# 30,000-Foot-Level Performance Metric Reporting

UNDERSTANDING AND IMPROVING PROCESSES FROM A BIRD'S-EYE VIEWPOINT

*By Forrest W. Breyfogle III, Smarter Solutions Inc.*

Consider the view from an airplane. When the airplane is at an elevation of 30,000 feet, passengers can see a big-picture view of the landscape from a window. When the airplane is at 50 feet during landing, passengers view a much smaller, more detailed portion of the landscape. Similarly, in the relationship $Y = f(X)$, a 30,000-foot-level statistical control chart could be used to provide a macro view of how a key process output variable, or $Y$, is performing over time. Meanwhile, a 50-foot-level control chart could give a time-series micro view of a key process input variable.

Statistical control charts are typically used to identify in a timely fashion, as a control mechanism, special cause conditions at a low level, or a 50-foot level. An example of this form of control would be to identify when temperature (an important $X$ to a process) significantly changes so that the process can be adjusted before a large amount of unsatisfactory product is produced.

In contrast, the 30,000-foot-level approach uses infrequent subgrouping or sampling to capture from a high level how a process is performing relative to overall customer needs. A sampling frequency that is long enough to span all short-term process noise inputs—such as between-day differences from raw material lots and personnel or machine differences—provides a high-level perspective that is often predictive. With this process-tracking and reporting approach, when an expected future performance statement for a process is undesirable, the improvement metric is "pulling" for a process improvement effort to change either the process's steps or its inputs so the process output transitions to an enhanced level of performance.

### Approaches to control charting

In the second half of the 1920s, Walter A. Shewhart of Bell Telephone Laboratories developed a theory of statistical quality control in which he concluded there were two components to variations displayed in all manufacturing processes.[1, 2]

The first component was a steady component—random variation that appeared to be inherent in the process. The second component was an intermittent variation to assignable causes. He concluded assignable causes could be economically discovered and removed with an effective diagnostic program, but random causes could not be removed without making basic process changes.

Shewhart is credited with developing the standard control chart test based on three standard deviation limits to separate the steady component of variation from assignable causes, in which calculated control chart upper control limits (UCL) and lower control limits (LCL) are a function of process variability and are independent of specification limits. Shewhart control charts came into wide use in the 1940s because of war production efforts. Western Electric is credited with the addition of other tests based on sequences or runs.[3]

In his book *Out of the Crisis*, W. Edwards Deming wrote: "A fault in the interpretation of observations, seen everywhere, is to suppose that every event (defect, mistake, accident) is attributable to someone (usually the one nearest at hand), or is related to some special event. The fact is that most troubles with service and production lie in the system."

"We shall speak of faults of the system as common causes of trouble, and faults from fleeting events as special causes," he wrote. "Confusion between common causes and special causes leads to frustration of everyone, and leads to greater variability and higher costs, exactly contrary to what is needed. I should estimate that in my experience most troubles and most possibilities for improvement add up to proportions something like this: 94% belong to the system (responsibility of management), 6% special."[4]

Control charts offer the study of variation and its source over time. Control charts are used to not only monitor and control, but also identify improvement opportunities. But there is more than one way to use control charts for process-response enhancement.

To illustrate this point, consider the following approaches:

- **Shewhart:** Control charts can identify assignable causes that could be internal or external to the system.
- **Deming:** Control charts can separate a process's special causes from common causes. Special causes originate from fleeting events experienced by the system and common causes originate from the natural variation of the process that is internal and external to the system.

The traditional Shewhart control charting approach emphasizes identification and resolution of assignable causes. This can be beneficial to process performance when it is important to chart the process input so it does not drift or have an unusually low or high response, which could affect the overall output of the process. Using this traditional approach, special causes must be identified early so issues can be resolved in a timely fashion, before many poor-quality parts or transactions are produced. However, there can be issues with these charts when attempting to use the Shewhart control chart approach to track the output of a process. More on this point will be elaborated upon later in this article.

### Calculating capability and performance indexes

Process capability and performance studies often use indexes to describe how a process is performing relative to specification criteria. A customer might set process capability and process performance indexes targets, and then ask suppliers to report on how well they meet these targets.

The Automotive Industry Action Group (AIAG)[5] provides the following definitions:

- **$C_p$:** The capability index, defined as the tolerance width divided by the process capability, irrespective of process centering.
- **$C_{pk}$:** The capability index that accounts for process centering. It relates the scaled distance between the process mean and the closest specification limit to half the total process spread.
- **$P_p$:** The performance index, defined as the tolerance width divided by the process performance, irrespective of process centering. Typically, it is expressed as the tolerance width divided by six times the sample standard deviation. It should be used only to compare to $C_p$ and $C_{pk}$ and to measure and prioritize improvement over time.
- **$P_{pk}$:** This is the performance index, which accounts for process centering. It should be used only to compare to $C_p$ and $C_{pk}$ and to measure and prioritize improvement over time.

These definitions are not followed by all organizations. Some organizations interpret process capability as how well a product performs relative to customer needs (or specification). This interpretation is closer to the definition given earlier for process performance. Organizations may require or assume that processes are in control before conducting process capability and performance assessments. Other organizations lump all data together, resulting in special-cause data increasing the value for long-term variability; however, a process should be considered stable before any process capability or performance statement is made. The term "process performance" is not always used to describe $P_p$ and $P_{pk}$.

The equations for process capability and performance indexes are quite simple but sensitive to the input value for standard deviation ($\sigma$). There are various opinions on how to determine standard deviation in a given situation. AIAG[6] provides guidance on how to calculate this standard deviation term from an $\bar{x}$ and $R$ control chart through use of the following definitions:

**Inherent process variation:** The portion of process variation due to common causes only. This variation can be estimated from an $\bar{x}$ and $R$ control chart by dividing the average range of the control chart $\bar{R}$ by $d_2$, which is a constant that depends on the number of samples in the subgroup. $\bar{R}/d_2$ is shown in Table 1.

**Total process variation:** Process variation due to common and special causes. This variation may be estimated by *s*, the sample standard deviation, using all the individual readings obtained from either a detailed control chart or a process study; that is:

$$s = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}} = \hat{\sigma}_s$$

in which $x_i$ is an individual reading, $\bar{x}$ is the average of individual readings and $n$ is the total number of all of the individual readings.

**Process capability:** The $6\sigma$ range of a process's inherent variation; for statistically stable processes only, where $\sigma$ is usually estimated by $\bar{R}/d_2$.

**Process performance:** The $6\sigma$ range of a process's total variation, in which $\sigma$ is usually estimated by *s*, the sample standard deviation.

The process capability index $C_p$ represents the allowable tolerance interval spread in relation to the actual spread of the data when the data follow a normal distribution. This equation is:

$$C_p = \frac{\text{USL - LSL}}{6\sigma}$$

in which USL and LSL are the upper specification limit and lower specification limit, respectively, and $6\sigma$ describes the range or spread of the process. Data centering is not taken into account in this equation. Other options to determine ($\sigma$) are determined later in this article.

$C_p$ addresses only the spread of the process, while $C_{pk}$ is used concurrently to consider the spread and mean shift of the process. Mathematically, $C_{pk}$ can be represented as the minimum value of the two quantities

$$C_{pk} = \min\left[\frac{\text{USL} - \mu}{3\sigma}, \frac{\mu - \text{LSL}}{3\sigma}\right].$$

An often recommended minimum acceptable process capability index[7] is 1.33 ($4\sigma$); however, Six Sigma programs have suggested striving to obtain a minimum individual process step $C_p$ value of 2 and a $C_{pk}$ value of 1.5.[8]

Process capability and performance indexes $P_p$ and $P_{pk}$ are sometimes referred to as long-term capability and performance indexes. Not all organizations report information as $P_p$ and $P_{pk}$. Some organizations calculate $C_p$ and $C_{pk}$ so they report information that is similar to $P_p$ and $P_{pk}$.

The mathematical calculation of $P_p$ and $P_{pk}$ is similar to that of $C_p$ and $C_{pk}$. For a given situation, the only quantitative difference is the value used for standard deviation.

Table 1. **Factors for constructing variables control charts**

| Subgroup size | $A_2$ | $c_4$ | $d_2$ |
|---|---|---|---|
| 2 | 1.880 | 0.7979 | 1.128 |
| 3 | 1.023 | 0.8862 | 1.693 |
| 4 | 0.729 | 0.9213 | 2.059 |
| 5 | 0.577 | 0.9400 | 2.326 |
| 6 | 0.483 | 0.9515 | 2.534 |

### Confusion with indexes

Practitioners must be careful about the methods they use to calculate and report process capability and performance indexes.[9] Customers might be asking for $C_p$ and $C_{pk}$ metrics when the documentation really stipulated the use of a long-term estimate for standard deviation. The supplier might not understand this and initially operate under the assumption that $C_p$ and $C_{pk}$ measure short-term variability. A misunderstanding of this type between customer and supplier could be costly.

Another possible source of confusion is the statistical computer program package used to calculate these indexes. Consider an instance in which a supplier entered randomly collected data into a computer program, thinking the usual sampling standard deviation formula would be the source of the standard deviation value used in the capability computations. The computer program presumed by default that the data were collected sequentially. The computer program estimated a short-term standard deviation by calculating the average moving range of the sequential entries and converting this moving range value to a standard deviation.

The computer program listed the response as $C_p$ and $C_{pk}$. The practitioner thought that he or she had used the program correctly because the output ($C_p$ and $C_{pk}$) was consistent with the customer's request. However, the data were not generated in sequence. If the same data were reentered in a different sequence, a different $C_p$ and $C_{pk}$ metric would probably result. For non-sequentially generated data, the practitioner should limit calculations to options of this program that lead to a $P_p$ and $P_{pk}$ type computation. The underlying assumption with this approach is that the data are collected randomly over a long period of time and accurately describe the population of interest.

Process capability and performance index metrics require good communication and agreement on the techniques used for calculation. These agreements also should include sample size and measurement considerations. You can avoid many of these issues by reporting in process capability and performance noncompliance rate units, not $C_p$, $C_{pk}$, $P_p$ and $P_{pk}$ indexes.

A great deal of confusion and difference of opinion exists related to short term and long-term variability. Consider the basic differences, which can be grouped under two main categories.

**Opinion one.** Process capability describes the capability or the best a process could currently be expected to work. It does not address directly how well a process is running relative to the needs of the customer. Rather, it considers short-term variability. A long-term variability assessment attempts to address directly how well the process is performing relative to customer needs. Special causes, which have the most impact on long-term variability estimates from a control chart, might be included in the analyses. Some might object that predictions couldn't be made without process stability. Processes can appear to be out of control from day-to-day variability effects such as raw material, which, they would argue, is common cause variability.

Standard deviation input-to-process capability and process performance equations can originate from short term or long-term considerations. In determining process capability indexes from $\bar{x}$ and $R$ control chart data, the standard deviation within subgroups is said to give an estimate of the short-term variability of the process, while the standard deviation of all the data combined is said to give an estimate of its long-term variability.

In a manufacturing process, short-term variability typically does not include, for example, raw material lot-to-lot variability and operator-to-operator variability. Within a business process, short-term variability might not include day-to-day variability or department-to-department variability. Depending on the situation, these long-term variability sources might be considered special causes and not common causes.

Process capability indexes $C_p$ and $C_{pk}$ typically assess the potential short-term capability by using a short-term standard deviation estimate, while $P_p$ and $P_{pk}$ typically assess overall long-term capability by using a long-term standard deviation estimate. Sometimes, the relationship $P_p$ and $P_{pk}$ is referred to as process performance.

Some organizations require or assume that processes are in control before conducting process capability and performance index assessments. Other organizations lump all data together, which results in special-cause data increasing the estimates of long-term variability. These organizations might try to restrict the application of control charts to monitoring process inputs.

**Opinion two.** Process capability describes how well a process is executing relative to the needs of the customer. The terms "short term" and "long term" are not typically considered separately as part of a process capability assessment.

The quantification for the standard deviation term within process capability calculations describes the overall variability of a process. When determining process capability indexes from $\bar{x}$ and $R$ control chart data, an overall standard deviation estimate would be used in the process capability equations. Calculation procedures for standard deviations differ from one practitioner to another, ranging from combining all data together to determining total standard deviation from a variance components model.

This opinion takes a more long-term view of variability. It involves a different view of factors in an in-control process. In manufacturing, raw material lot-to-lot variability and operator-to-operator variability are more likely considered common causes. In a business process, day-to-day variability or department-to-department variability are more likely considered common causes.

Process capability indexes $C_p$ and $C_{pk}$ typically address the needs of customers and have a total standard deviation estimate within the calculations. $P_p$ and $P_{pk}$ are not typically used as a metric in this approach.

## Calculating standard deviation

Confusion can also often be encountered with regard to the calculation of the seemingly simple standard deviation statistic.[10] Though standard deviation is an integral part of the calculation of process capability, the method used to calculate it is rarely adequately scrutinized. In some cases, it is impossible to get a specific desired result if data are not collected in the appropriate fashion. Consider the following three sources of continuous data:

1. **Situation A:** an $\bar{x}$ and $R$ control chart with subgroups with a sample size of five.
2. **Situation B:** an $X$ chart with individual measurements.
3. **Situation C:** a random sample of measurements from a population.

All three are real possible sources of information, but no one method is correct for obtaining an estimate of standard deviation $\sigma$ in all three scenarios. Author Thomas Pyzdek[11] presents five methods of calculating standard deviation. Figure 1 (p. 23) illustrates six approaches to make this calculation.

### Method one: long-term estimate of σ

One approach for calculating the standard deviation of a sample ($s$) is to use the formula:

$$\hat{\sigma} = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}}$$

in which $\bar{x}$ is the average of all data, $x_i$ are the data values and $n$ is the overall sample size.

Sometimes, computer programs apply an unbiasing term to this estimate, dividing the above by $c_4(n-1)$.[12] Tabulated values for $c_4$ at $n-1$ are shown in Table 1.

1. **Situation A:** When data come from an $\bar{x}$ and $R$ chart, this traditional estimate of standard deviation is only valid when a process is stable, though some use this method even when processes are not stable. Shewhart showed that this method overestimates scatter if the process is influenced by a special cause. This estimate should never be used to calculate control limits. Control limits are calculated using sampling distributions.
2. **Situation B:** When data are from an individuals control chart, this approach can give an estimate of process variability from the customer's point of view.
3. **Situation C:** For a random sample of data from a population, this is the only method that makes sense because the methods that follow all require the sequence of part creation.

### Method two: short-term estimate of σ

A standard method for estimating standard deviation from $\bar{x}$ and $R$ control chart data is:

$$\hat{\sigma} = \frac{\bar{R}}{d_2}$$

in which $\bar{R}$ is the average of the subgroup range values from a control chart and $d_2$ is a value from Table 1 that depends on subgroup sample size.

1. **Situation A:** When data come from an $\bar{x}$ and $R$ chart, the problem of the standard deviation being inflated by special causes is alleviated because it does not include variation between time periods. Shewhart proposed using a rational subgroup to achieve this, in which the subgroup sample is chosen so that the opportunity for special causes is minimized. Often, this is accomplished by selecting consecutively produced units from a process. The method of analysis is inefficient when range is used to estimate standard deviation because only two data values are used from each subgroup. This inefficiency increases as the subgroup size increases. Efficiency increases when subgroup standard deviation is used.
2. **Situation B:** When data are from an individuals control chart, this calculation is not directly possible because the calculation of $\bar{R}$ for a subgroup size of one is not possible.
3. **Situation C:** For a random sample of data from a population, this calculation is not possible because the sequence of unit creation is not known.

### Method three: short-term estimate of σ

The following equation is derived from the equation used to determine the centerline of a control chart when the process standard deviation is known:

$$\hat{\sigma} = \frac{\bar{s}}{c_4}$$

in which $\bar{s}$ is the average of the subgroup standard deviation values from a control chart, and $c_4$ is a value from Table 1 that depends on subgroup sample size. Subgroup standard deviation values are determined by the formula shown in method one.

1. **Situation A:** When data are from an $\bar{x}$ and $R$ or $s$ chart, the comments relative to this situation are similar to the comments in method two. Compared to method two, this approach is more involved but more efficient.
2. **Situation B:** When data are from an individuals control chart, this calculation is not possible because the calculation of $\bar{s}$ for a subgroup size of one is not possible.
3. **Situation C:** For a random sample of data from a population, this calculation is not possible because the sequence of unit creation is not known.

### Method four: short-term estimate of σ

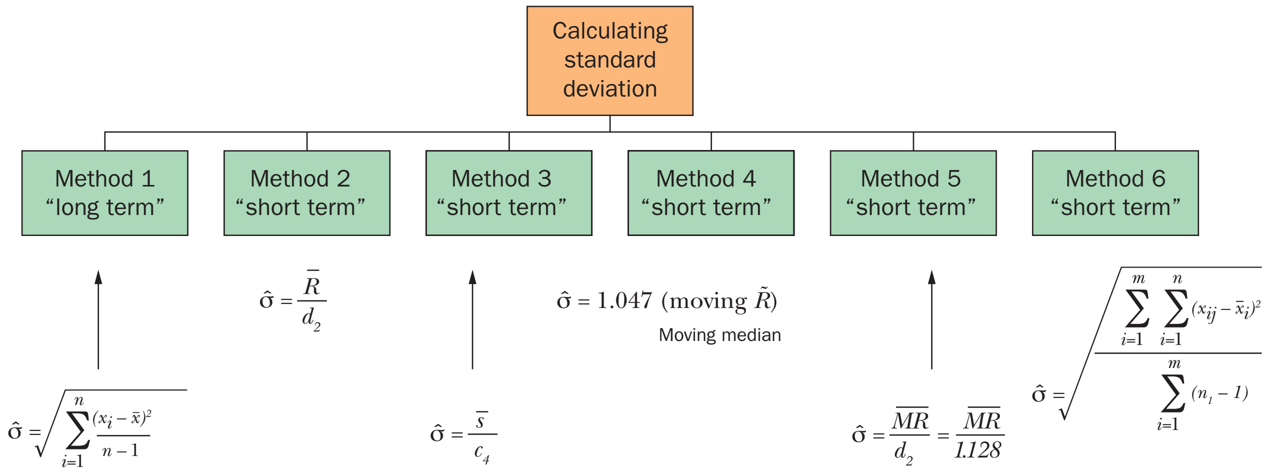
The following relationship is taken from one of the equation options used to determine the centerline of an individual control chart:

$$\hat{\sigma} = 1.047 \text{ (Moving } \tilde{R}\text{)}$$

in which a correction factor of 1.047 is multiplied by the median of the moving range (Moving $\tilde{R}$).

1. **Situation A:** When data are from an $\bar{x}$ and $R$

Figure 1. **Various ways to calculate process standard deviation**




chart, this approach is not directly applicable.

2. **Situation B:** When data are from an individuals control chart, this calculation is an alternative. If the individuals control chart values are samples from a process, you would expect a higher value if there is less variability from consecutively created units when compared to the overall variability experienced by the process between sampling periods of the individuals control chart. Research has recently indicated that this approach gives good results for a wide variety of out-of-control patterns.
3. **Situation C:** For a random sample of data from a population, this calculation is not possible because the sequence of unit creation is not known.

### Method five: short-term estimate of σ

The following equation derives from one of the equations used to determine the centerline of an individual control chart:

$$\hat{\sigma} = \frac{\overline{MR}}{d_2} = \frac{\overline{MR}}{1.128}$$

in which $\overline{MR}$ is the moving range between two consecutively produced units and $d_2$ is a value from the table of factors for constructing control charts using a sample size of two.

1. **Situation A:** When data are from an $\bar{x}$ and $R$ chart, this approach is not directly applicable.
2. **Situation B:** When data are from an individuals control chart, this calculation is an alternative. Most of the method four comments for this situation are similarly applicable. This is the method suggested by AIAG.[13] Some practitioners prefer method four over method five.[14]
3. **Situation C:** For a random sample of data from a population, this calculation is not possible because the sequence of unit creation is not known.

### Method six: short-term estimate of σ

The following relationship is sometimes used by computer programs to pool standard deviations when there are $m$ subgroups of sample size $n$:

$$\hat{\sigma} = \frac{s_p}{c_4(d)}$$

in which

$c_4(d)$ is a value that can be determined from Table 1 and

$$s_p = \sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \bar{x}_i)^2}{\sum_{i=1}^{m}(n_i - 1)}}$$

and

$$d = \left(\sum_{i=1}^{m} n_i\right) - m + 1 .$$

The purpose of using $c_4(d)$ when calculating $\hat{\sigma}$ is to reduce bias to this estimate. Values for $c_4$ are given in Table 1.

1. **Situation A:** When data come from an $\bar{x}$ and $R$ or $s$ chart, the comments relative to this situation are similar to the comments in methods two and three. If all groups are to be weighed the same—regardless of the number of observations—the $\bar{s}$

(or $\bar{R}$) approach is preferred. If the variation is to be weighted according to subgroup size, the pooled approach is appropriate.

2. **Situation B:** When data are from an individuals control chart, this calculation is not directly possible because the calculation of $\bar{R}$ for subgroup size of one is not possible.
3. **Situation C:** For a random sample of data from a population, this calculation is not possible because the sequence of unit creation is not known.

The concepts of this variability discussion will now be applied to control charting and process capability reporting.

### Shewhart vs. Deming

An $\bar{x}$ and $R$ control chart are often the preferred method for tracking continuous data over time. Classically, it is stated that $\bar{x}$ and $R$ control charting subgroup intervals should be selected so that variation among the units within a subgroup is small. The thought is that if variation within a subgroup represents the piece-to-piece variability over a very short period of time, any unusual variation between subgroups would reflect changes in the process that should be investigated for appropriate action. With this approach, process enhancements are often undertaken through the identification and resolution of an assignable cause that caused an out-of-control condition.

This form of control can beneficial. However, organizations often give focus on the output of a process when applying control charts. This type of measurement is not really controlling the process and may not offer timely problem identification. To control a process, using $\bar{x}$ and $R$ control charts are most beneficial when monitoring key process input variables, where process flow is stopped for timely resolution when the variable goes out of control.

However, Deming's description of the terms "special cause" and "common cause" variability can lead to a different point of view for the construction and use of control charts.[15, 16] Deming's description of common cause and special cause could be applied to not only a process's input but also the output of a process.

To illustrate this point, a situation will be examined where an $\bar{x}$ and $R$ control chart process was selected that has one operator per shift, and batch-to-batch raw material changes occur daily. In this illustration, consider also that there are some slight operator-to-operator differences and raw material differences from batch to batch, but raw material is always within specification limits.

If a control chart were established in which five pieces are taken in a row for each day, the variability used to calculate $\bar{x}$ and $R$ control chart limits does not consider operator-to-operator and batch-to-batch raw material variability. If the variability between operators and batch to batch is large relative to five pieces in a row, the process could often have out-of-control signals, where intervention is to investigate and resolve an assignable cause. Much frustration can occur in manufacturing when time is spent to no avail trying to fix a problem over which operators may have little, if any, control.

Let's examine this situation using the terminology of Shewhart and Deming. With Shewhart's terminology, control charts are examined to identify out-of-control signals that are the result of assignable causes that are to be resolved. Meanwhile, Deming's point of view suggests that a process inherently has common-cause variability from internal and external sources, and special cause events can cause out-of-control conditions.

From a Deming point of view, the variability between operators and batch to batch in the example could be considered sources of common-cause variability. With this perspective, operations should not react to operator and batch difference as though they were special cause, even though they are assignable, using Shewhart's terminology.

From the Deming perspective, that's not to say the raw material does not cause a problem to the process, despite being within its specified tolerance. The point is whether the variability between raw material lots should be treated as a special cause. It seems that even though the raw material lot variability could be classified as assignable (Shewhart's terminology), a strong argument can be made to treat this type of variability as common cause (Deming's terminology).

From a Deming perspective, you could conclude that control limits then should be created so that raw material lot-to-lot variability and differences between operations are included in the calculated control limits. This charting approach provides a higher-elevated view of the process than traditional control charts. If we consider an airplane's in-flight view of the earth, traditional control charts as proposed by Shewhart provide a 50-foot elevation view of earth, while a high-level planet view could be from 30,000 feet.

Because of this analogy, refer to this high level, time-series process output reporting as a 30,000-foot level, while traditional process input tracking would be at the 50-foot level. Unlike 50-foot-level control charts, the process measurements from a high-level perspective suggests infrequent subgrouping and sampling to capture how the process is performing relative to overall customer needs. The subgrouping frequency when creating 30,000-foot-level control charts must be long enough to span all short-term process noise inputs, such as raw material differences and between days or daily cycle differences.

Now examine control chart equation options for continuous data to determine a method for 30,000-foot-level time-series reporting. Figure 2 provides a visual of the previously described situation with its overall process's two components of variation—within day and between-day variability. Again, what is desired is that the magnitude of both variation components are involved in the control chart upper and lower calculated values, if the process were tracked through daily subgroupings.

The traditional approach for tracking the previous described situation would be an $\bar{x}$ and $R$ control chart using the equation:

$$UCL = \bar{\bar{x}} + A_2\bar{R} \qquad LCL = \bar{\bar{x}} - A_2\bar{R}$$

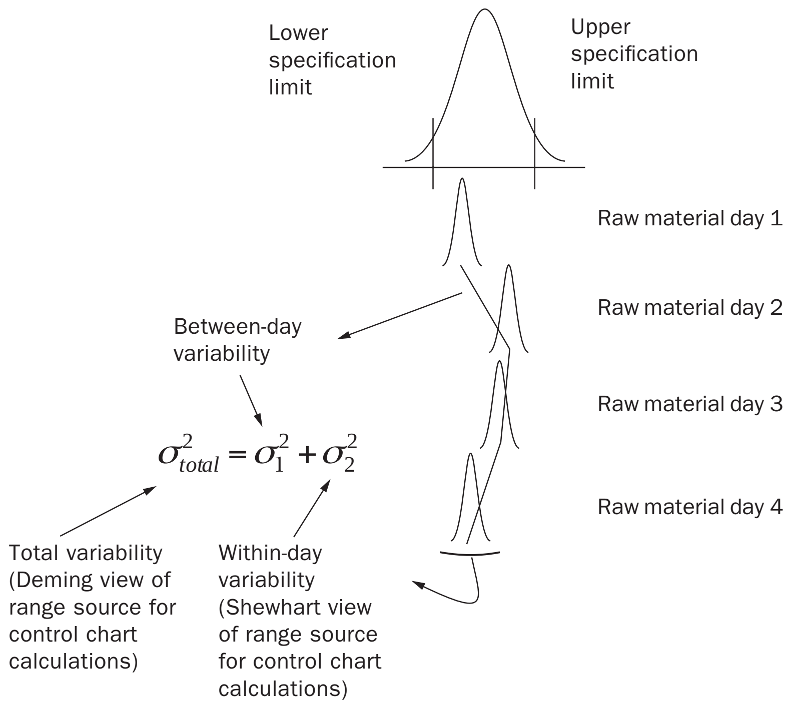
in which $\bar{\bar{x}}$ is the overall average of the subgroups, $A_2$ is a constant depending upon subgroup size (see Table 1) and $\bar{R}$ is the average range within subgroups. This $\bar{R}$ includes only a $\sigma^2_{within}$ influence.

An alternative to tracking multiple samples from each subgroup distribution is the selection of only one sample per unit of time. For this situation, an individuals chart ($X$ chart) or individuals-moving-range chart ($XmR$ chart) would be appropriate process tracking alternatives. For presentation simplicity, only the $X$ chart will be used in the following discussion because the addition of a moving-range chart would only add the inclusion of data-point-transitions tracking, which is included in the individuals chart. Now examine the individuals or $X$ chart calculation for the UCL and LCL in the following equations:

$$UCL = \bar{x} - + \frac{3}{d_2}(\overline{MR}) = \bar{x} + \frac{3}{1.128}(\overline{MR}) = \bar{x} + 2.66(\overline{MR})$$

$$LCL = \bar{x} - \frac{3}{d_2}(\overline{MR}) = \bar{x} - \frac{3}{1.128}(\overline{MR}) = \bar{x} - 2.66(\overline{MR})\,.$$

Figure 2. **Time series data with calculations for 30,000-foot-level control chart creation**




In these equations, $\overline{MR}$ is the average moving range between subgroups, which is added or subtracted from the process mean after being adjusted by a multiple of three standard deviations divided by a 1.128 $d_2$ constant for adjacent subgroups. This $\overline{MR}$ includes $\sigma^2_{total}$ influence.

Which control charting technique is most appropriate? It depends on how the source of variability is considered relative to common and special causes, as previously described. The 30,000-foot-level tracking approach considers that variability between subgroups should affect the control limits. For this to occur, a sampling plan is needed where the impact from this type of common-cause noise variable occurs between subgroupings.

From the earlier $\bar{x}$ control charting limits equation, it is noted that between subgroups, common-cause variability has no impact on calculated control limits. Therefore, whenever common-cause variability does occur between subgroups, an $\bar{x}$ and $R$ control chart can trigger this occurrence as a special cause. This behavior is different with the individuals control chart, where control chart limits are a function of the mean moving range, which is an estimate from between subgroup variability. The between subgroup needs for calculating 30,000-foot-level control charts is achieved with individuals control charts and not with $\bar{x}$ and $R$ control charting. Because of this, 30,000-foot-level reporting will use individuals control charts to determine if a process's within subgroup mean and standard deviation remains stable over time.

### Deming's view, Shewhart's charts

Time-series data, like that shown in Table 2,[17] could have been generated from measuring a manufactured part's dimension, response time in a call center, or the time it took to complete a transaction. In any case, five daily sample responses were documented for a 10-day period. The tabular mean and range value calculations in the table will be used to determine the upper and lower control chart limits.

A process capability and performance statement is desired for this time-series data relative to specification limits of 95 to 105.

The described scenarios where this data could have originated would suggest the application of a 30,000-foot level performance metric chart. However, an $\bar{x}$ and $R$ control chart would traditionally be used for time-series tracking data of this type. For purposes of illustration, these two charting differences will be compared for this set of data to illustrate how different conclusions could be made relative to what should be done issues within this process.

An $\bar{x}$ and $R$ control chart of this data is shown in Figure 3, in which the UCL and LCL were determined using the following equations. Note: slight differences may occur between the computer generated control chart limits and calculated values because of the number of digits that was used for the constant for the constants or rounding differences.

$$UCL = \bar{\bar{x}} + A_2\bar{R} = 101.044 + 0.577(2.1) = 102.26$$

$$LCL = \bar{\bar{x}} - A_2\bar{R} = 101.044 - 0.577(2.1) = 99.83$$

Some courses may teach that you should not generate a control chart with only 10 subgroups, which is intended to reduce the uncertainty of the standard

Table 2. **Time series data with multiple samples in daily subgroupings and calculations for $\bar{x}$ and $R$ control chart creation**

| Day | Sample one | Sample two | Sample three | Sample four | Sample five | Mean | Range |
|---|---|---|---|---|---|---|---|
| 1 | 102.7 | 102.2 | 102.7 | 103.3 | 103.6 | 102.9 | 1.4 |
| 2 | 108.2 | 108.8 | 106.7 | 106.6 | 109.1 | 107.9 | 2.5 |
| 3 | 101.9 | 103.0 | 100.6 | 101.4 | 101.3 | 101.6 | 2.4 |
| 4 | 103.9 | 105.5 | 104.3 | 104.5 | 104.5 | 104.5 | 1.6 |
| 5 | 97.2 | 99.0 | 96.5 | 94.9 | 96.5 | 96.8 | 4.1 |
| 6 | 94.4 | 93.0 | 93.0 | 95.2 | 93.6 | 93.8 | 2.2 |
| 7 | 104.7 | 103.6 | 103.7 | 104.7 | 104.5 | 104.2 | 1.1 |
| 8 | 102.5 | 102.7 | 101.2 | 100.6 | 103.1 | 102.0 | 2.5 |
| 9 | 101.9 | 103.1 | 101.0 | 101.2 | 101.4 | 101.7 | 2.1 |
| 10 | 95.0 | 95.3 | 95.3 | 94.4 | 94.2 | 94.8 | 1.1 |
| | | | | | Average | 101.0 | 2.1 |

Table 3. **Time series data of only the first sample from Table 2**

| Day | Sample one | Moving range |
|---|---|---|
| 1 | 102.7 | |
| 2 | 108.2 | 5.5 |
| 3 | 101.9 | 6.3 |
| 4 | 103.9 | 2.0 |
| 5 | 97.2 | 6.7 |
| 6 | 94.4 | 2.8 |
| 7 | 104.7 | 10.3 |
| 8 | 102.5 | 2.2 |
| 9 | 101.9 | 0.6 |
| 10 | 95.0 | 6.9 |
| Average | 101.24 | 4.81 |

Table 4. **Time series data with calculations for 30,000-foot-level control chart creation**

| Day | Sample one | Sample two | Sample three | Sample four | Sample five | Mean | MR mean | Std. Dev. | MR Std. Dev. |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 102.7 | 102.2 | 102.7 | 103.3 | 103.6 | 102.9 | | 0.55 | |
| 2 | 108.2 | 108.8 | 106.7 | 106.6 | 109.1 | 107.9 | 4.98 | 1.17 | 0.62 |
| 3 | 101.9 | 103.0 | 100.6 | 101.4 | 101.3 | 101.6 | 6.24 | 0.89 | 0.28 |
| 4 | 103.9 | 105.5 | 104.3 | 104.5 | 104.5 | 104.5 | 2.90 | 0.59 | 0.30 |
| 5 | 97.2 | 99.0 | 96.5 | 94.9 | 96.5 | 96.8 | 7.72 | 1.48 | 0.89 |
| 6 | 94.4 | 93.0 | 93.0 | 95.2 | 93.6 | 93.8 | 2.98 | 0.95 | 0.53 |
| 7 | 104.7 | 103.6 | 103.7 | 104.7 | 104.5 | 104.2 | 10.40 | 0.55 | 0.41 |
| 8 | 102.5 | 102.7 | 101.2 | 100.6 | 103.1 | 102.0 | 2.22 | 1.07 | 0.52 |
| 9 | 101.9 | 103.1 | 101.0 | 101.2 | 101.4 | 101.7 | 0.30 | 0.84 | 0.23 |
| 10 | 95.0 | 95.3 | 95.3 | 94.4 | 94.2 | 94.8 | 6.88 | 0.51 | 0.33 |
| | | | | | **Average** | **101.04** | **4.96** | **0.86** | **0.46** |

MR = moving range
Std. Dev. = standard deviation

deviation estimate, but having fewer than 25 subgroups or data points will only reduce the chance of detecting an out-of-control condition. Using a smaller number of data points in a control chart increases a beta-risk equivalent, in which a true out-of-control condition may not be detected. Smaller sample control charts that show an out-of-control condition should still be investigated because it is likely that an out-of-the-norm event occurred, given underlying assumptions used to create the chart.

Whenever a measurement on a control chart is beyond the UCL or LCL, the process is said to be out of control. Out-of-control conditions are considered special-cause conditions, and out-of- control conditions can trigger a causal problem investigation.

As an alternative when creating a sampling plan, you might have selected only an individual sample instead of several samples for each subgroup. Imagine this is what happened for this same process and only the first measurement was observed for each of the 10 subgroups. Table 3 shows this set of data with the inclusion of some additional information for later determination of upper and lower control chart limits.

For this set of data, you could create an individuals control chart or and individuals-moving-range chart (*XmR*). The moving range (*MR*) chart was not included for the sake of report-out simplicity because the *MR* portion of the *XmR* chart only provides additional visual information for large swings in adjacent individuals chart values, which could be visually observed in a stand-alone individuals chart. Calculations to determine the UCL and LCL are:

$$\text{UCL} = \bar{x} - 2.66(\overline{\text{MR}}) = 101.24 - 2.66(4.81) = 114.03$$

$$\text{LCL} = \bar{x} + 2.66(\overline{\text{MR}}) = 101.24 + 2.66(4.81) = 88.45$$

This individuals control chart shown in Figure 4 provides a different view than the $\bar{x}$ and $R$ control chart

Figure 3. **$\bar{x}$ and *R* control chart of the data in Table 2**

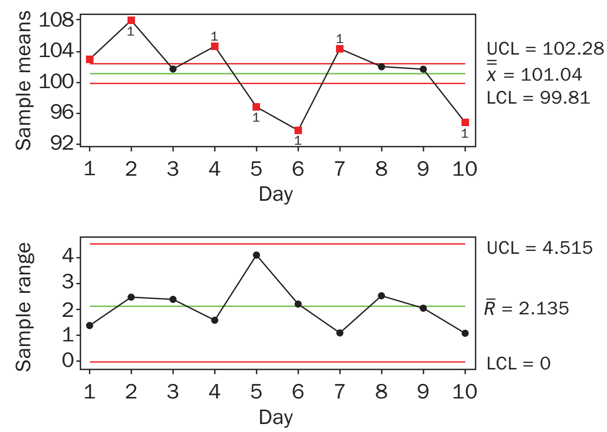
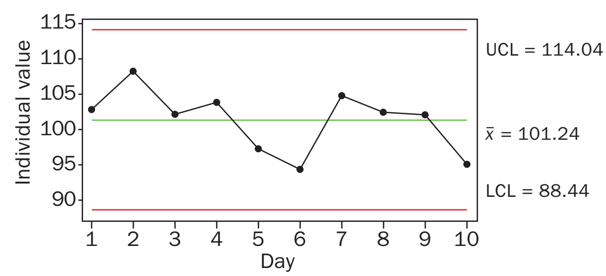


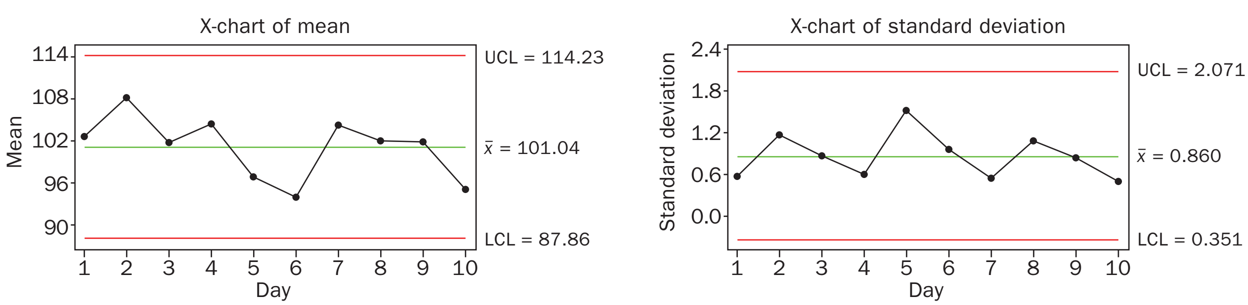

Figure 4. **Individuals chart of sample 1**




LCL = lower control limit
UCL = upper control limit

shown in Figure 3 relative to whether the process is experiencing special-cause conditions or not. The Figure 4 charting alternative is not unlike the creation of a 30,000-foot-level performance-tracking chart. Because the plotted values are within the control limits, we would conclude from a 30,000-foot-level perspective that only common-cause variability exists and the process should be considered in control. The dramatic difference between the limits of these two control charts is

Figure 5. **30,000-foot-level control chart of data in Table 4**




LCL = lower control limit
UCL = upper control limit

caused by the two differing approaches to determining sampling standard deviation in the control charting equation shown previously; that is, the $\bar{x}$ and $R$ control chart equations do not take into account the variation between subgroups.

Figure 4 illustrates one form of reporting, a 30,000-foot-level performance tracking chart. Another approach is to consider each subgroup average and standard deviation as an individual random sample of the population mean and standard deviation. Using this approach, you can report out the subgroup average and standard deviation using an individuals chart for the creation of each plot. Table 4 (p. 26) shows the calculation of inputs to the UCLs and LCLs for the creation of these charts.

Figure 6. **Illustration of 30,000-foot-level process capability and performance determination**

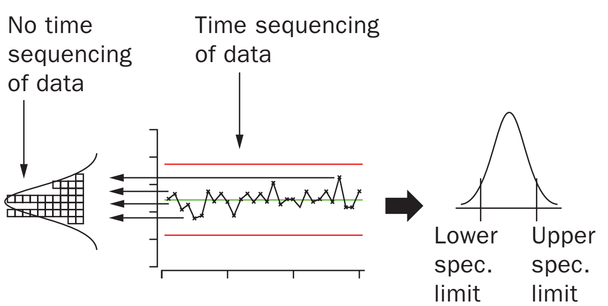
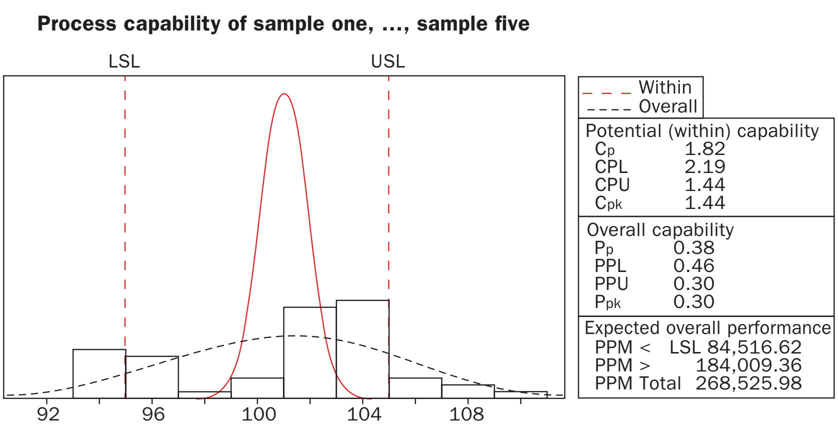



Figure 7. **Process capability indexes report-out**




$C_p$ = capability index
$C_{pk}$ = capability index, accounting for process centering
CPL = $C_p$ lower (for $C_{pk}$ calculation)
CPU = $C_p$ upper (for $C_{pk}$ calculation)
LSL = lower specification limit
$P_p$ = performance index, irrespective of process centering
$P_{pk}$ = performance index, accounting for process centering
PPL = $P_p$ lower (for $P_{pk}$ calculation)
PPM = parts per million
PPU = $P_p$ upper (for $P_{pk}$ calculation)
USL = upper specification limit

Mean individuals control chart UCL and LCL calculations:

$$\text{UCL} = \bar{x} + 2.66(\overline{\text{MR}}) = 101.04 + 2.66(4.96) = 114.23$$

$$\text{LCL} = \bar{x} - 2.66(\overline{\text{MR}}) = 101.04 - 2.66(4.96) = 87.85\ .$$

Standard deviation control chart UCL and LCL calculations:

$$\text{UCL} = \bar{x} + 2.66(\overline{\text{MR}}) = 0.86 + 2.66(0.46) = 2.08$$

$$\text{LCL} = \bar{x} - 2.66(\overline{\text{MR}}) = 0.86 - 2.66(0.46) = -0.36\ .$$

From this two-chart assessment of the process shown in Figure 5 (p. 27), you can conclude that this process is in control or stable, a different conclusion from the traditional $\bar{x}$ and $R$ control chart of this data shown in Figure 3.

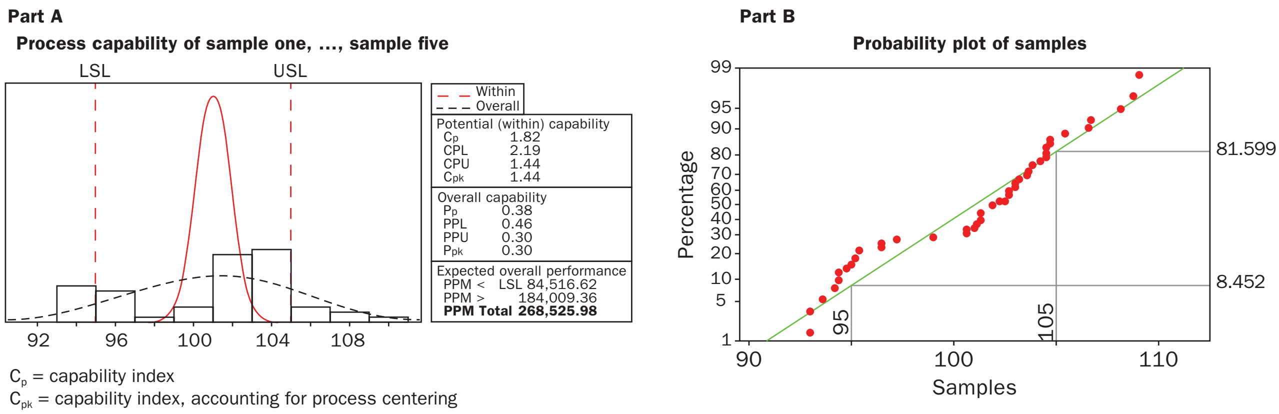
Upon examination, you can notice that the standard deviation individuals control chart has an LCL below zero. Because it is not physically possible to have a minus standard deviation, one of two things could be done to resolve this issue. One would be to move the LCL to zero. The other is to consider implementing a data transformation that makes physical sense.

Because standard deviation has a lower bound of zero, a normal distribution may not adequately model standard deviation subgrouping values from tracking a process over time. The natural skewness of the lognormal distribution can be used to overcome these shortcomings through tracking the log of the standard deviation values over time to determine whether within subgroup variability has changed over time. In this article, no adjustments will be made to the standard deviation chart or the LCL.

### Process performance statements

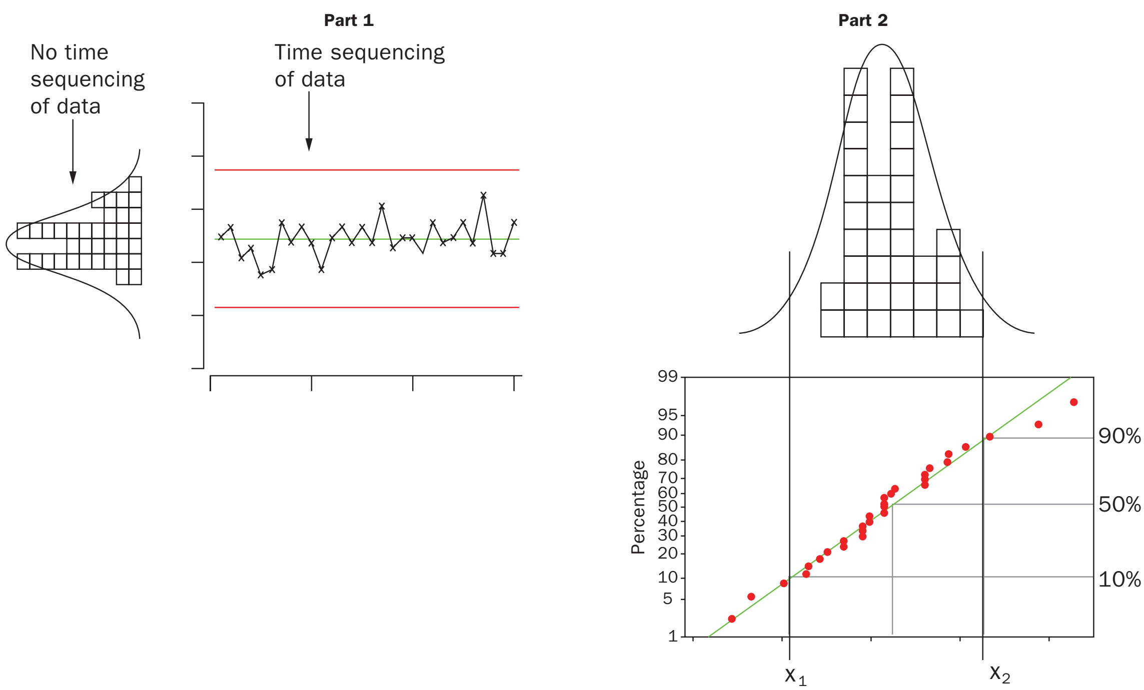
Over time, a 30,000-foot-level individuals control chart could have several regions of stability. If there is a recent region of stability, a process can be said to be predictable—unless something changes, the same basic level of performance could be expected in the future. However, process stability does not mean that the process performance is satisfactory relative to customer needs. To address whether the batch-to-batch

Figure 8. **Comparing response from process capability reporting methods**




$C_p$ = capability index
$C_{pk}$ = capability index, accounting for process centering
CPL = $C_p$ lower (for $C_{pk}$ calculation)
CPU = $C_p$ upper (for $C_{pk}$ calculation)
LSL = lower specification limit
$P_p$ = performance index, irrespective of process centering
$P_{pk}$ = performance index, accounting for process centering
PPL = $P_p$ lower (for $P_{pk}$ calculation)
PPM = parts per million
PPU = $P_p$ upper (for $P_{pk}$ calculation)
USL = upper specification limit

Figure 9. **Reporting process capability and performance when there is no specification**




variability or other variability sources are causing a problem in a stable region of a 30,000-foot-level chart, the total variability of the process could then be compared to specification needs, as illustrated in Figure 6.

For a stable process, the level of performance relative specification limits could be reported using process capability and performance $C_p$, $C_{pk}$, $P_p$ and $P_{pk}$ indexes, as shown in Figure 7. However, the interpretation of these indices can be confusing. A measurement that is easier to understand is the estimated area under the probability density curve beyond the specification range, which would be an estimate for the predicted nonconformance percentage. In Figure 7, this measurement is presented as parts per million (PPM) total, but an alternative approach for presenting this information will be described later.

If someone is using an $\bar{x}$ and $R$ control chart, as shown in Figure 3, to assess process stability, no capability and performance statement should be made until the process is brought into control, which could be difficult to accomplish because, in the real world, there might not be much that can be done to significantly reduce the variability between raw material lots, for example. However, because the 30,000-foot-level time-series-tracking assessment, as shown in Figure 5, considers people and material differences as a source for common-cause input variability, the process is considered stable, and this form of reporting could be used to describe how the process is performing relative to specification limits. The data for creating this process performance estimate is not to be bounded by any calendar considerations. Instead, shifts in process performance should lead to the staging of the control chart, noting that recent region of stability data and predictability statement could be three weeks, three months or three years of data.

With 30,000-foot-level reporting, it would be understood that if the process is not performing satisfactorily relative to this criterion, process improvement effort is needed to address the common-cause variability sources so that enhancements can be made in these areas of the system.

In 30,000-foot-level reporting, an alternative approach to process capability and performance is used where a best estimate for percent beyond specification or in ppm defect rates is reported in lieu of process capability indexes, which can be confusing and depend on how data are sampled from the process.

With 30,000-foot-level reporting of continuous data, a probability plot can provide a visual representation of how well the data are performing relative to specification limits. With this form of reporting, you can also determine whether other adjustments or investigations are needed relative to outlier data points or multi-modal distributions. Results of this form of reporting provide a similar result to a traditional process capability and performance analysis in the area of expected overall performance. Figure 8 (p. 29) illustrates this point in which PPM total from Part A of the figure when converted to a percentage value of 26.8% is approximately equal to the calculated percentage beyond specification from the probability plot, or 8.452 + 100(100 - 81.599) = 26.853.

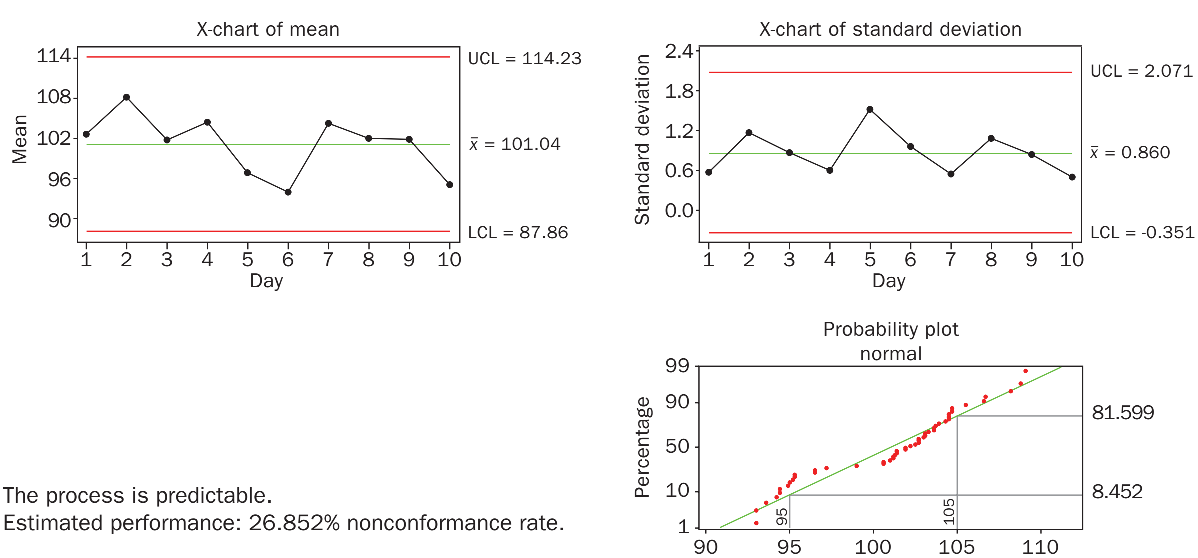
One other point that should be highlighted is that this fabricated data actually consisted of two variance components, where one component was within-day and the other was between-day. Because of this, the data did not have as good a normal probability plot fit as you would like to see. However, even with this extreme fabricated data set, you could still achieve a basic understanding of how the process is performing—the process is stable with an expected percentage that is about 27% nonconformance.

When no specification exists, as might be the case in transactional or service processes, the 30,000-foot-level process capability and performance estimate could be expressed as a median and percentage of occurrence, in which a probability plot is used determine these values, as illustrated in Figure 9 (p. 29). An example of this form of reporting is the process is predictable with an estimated median process order time of 30 days with 80% of orders being filled between 10 and 50 days.

The basic approach for 30,000-foot-level reporting is:[18]

1. Determine an infrequent subgrouping or sam-

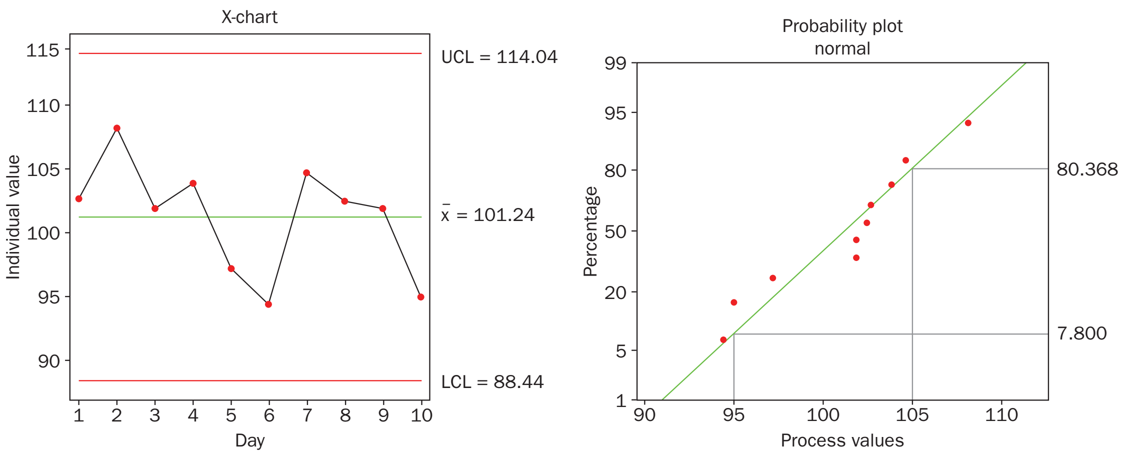
Figure 10. **30,000-foot-level performance metric report-out of data in Table 4**



pling plan so that the typical variability from process input factors occurs between subgroups—subgroup by day, week or month.
2. Analyze the process for predictability using an individuals control chart.
3. When the process is considered predictable, formulate a prediction statement for the latest region of stability. The usual reporting format for this prediction statement is:
   - When there is a specification requirement: nonconformance percentage or defects per million opportunities (DPMO).
   - When there are no specification requirements: median response and 80% frequency of occurrence rate.

This process for metric reporting would lead to a 30,000-foot-level report-out format for the data in Table 2 to be that shown in Figure 10, where an easy-to-understand prediction statement is included at the bottom of the process's report-out.

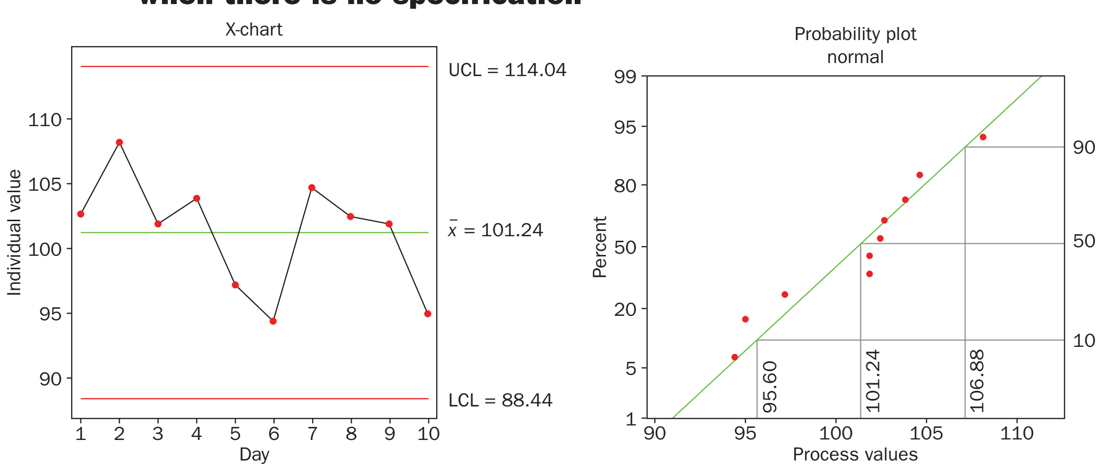
Figure 11. **30,000-foot-level performance metric report-out when there is a specification**




The process is predictable.
Estimated performance: 27.432% nonconformance rate.

LCL = lower control limit
UCL = upper control limit

Figure 12. **30,000-foot-level performance metric report-out when there is no specification**




The process is predictable.
Estimated performance: Median is 101.24 with 80% of the occurrences between 95.60 to 106.88.

UCL = upper control limit
LCL = lower control limit

Some could rightfully challenge that the data shown in Figure 10 are not normally distributed. People who take this position might conclude that a nonconformance estimate should not be made from the probability plot. A non-normality conclusion for this data set can't be disagreed with, especially because, by the nature of this simulated example, there are two components of variation—between and within subgroup variability. However, George Box[19] noted that while all models may be wrong, some are useful. For this illustrative example, take Box's position that this model is wrong but, visually, the straight line in the probability plot appears to provide a fair representation of the population from which the data were taken. Therefore, the estimate is useful.

In general, other types of probability plots—such as a log-normal probability plot—may provide a better 30,000-foot-level performance estimate than a normal probability plot. But take care that the selected probability plot makes physical sense. For example, data from a situation that has a lower bound of zero (for example, hold time in a call center) might better be represented by a log-normal probability plot than a normal probability plot.

Figures 11 and 12 (p. 31) illustrate a 30,000-foot-level report out for the Sample 1 individual values data (Table 3), with and without a specification.

If a process is not capable of meeting specification needs, improvements could be implemented, such that the process' 30,000-foot-level individuals control chart transitions to a new, improved level of performance. This may seem obvious; however, often organizational metric reporting does not encourage this form of thinking. When an organization reacts to the problems of the day or week, the group is not structurally assessing this form of problem solving. Typically, 30,000-foot-level metric reporting includes a measurement variability component, which could provide the largest opportunity for improving the high-level reported performance, if something can be done to reduce the amount of variation caused by individual measurement reporting differences.

## A higher level

At the 30,000-foot level, control charts are used to examine processes from a high viewpoint, which is consistent with Deming's philosophy. His statement that 94% of the troubles in a process belong to the system (common cause); only 6% are special cause suggests an individuals control chart might illustrate to management and others that past firefighting activities have been the result of common cause issues. Because most of these issues were not special-cause issues, this expensive approach to issue resolution had no lasting value.

The 30,000-foot-level performance metric tracking approach does not view short-term typical process variability as special-cause excursion control chart issues but as noise to the overall process response. If the process is shown to be predictable and the process overall does not yield a desired response, there is a "pull" for project creation.

Within the improvement projects, inputs and their levels are examined collectively over the process's stable or predictable period of time to determine opportunities for process improvement. If an organization chose the process that produced the nonconformance rate shown in Figure 10 for targeting their improvement efforts, an analysis of the data in Table 2 would show that there is significantly more variability between days than within days. This insight could be helpful to determine what should be changed in the process so that the process's 30,000-foot-level chart is transitioned to a new, enhanced level of performance after the change was made. σ

### REFERENCES

1. Walter A. Shewhart, *Economic Control of Quality of Manufactured Product*, D. Van Nostrand Co., 1931.
2. Walter A. Shewhart, *Statistical Method From the Viewpoint of Quality Control*, United States Department of Agriculture, 1939.
3. Western Electric, *Statistical Quality Control Handbook*, Western Electric Co., 1956.
4. W. Edwards Deming, *Out of the Crisis*, Massachusetts Institute of Technology, 1986.
5. Automotive Industry Action Group, *Advanced Product Quality Planning and Control Plan: Reference Manual*, 1995.
6. Ibid.
7. Joseph M. Juran, Frank M. Gryna and Richard S. Bingham, *Quality Control Handbook*, third edition, McGraw-Hill, 1976.
8. Davis R. Bothe, *Measuring Process Capability: Techniques and Calculations for Quality and Manufacturing Engineers*, McGraw-Hill, 1997.
9. Forrest W. Breyfogle III, *Integrated Enterprise Excellence Volume III—Improvement Project Execution: A Management and Black Belt Guide for Going Beyond Lean Six Sigma and the Balanced Scorecard*, Bridgeway Books/Citius Publishing, 2008.
10. Ibid.
11. Thomas Pyzdek, "How Do I Compute σ? Let Me Count the Ways," *Quality Digest*, May 1998.
12. Minitab, Minitab Statistical Software, Release 15, 2007.
13. Automotive Industry Action Group, *Advanced Product Quality Planning and Control Plan: Reference Manual*, see reference 5.
14. Pyzdek, "How Do I Compute σ? Let Me Count the Ways," see reference 11.
15. W. Edwards Deming, *Some Theory of Sampling*, Dover Publications, 1966.
16. Deming, *Out of the Crisis*, see reference 4.
17. Breyfogle, *Integrated Enterprise Excellence Volume III—Improvement Project Execution: A Management and Black Belt Guide for Going Beyond Lean Six Sigma and the Balanced Scorecard*, see reference 9.
18. Ibid.
19. George E. P. Box and Norman R. Draper, *Empirical Model-Building and Response Surfaces*, Wiley, 1987.

### BIBLIOGRAPHY

Montgomery, Douglas C., *Introduction to Statistical Quality Control*, Wiley, 1985.