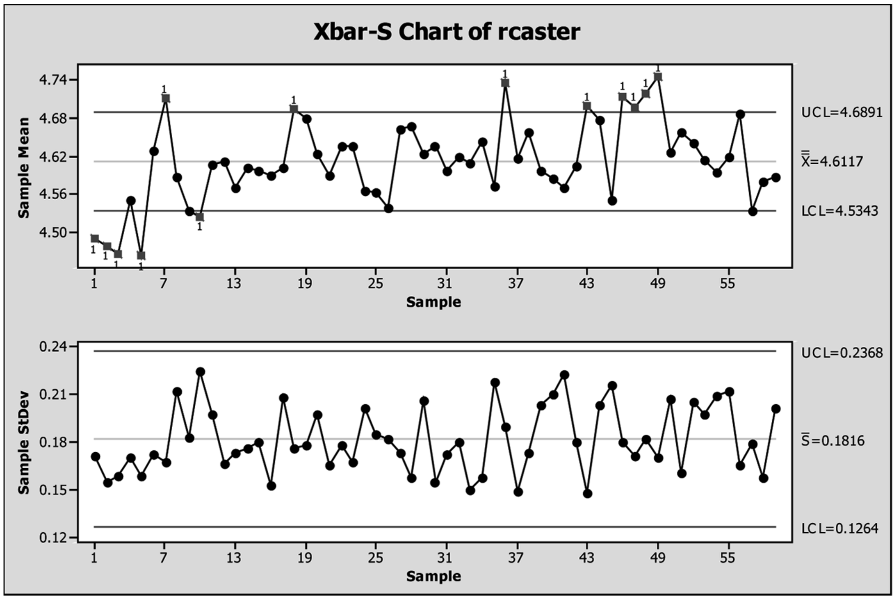

FIGURE 7 $\overline{X}$ **and s charts based on subgroup mean and standard deviation (ignores gauges).**

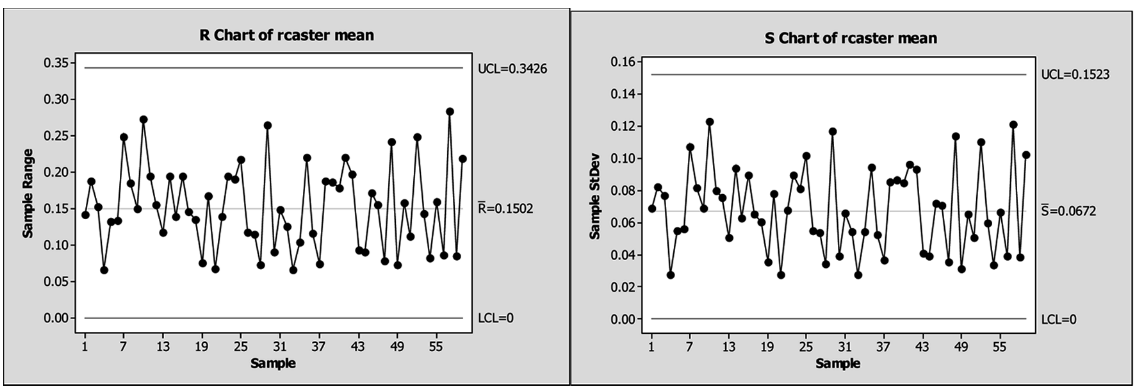

FIGURE 8 **Runger R chart on left, gauge average S charts for the gauge averages in each subgroup. Mortell and Runger *r* chart on left, *s* chart on right.**

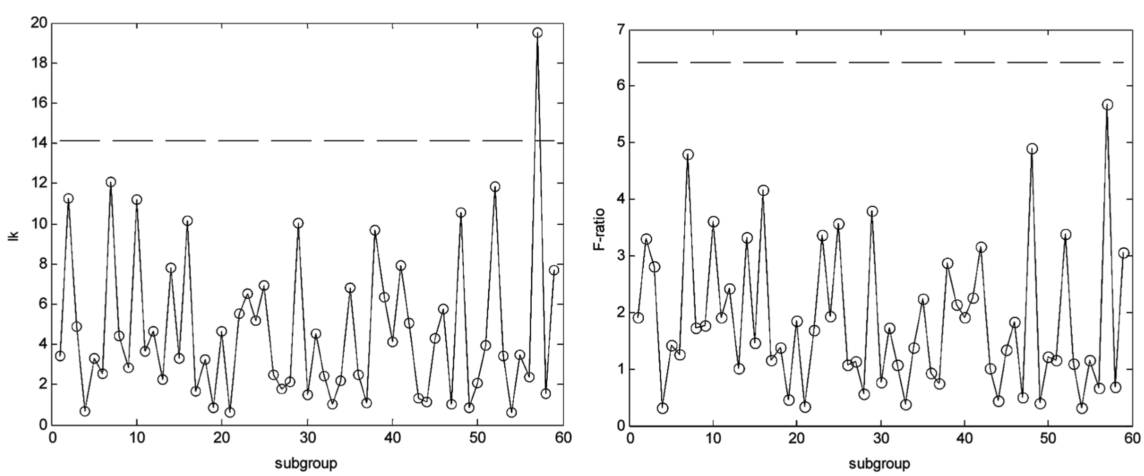

FIGURE 9 ***l* and *f* charts based on the subgroup data stratified by gauge. Likelihood ratio statistic: (*l*) chart on left, F-test (*f*) chart on right.**

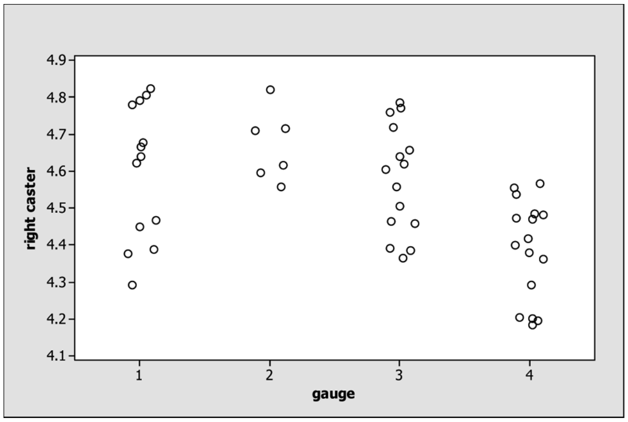

FIGURE 10 **Right caster by gauge for subgroup 57.**

depends on the number of gauges and the subgroup size for each gauge. The likelihood ratio chart statistic with unknown variance and F test chart statistic described here have an advantage over other methods since they do not require any historical data. Not requiring phase 1 will save time and money. Therefore, we recommend these two methods when there are either a large number of gauges or streams or a large number of observations in each subgroup.

In the range method, Mortell and Runger assumed that the process variance is large relative to the within-gauge/stream variance. The discussion in this article is also based on this assumption. However, in practice, this condition may not hold. Then, we believe that the likelihood chart statistic is also applicable if we change some assumptions in our model. We can use the likelihood chart statistic for monitoring measurement system bias and variance consistency at the same time.

Extensions of this work could include adapting any of the proposed test statistics to a sequential control chart such as a cumulative sum (CUSUM) or exponentially weighted moving average (EWMA). For example, to check for small persistent changes in a single gauge we could calculate a CUSUM for each l_k separately, and then plot the max of the m CUSUM statistics. Sequential charts should provide quicker detection of small sustained shifts in the gauge biases than the Shewhart type charts presented in this article.

ACKNOWLEDGEMENTS

The authors thank the referees and associate editor for valuable comments that led to a substantial improvement in the presentation of this material. This research was supported, in part, by the Natural Sciences and Engineering Research Council of Canada.

ABOUT THE AUTHORS

Xuyuan Liu is a Ph.D. student at the School of Industrial and System Engineering at Georgia Institute of Technology. Prior to coming to the United States, Xuyuan Liu worked as a graduate assistant at the University of Waterloo. Xuyuan Liu holds a masters degree in statistics from the University of Waterloo and a bachelor's degree in statistics from Peking University.

Stefan Steiner and Jock MacKay are both associate professors in the Statistics and Actuarial Science Department of the University of Waterloo. They are also active consultants who have worked with organizations from a wide range of industries, including automotive, telecommunications, aerospace, government, and more.

REFERENCES

Boyd, D. F. (1950). Applying the group chart for $\bar{X}$ and R. *Industrial Quality Control*, *7*:22–25.

Chang, T. C. Gan, F. F. (2006). Monitoring linearity of measurement gauges. *Journal of Statistical Computation and Simulation*, *76*(10):889–911.

Harter, H. L. (1960). Critical values for Duncan's new multiple range test. *Biometrics*, *16*(4):671–685.

Montgomery, D. C. (2004). *Introduction to Statistical Quality Control*, 5th ed. New york: John Wiley & Sons.

Mortell, R. R. and Runger, G. C. (1995). Statistical process control of multiple stream processes. *Journal of Quality Technology*, *27*:1–12.

Nelson, L. S. (1986). Control chart for multiple stream processes. *Journal of Quality Technology*, *18*:255–256.