3.4 PER MILLION

BY FORREST W. BREYFOGLE III

Insight or Folly?

Resolve issues with process capability indexes, business metrics

IN LEAN SIX SIGMA, much training effort is spent on conveying the importance of having a measurement system so that consistent and correct decisions are made relative to assessing part quality and process attributes. In this training, measurement systems analysis (MSA) and associated gage repeatability and reproducibility (R&R) studies are integral.

It seems MSA should be a reporting consideration for all forms of measurement, including business performance metrics. It also seems a focus should be placed on metric statements that are in a clear language. In the real world, however, how often are these goals achieved?

Every organization's goal should be to achieve the three Rs of business: Everyone does the right things in the right way at the right time. One tool that provides direction for the three Rs goal is process performance metrics; that is, a process performance report-out should lead to the most appropriate action or inaction, which is independent of the person compiling the information. This basic right-behavior objective is like an inspection gage MSA, which ensures inspectors can adequately determine whether a manufactured component should be accepted or rejected.

Because of this performance-reporting need, it seems that management and practitioners—from a conceptual MSA point of view—would be assessing the health of current scorecard and metric reporting systems. However, this doesn't seem to occur. The question is: Why don't we examine business metrics and process capability indexes reporting from an MSA point of view with the same level of intensity we do for product quality metrics?

To quantify the magnitude of this issue, consider different approaches someone might use to report process performance. You might choose a bar chart, a pie chart, a red-yellow-green scorecard, process capability indexes (that is, C_p, C_{pk}, P_p, P_{pk} and C_{pm}) or a table of numbers. For a given process, each of these reporting methods can provide a very different and somewhat subjective picture of how a process performs and whether any actions should be taken.

To illustrate the magnitude of this issue, an example later in this column will show how reported C_p, C_{pk}, P_p and P_{pk} process capability indexes can be sensitive to process sampling procedures—a conceptual MSA issue.

In addition, these reports describe historically what happened, which may not be representative of the future. What is really desired is a statement about what is expected in the future so changes can be made, if needed.

Metric reporting should lead to the most appropriate action or inaction. However, process-metric decisions are often a function of how an individual or a group chose its process sampling, data analysis and reporting procedures. From a conceptual MSA point of view, the reporting of process performance should be independent of the person doing the sampling and reporting. Also, it is most desirable to use a predictive measurement system in which the only difference between individual process reporting is from chance sampling variability.

Organizations benefit when managers use predictive measurement reporting throughout their business functional process map. Practitioners can enhance the understanding of the benefits of this system when providing illustrative report-outs that compare current scorecard metric reporting to predictive-performance metric reporting system.

Process capability indexes

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:

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

in which USL and LSL are the upper and

Process capability standard deviation calculations / TABLE 1

	X chart	$\bar{x}$ and R	Random sample
C_p and C_{pk}	$\hat{\sigma} = \frac{\overline{MR}}{d_2} = \frac{\overline{MR}}{1.128}$	$\hat{\sigma} = \frac{\overline{R}^{*}}{d_2}$	N/A
P_p and P_{pk}	$\hat{\sigma} = \sqrt{\sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{N-1}}$	$\hat{\sigma} = \sqrt{\sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{N-1}}$	$\hat{\sigma} = \sqrt{\sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{N-1}}$

* Statistical computer programs will sometimes pool standard deviations for unbiasing reasons when there are *m* subgroups of sample size *n*, resulting in a slightly different value for standard deviation.

lower specification limits, respectively, and the spread of the distribution is described as six times standard deviation; that is, 6σ.

C_p addresses only the spread of the process; C_{pk} is used to address the spread and mean (μ) shift of the process concurrently. Mathematically, C_{pk} can be represented as the minimum value of the two quantities:

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

P_p and P_{pk} indexes are sometimes referred to as long-term capability or performance indexes. The relationship between P_p and P_{pk} is similar to that between C_p and C_{pk}. The index P_p represents the allowable tolerance spread relative to the actual spread of the data when the data follow a normal distribution. This equation is:

$$P_p = \frac{\mathrm{USL}-\mathrm{LSL}}{6\sigma}$$

in which, again, USL and LSL are the upper and lower specification limits. No quantification for data centering is described within this P_p relationship.

Mathematically, P_{pk} can be represented as the minimum value of the two quantities:

$$P_{pk} = \min\left[\frac{\mathrm{USL}-\mu}{3\sigma}, \frac{\mu-\mathrm{LSL}}{3\sigma}\right]$$

Consider the confusion in calculating the seemingly simple statistic of standard deviation. Although standard deviation is an integral part of calculating process capability, the method used to calculate the value rarely seems to be 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 sources of continuous data:

- Situation one: An $\bar{x}$ and R control chart with subgroups of sample size of 5.
- Situation two: An X chart with individual measurements.
- Situation three: A random sample of measurements from a population.

Sample data / TABLE 2

Subgroup	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
1	102.7	102.2	102.7	103.3	103.6
2	108.2	108.8	106.7	106.6	109.1
3	101.9	103.0	100.6	101.4	101.3
4	103.9	105.5	104.3	104.5	104.5
5	97.2	99.0	96.5	94.9	96.5
6	94.4	93.0	93.0	95.2	93.6
7	104.7	103.6	103.7	104.7	104.5
8	102.5	102.7	101.2	100.6	103.1
9	101.9	103.1	101.0	101.2	101.4
10	95.0	95.3	95.3	94.4	94.2

For these three situations, C_p, C_{pk}, P_p and P_p, a standard deviation estimate ($\hat{\sigma}$) is determined through the relationships in Table 1.

In Table 1, $\bar{x}$ is the overall sample mean, x_i is an individual sample (i) from a total sample size N, $\bar{R}$ is the mean subgroup range, $\overline{MR}$ is the mean range between adjacent subgroups, N is the total sample size, and d_2 is a factor for constructing variables control charts; that is, $d_2 = 1.128$ for a two-observation sample and 2.326 for a five-operation sample.

The data set in Table 2 illustrates the impact different data collection techniques can have on reported process capability metrics. When reporting process capability indexes, it is important that the data from which the metric is calculated comes from a stable process. In other words, the process is in control.[1]

To quantify the capability of this process, you could have selected only one sample instead of five for each subgroup. These two scenarios would result in the following standard deviation calculations:

$$\hat{\sigma} = \sqrt{\sum_{i=1}^{N}\frac{(x_i-\bar{x})^2}{n-1}} = 4.40$$

$$\hat{\sigma} = \frac{\overline{MR}}{d_2} = \frac{\overline{MR}}{1.128} = 4.26$$

$$\hat{\sigma} = \frac{\bar{R}}{d_2} = \frac{\bar{R}}{2.326} = 0.90$$

In the second standard deviation calculation, consider that sample one in Table 2 was the individual reading for each subgroup.

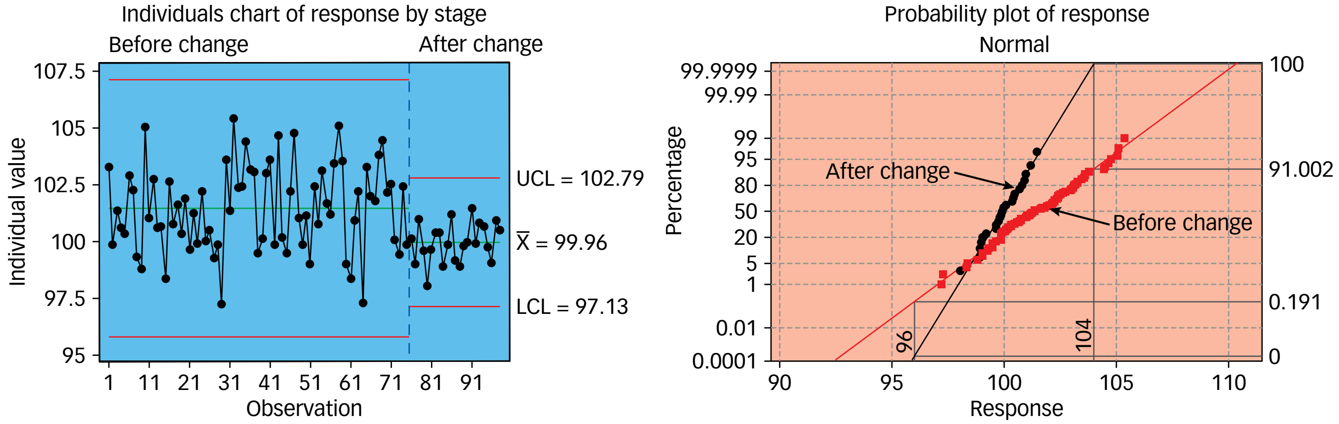
For a specification of 95 – 105, a statistical analysis program used similar standard deviations when determining the process capability results in Table 3.

In Table 3, there is large difference between the C_p and C_{pk} values for a subgrouping sample size of one sample versus five. An examination of the standard deviation equations provides the reason for the large difference: C_p and C_{pk} calculations—which used an $\bar{x}$ and R chart—had their equation's standard

Process capability results—sample data / TABLE 3

Subgroups	C_p	C_{pk}	P_p	P_{pk}
Five samples	1.82	1.44	0.38	0.30
One sample	0.39	0.29	0.38	0.28

Process response baseline, demonstrated improvement for continuous response / FIGURE 1



For specification of 100 +/-4
Before change: Process was stable with an approximate nonconformance rate of 9.2%.
After change: Process is predictable (stable) with an approximate nonconformance rate of 0%.

UCL = upper control limit
LCL = lower control limit

deviation determined by averaging within subgroups. For the individual's chart, standard deviation was calculated between subgroups.

Conceptual MSA issues

With an effective conceptual MSA system, process sampling plans should have no effect on process performance statements.

Because of the differences noted earlier, you can conclude process capability reporting can have MSA issues because a sample of five versus one did not provide similar answers; that is, the differences being only the result of luck-of-the-draw sampling.

Statistical business performance charting action options / TABLE 4

1. Is the process unstable, or did something out of the ordinary occur, which requires action or no action?
2. Is the process stable and meeting internal and external customer needs? If so, no action is required.
3. Is the process stable but does not meet internal and external customer needs? If so, process improvement efforts are needed.

In this analysis, you might notice P_p and P_{pk} are similar for the two sampling procedures. An $\bar{x}$ and R control chart analysis, however, would indicate the process was out of control; therefore, a process capability analysis would not be appropriate for this form of control-charting analysis.[2]

Other conceptual MSA issues with process capability indexes reporting include:

- The physical implication of reported process capability indexes is uncertain and possibly wrong.
- Without an accompanying statement of process stability from a control chart, all process capability indexes are questionable in value. Any process capability assessment of an unstable process is improper and often deceptive.
- Process capability indexes do not provide a predictive performance statement.

Predictive reporting alternative

From a conceptual MSA point of view, Table 4 describes three reasons for statistical business performance charting (SBPC), or 30,000-foot-level[3] tracking, and reporting for transactional and manufacturing process outputs.

Process performance reporting using process capability indexes, bar charts, pie charts, red-yellow-green scorecards or tables of numbers can provide differing process performance assessments, a conceptual MSA issue. In addition, process performance reporting does not structurally address the action options in Table 4.[4]

The following example illustrates a system for describing a process-output performance from a 30,000-foot-level. For this SBPC reporting, an individual's control chart subgrouping frequency is made so typical variability from input variables occurs between subgroups.

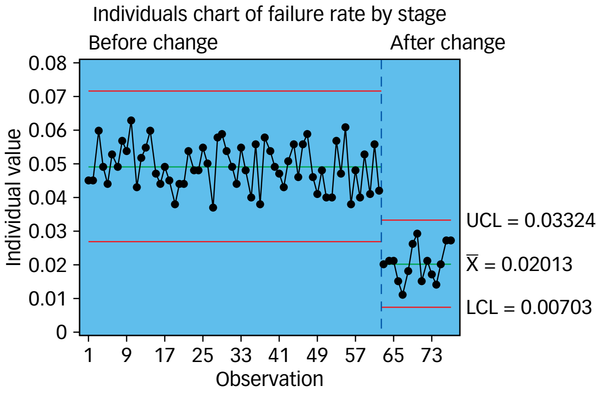
Data from regions of stability can be used to estimate the nonconformance rate of a process during those timeframes. If there is a recent region of stability, data from this region can be considered a random sample of the future from which a prediction can be made. This prediction presumes no fundamental positive or negative changes will occur in the future relative to the process inputs or its execution steps.

If, at some point in time, the output of a stable process is performing at an undesirable, nonconformance level, an organization can initiate a project to change process inputs or take steps to improve a process performance level.

For continuous data, a probability plot can provide an estimate of the process nonconformance rate in either percentage or defects per million opportunities units. For attribute data, the process estimated nonconformance rate is simply the overall combined subgroup failure rates in the region of process stability.

Figures 1 and 2 illustrate 30,000-foot-level reporting and the results of an improvement project for both continuous and attribute data.[5]

Process response baseline, demonstrated improvement for attribute response / FIGURE 2



For equal subgroup sizes of 1,000:
Before change: Process was stable with an approximate nonconformance rate of 5%.
After change: Process is predictable (stable) with an approximate nonconformance rate of 2%.

UCL = upper control limit
LCL = lower control limit

Integrating SBPC

When reporting how a process is performing using capability indexes, the magnitude of the reported metrics for a given situation can be a function of sampling procedures. For example, different conclusions could be made when process data are analyzed from an individuals chart report-out (one sample per subgroup) versus a $\bar{x}$ and R chart reporting (multiple samples per subgroup); that is, a conceptual process performance MSA issue.

Traditional organizational performance measurement reporting systems can use tables of numbers, stacked bar charts, pie charts and red-yellow-green, goal-based scorecards. For a given situation, one person may choose one reporting scheme, while another uses a completely different approach. These differences can lead to a different conclusion about what is happening and should be done, as shown in Table 4.

In addition, these traditional reporting methods provide only an assessment of historical data and make no predictive statements. Using this form of metric reporting to run a business is similar to driving a car by only looking at the rearview mirror—a dangerous practice.

When predictive SBPC system reporting is used to track interconnected business process map functions, an alternative, forward-looking dashboard performance reporting system becomes available. With this 30,000-foot-level metric system, organizations can systematically evaluate future expected performance and make appropriate adjustments if they don't like what they see.

Organizations can benefit when SBPC techniques are integrated within a business system that analytically and innovatively determines strategies with the alignment of improvement projects that positively impact the overall business.[6] QP

REFERENCES AND NOTES

1. Forrest Breyfogle III, "Control Charting at the 30,000-foot-level," *Quality Progress,* November 2003, pp. 67-70. The data used in Table 2 were also used in the 2003 article to compare traditional $\bar{x}$ and R process stability assessment to 30,000-foot-level operational-metric reporting. In the 2003 article, a traditional control chart indicated that the process was out of control, while 30,000-foot-level reporting indicated the process was in control. The article also described the advantages of a 30,000-foot-level assessment when compared to traditional reporting.
2. Ibid. The article also notes technical reasons why individual control charting is preferred over $\bar{x}$ and R control charting.
3. Forrest Breyfogle III, *Integrated Enterprise Excellence Volume II—Business Deployment: A Leaders' Guide for Going Beyond Lean Six Sigma and the Balanced Scorecard,* Bridgeway Books, 2008.
4. Ibid.
5. Forrest Breyfogle III, "Control Charting at the 30,000-foot-level, Part 2" *Quality Progress*, November 2004, pp. 85-87. This article describes the advantages of attribute-failure-rate individuals charting (as shown in Figure 2) over a p-chart.
6. Forrest Breyfogle III, "Control and Grow Your Enterprise," *Quality Progress*, February 2009, pp. 54-56.

FORREST W. BREYFOGLE III is founder and CEO of Smarter Solutions, Inc. in Austin, TX. He earned a master's degree in mechanical engineering from the University of Texas. Breyfogle is the author of a series of books on the Integrated Enterprise Excellence System. He is an ASQ fellow and recipient of the 2004 Crosby Medal.

About the Author
Forrest Breyfogle, III
Integrated Enterprise Excellence

In a professional career spanning over a quarter century, Forrest Breyfogle has established himself as a leading edge thinker, a prolific author, an innovative consultant, a world-class educator, and a successful business executive. His work is documented in eleven books and over ninety articles on the topic of quality improvement.

A professional engineer, Forrest is also a member of the board of advisors for the University of Texas Center for Performance Excellence. He is the founder and CEO of Smarter Solutions, Inc., an Austin, Texas based consulting firm offering business measurement and improvement consultation and education to a distinguished list of clients worldwide, including BAMA, CIGNA, Dell, HP, IBM, Oracle Packaging, Sherwin Williams, Cameron, TIMET, and TATA. He served his country on active duty in the US Army for 2 years, and has played an active leadership role in professional and educational organizations. Forrest received the prestigious Crosby Medal from the American Society for Quality (ASQ) in 2004 for his book, *Implementing Six Sigma* (second edition). This award is presented annually by the American Society for Quality to the individual who has authored a distinguished book contributing significantly to the extension of the philosophy and application of the principles, methods, or techniques of quality management

He is a widely recognized authority in the field of management improvement and is a frequent speaker before professional associations and businesses. His earlier work in the field of management science has been widely acclaimed. A previous book, *Implementing Six Sigma*, sold over 40,000 copies and still ranks among the top Amazon books in Applied Mathematics/Engineering Statistics and Industrial Engineering /Quality Control.

He founded Smarter Solutions in 1992 after a 24-year career at IBM. The associates of Smarter Solutions specialize in helping companies throughout the world improve their bottom line and customer satisfaction through the implementation of techniques that are beyond traditional Lean Six Sigma and the balanced scorecard methodologies. His latest and most extensive work has been in the documentation of a new system of enterprise management, the Integrated Enterprise Excellence (IEE) system, in a series of four books. IEE provides a detailed roadmap that builds on and integrates the best practices of earlier disciplines like Six Sigma, Lean, TQM , PDCA, DOE, and TPS combined with innovative analytical tools to produce improvements at the highest level of an enterprise.

In addition to assisting hundreds of major clients in the wise implementation of improvement systems worldwide, Forrest has also developed over 300 hours of classroom instruction used to train executives, managers, and Black Belt practitioners to plan for, implement, and manage IEE systems. He also leads formal seminars and workshops worldwide.

Forrest Breyfogle
forrest@smartersolutions.com
512-918-0280 x401
www.smartersolutions.com