IE 361 Module 17

Process Capability Analysis: Part 1

Prof.s Stephen B. Vardeman and Max D. Morris

Reading: Section 5.1, 5.2 *Statistical Quality Assurance Methods for Engineers*

Normal Plotting for Process Characterization

If (by virtue of process monitoring and wise intervention) one is willing to say that a data set represents a *stable process*, it may be used to characterize process output. Section 5.1 discusses several graphical techniques for summarizing a sample and therefore representing the process that stands behind it. Here we emphasize one of these, so called "normal plotting," a tool for invesitgating the extent to which a data set (and thus the process that produced it) can be described using a normal distribution.

Normal plots are made using so called **quantiles**. The p quantile (or $100 \times p$th percentile) of a distribution is a number such that a fraction p of the distribution lies to the left and a fraction $1 - p$ lies to the right. If one scores at the .8 quantile (80th percentile) on an exam, 80% of those taking the exam had lower marks and 20% had higher marks. Or, since 95% of the standard normal distribution is to the left of 1.645, 1.645 is the .95 quantile of that distribution. We will use the notation $Q(p)$ to stand for the p quantile of any distribution.

For a data set consisting of n values $x_1 \leq x_2 \leq \cdots \leq x_n$ (x_i is the ith smallest data value), we'll adopt the convention that x_i is the $p = (i - .5)/n$ quantile of the data set, that is

$$Q_{\text{data}}\left(\frac{i - .5}{n}\right) = x_i$$

For $Q_z(p)$ the standard normal quantile function, a normal plot is then made by plotting ordered pairs

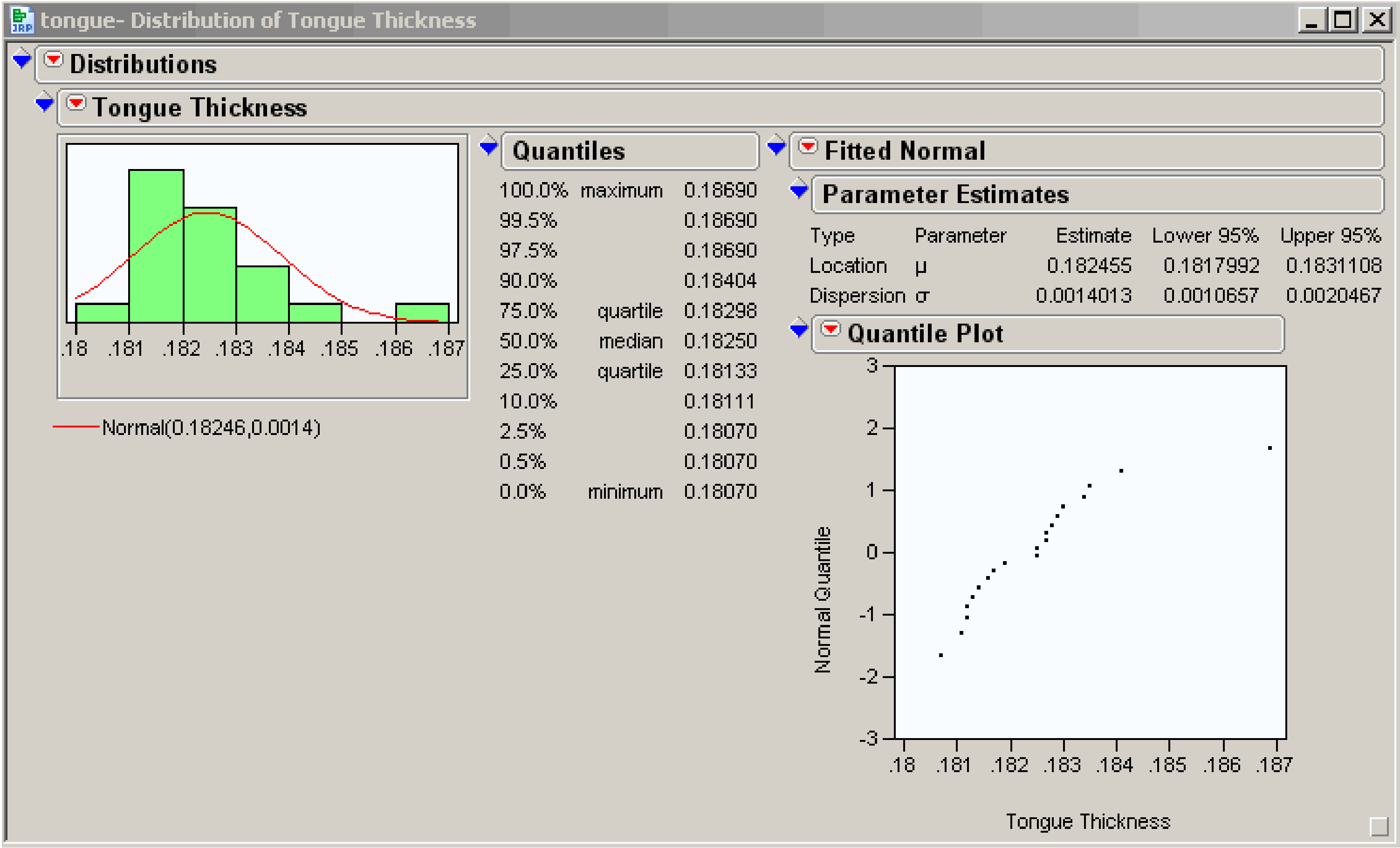
$$\left(Q_{\text{data}}\left(\frac{i - .5}{n}\right), Q_z\left(\frac{i - .5}{n}\right)\right)$$

i.e.

$$\left(x_i, Q_z\left(\frac{i - .5}{n}\right)\right)$$

(Standard normal quantiles $Q_z(p)$ can be found by locating values of p in the body of a typical cumulative normal probability table and then reading corresponding quantiles from the table's margin. And statistical packages like `JMP` provide "inverse cumulative probability" functions and "normal plotting" functions that can be used to automate this.) This plot allows comparison of data quantiles and (standard) normal ones. A "straight line" normal plot indicates that a data set has the same shape as the normal distributions, and suggests that the process that stands behind the data set can be modeled as producing normally distributed observations. (Section 5.1 has a careful discussion of interpretation of such Q-Q plots for those who need a review of this Stat 231 material.)

Example 17-1 Table 5.7 of *SQAME* contains measured "tongue thicknesses" for $n = 20$ steel levers. Figure 1 shows a `JMP` report including a normal plot for the data of Table 5.7.



tongue- Distribution of Tongue Thickness

Distributions

Tongue Thickness

Quantiles

100.0%	maximum	0.18690
99.5%		0.18690
97.5%		0.18690
90.0%		0.18404
75.0%	quartile	0.18298
50.0%	median	0.18250
25.0%	quartile	0.18133
10.0%		0.18111
2.5%		0.18070
0.5%		0.18070
0.0%	minimum	0.18070

Fitted Normal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	0.182455	0.1817992	0.1831108
Dispersion	σ	0.0014013	0.0010657	0.0020467

Quantile Plot

Figure 1: `JMP` Report for the Data of Table 5.7 of *SQAME* Including a Normal Plot

The normal plot in Figure 1 shows that the largest tongue thickness in the data set is much too large to "fit" with the other observations. It would have to be pulled substantially "back to the left" in order to make a linear plot. There is some important departure from a "normal"/Gaussian shape indicated on this plot.

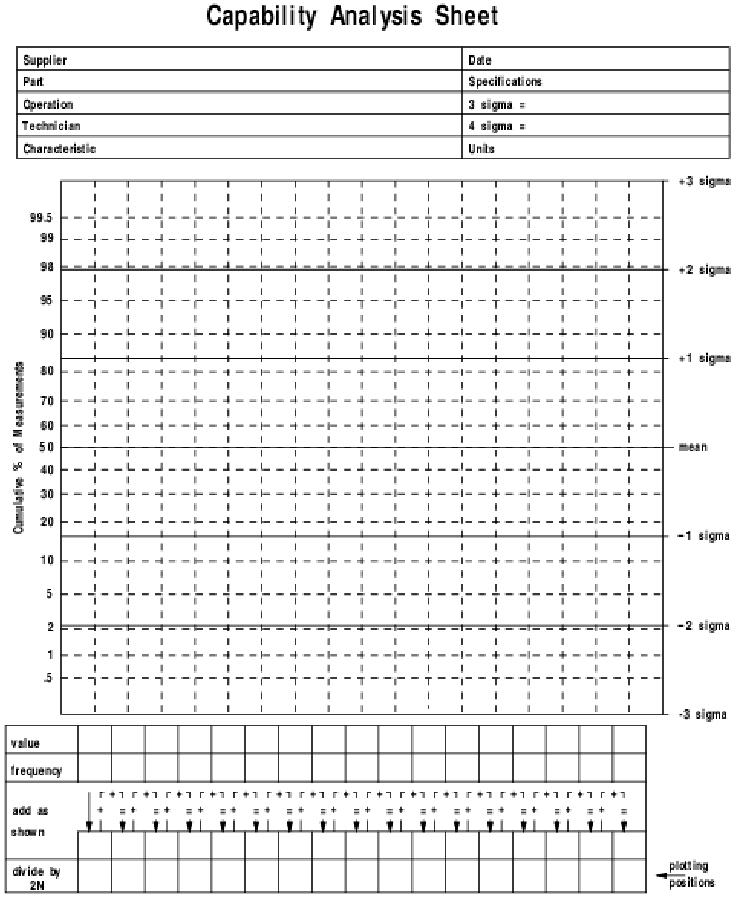
Probability plotting is important for several reasons. First, it helps one judge how much faith to place in calculations based on a normal distribution, and suggests in what ways the calculations might tend to be wrong. For example, the normal plot for the tongue thicknesses suggests that if the mechanism that operated to produce the single very large value is truly "part of the process," using a normal distribution to describe manufactured thickness will likely underpredict the frequency of large data values.

Probability plotting is also sometimes helpful is in providing graphical estimates of distribution parameters. For example, if one makes a normal plot of an exactly normal distribution, the slope of the plot is the reciprocal of σ and the horizontal intercept is μ. That suggests that for a real data set whose normal plot is fairly linear,

1. the horizontal intercept of an approximating line is a sensible estimate of the mean of the process generating the data, and

2. the reciprocal of the slope is a sensible estimate of the standard deviation of the process generating the data.

The facts that (for bell-shaped data sets) normal plotting provides a simple way of approximating a standard deviation and that 6σ is often used as a measure of the intrinsic spread of measurements generated by a process, together lead to the common practice of basing **process capability analyses** on normal plotting. The next figure shows a very common type of industrial form that essentially facilitates the making of a normal plot by removing the necessity of evaluating the standard normal quantiles $Q_z(p)$. (On the special vertical scale one may simply use the plotting position p rather than $Q_z(p)$, as would be required when using regular graph paper.) After plotting a data set and drawing in an approximating straight line, 6σ can be read off the plot as the difference in horizontal coordinates for points on the line at the "$+3\sigma$" and "-3σ" vertical levels (i.e., with $p = .0013$ and $p = .9987$).



Capability Analysis Sheet

Supplier	Date
Part	Specifications
Operation	3 sigma =
Technician	4 sigma =
Characteristic	Units

Figure 2: A "Capability Analysis Sheet" (That is Essentially a Piece of Normal Probability Paper) (Page 211 *SQAME*)

Forms like the one in the figure encourage the *plotting* of process data (always a plus) and also allow even fairly nonquantitative people to easily estimate and develop some intuition about "the process spread."

Process Capability Measures and Their Estimation

Graphical methods provide a visual representation of the pattern of variation associated with a process. Often, in addition to these tools, it is convenient to have some numerical "capability index" summary measures to quote.

We discuss the "process capability" and the two "capability ratios," C_p and C_{pk}, and methods for making confidence intervals for them. But it is important to begin with a disclaimer: *Unless a normal distribution makes sense as a description of process output, these measures are of dubious relevance. Further, the confidence interval methods presented here are completely unreliable unless a normal model is appropriate. So the normal plotting idea just presented is a very important prerequisite for using these methods.*

It is well known that the majority of a normal distribution is located within three standard deviations of its mean. The following figure illustrates this elementary point, and in light of the picture, it makes some sense to say that (for a normal distribution) 6σ is a measure of process spread, and to call 6σ the **process capability** for a stable process generating normally distributed measurements.

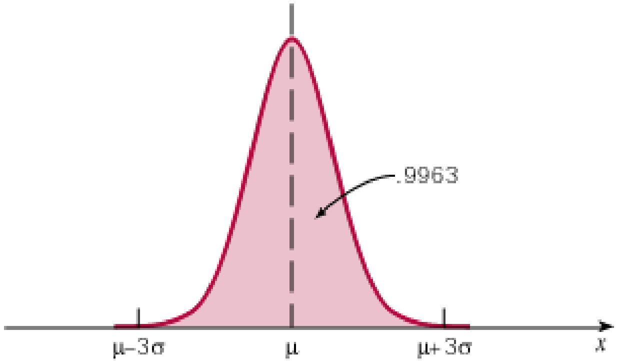


Figure 3: "Most" of a Normal Distribution is Within 3σ of the Mean μ

The fact that there are methods for estimating the standard deviation of a normal distribution implies that it is easy to give confidence limits for the process capability. That is, if one has a sample of n observations with corresponding sample standard deviation s, then confidence limits for 6σ are simply 6 times the limits for σ (met first in Stat 231 and used in this course beginning already in Module 2) namely

$$6s\sqrt{\frac{n-1}{\chi^2_{\text{upper}}}} \text{ and/or } 6s\sqrt{\frac{n-1}{\chi^2_{\text{lower}}}}$$

where χ^2_{upper} and χ^2_{lower} are upper and lower percentage points of the χ^2 distribution with $n-1$ degrees of freedom.

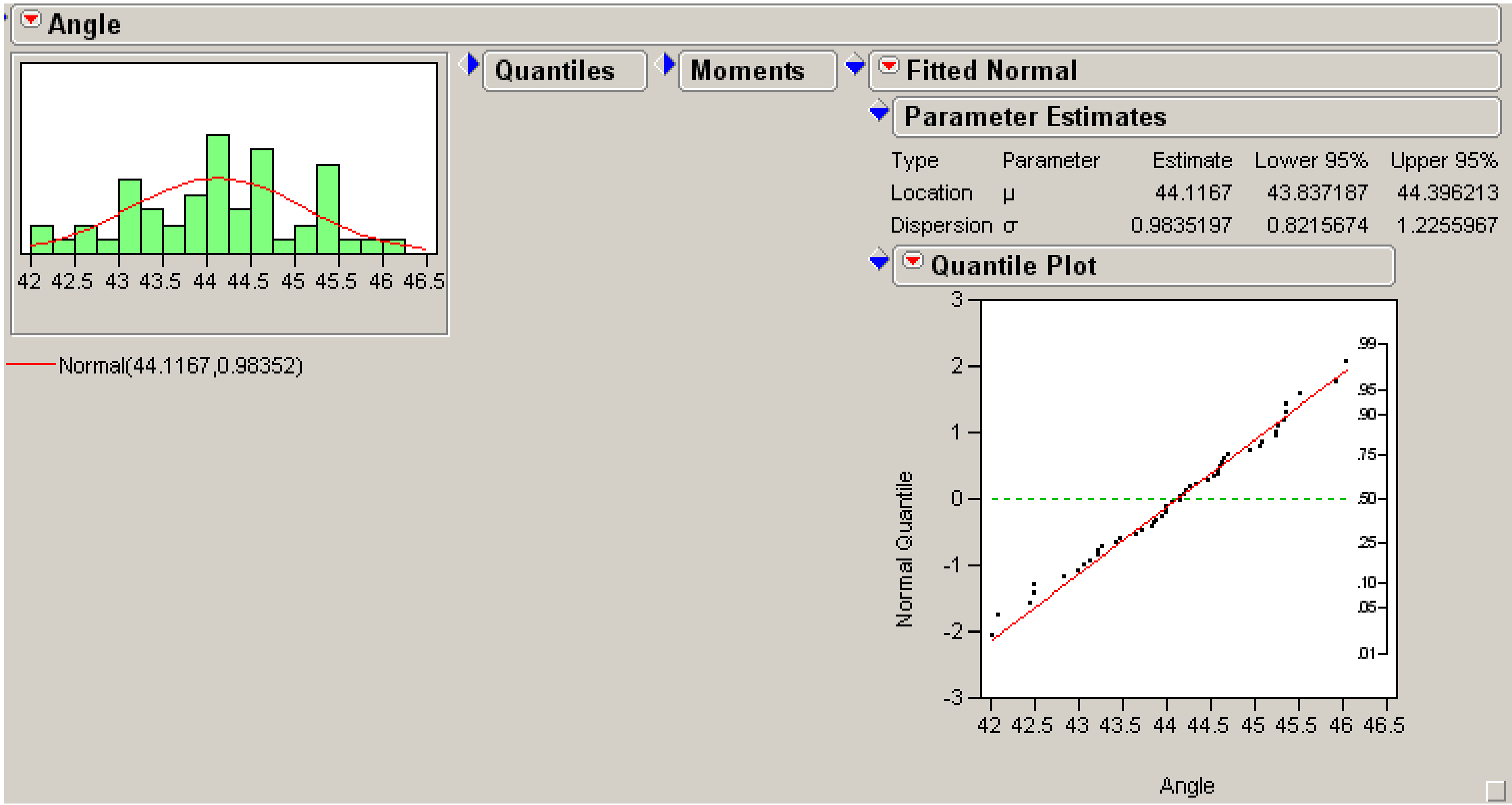
Example 17-1 An IE 361 group did some measuring of angles with a flat surface made in the EDM drilling of holes on a high precision metal part. The $n = 50$ data values they collected are on page 209 of *SQAME*. The sample mean for these data is $\bar{x} = 44.117^\circ$ and the sample standard deviation is $s = .984^\circ$. The next figure is a `JMP` report for these data. It includes a normal plot for the data that (as it is very linear) indicates that a normal model for angles produced by this process is quite sensible. It also includes 95% confidence limits for σ, namely

$$.822^\circ \text{ and } 1.226^\circ$$

These limits translate to limits

$$6\,(.822) = 4.929^\circ \text{ and } 6\,(1.226) = 7.536^\circ$$

for the "process capability."



Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	44.1167	43.837187	44.396213
Dispersion	σ	0.9835197	0.8215674	1.2255967

Figure 4: `JMP` Report for the Hole Angle Data of Table 5.7 of *SQAME*

Where there are both an upper specification U and a lower specification L for measurements generated by a process, it is common to compare process variability to the spread in those specifications. One way of doing this is through **process capability ratios.** And a popular process capability ratio is

$$C_p = \frac{U - L}{6\sigma}$$

When this measure is 1, process output will fit more or less exactly inside specifications *provided the process mean is exactly on target at* $(U + L)/2$. When C_p is larger than 1, there is some "breathing room" in the sense that a process would not need to be perfectly aimed in order to produce essentially all measurements inside specifications. On the other hand, where C_p is less than 1, no matter how well a process producing normally distributed observations is aimed, a significant fraction of the output will fall outside specifications.

The very simple form of C_p makes it clear that once one knows how to estimate 6σ, one may simply divide the known difference in specifications by confidence limits for 6σ in order to find confidence limits for C_p. That is, lower and upper confidence limits for C_p are respectively

$$\frac{(U-L)}{6s}\sqrt{\frac{\chi^2_{\text{lower}}}{n-1}} \quad \text{and/or} \quad \frac{(U-L)}{6s}\sqrt{\frac{\chi^2_{\text{upper}}}{n-1}}$$

where again χ^2_{upper} and χ^2_{lower} are upper and lower percentage points of the χ^2 distribution with $n-1$ degrees of freedom.

Example 17-1 continued Specifications on the angles in the EDM drilling application were $45^\circ \pm 2^\circ$. That means that for this situation $U - L = 4^\circ$. Based on the measurement of $n = 50$ parts, the students found $s = .984$ and 95% two-sided confidence limits for 6σ of 4.929° and 7.536°. Thus, one can be 95% confident that C_p is between

$$\frac{4}{7.536} \quad \text{and} \quad \frac{4}{4.929}$$

that is, between

$$.53 \quad \text{and} \quad .81.$$

The main message conveyed by this interval is that even at its most optimistic (C_p as large as .81) this EDM drilling process is not capable of meeting the $\pm 2^\circ$ engineering specifications.

C_p is more a measure of process *potential* than it is a measure of current performance. Since process aim is not considered in the computation of C_p, it is possible for a misaimed process with very small intrinsic variation to have a huge value of C_p and yet currently be turning out essentially no product in specifications. C_p attempts only to measure "what could be" were the process perfectly aimed. This is not necessarily an undesirable feature of C_p, but it is one that users need to understand.

Another process capability index that does take account of the process mean (and is more a measure of current process performance than of potential performance) is C_{pk}. This measure can be described in words as "the number of 3σ's that the process mean is to the good side of the closest specification." For example, if $U - L$ is 10σ, and μ is 4σ below the upper specification, then C_{pk} is $4\sigma/3\sigma = 1.33$. On the other hand, if $U - L$ is 10σ and μ is 4σ above the upper specification, then C_{pk} is -1.33.

In symbols,

$$C_{pk} = \min\left\{\frac{U-\mu}{3\sigma}, \frac{\mu - L}{3\sigma}\right\} = \frac{U - L - 2\left|\mu - \frac{U+L}{2}\right|}{6\sigma}$$

This quantity will be positive as long as μ is between L and U. It will be large if μ is between L and U (preferably centered between them) and $U - L$ is large compared to σ. It is always true that

$$C_{pk} \leq C_p$$

and the two measures are equal only when $\mu = (U + L)/2$ exactly.

The best currently available confidence interval method for C_{pk} is only appropriate for large samples and provides a real confidence level that only approximates the nominal one. The method is based on the natural single number estimate of C_{pk},

$$\hat{C}_{pk} = \min\left\{\frac{U-\bar{x}}{3s}, \frac{\bar{x}-L}{3s}\right\} = \frac{U - L - 2\left|\bar{x} - \frac{U+L}{2}\right|}{6s}$$

Then for z an appropriate standard normal upper percentage point, approximate confidence limits for C_{pk} are

$$\hat{C}_{pk} \pm z\sqrt{\frac{1}{9n} + \frac{\hat{C}_{pk}^2}{2n-2}}$$

Example 17-1 continued The $n = 50$ EDM hole angles had corresponding sample mean $\overline{x} = 44.117$. So

$$\hat{C}_{pk} = \min\left\{\frac{47 - 44.117}{3(.984)}, \frac{44.117 - 43}{3(.984)}\right\} = \min\{.98, .38\} = .38.$$

So, for example, since the .975 quantile (upper 2.5% point) of the standard normal distribution is 1.96, approximate 95% confidence limits for C_{pk} are

$$.38 \pm 1.96\sqrt{\frac{1}{9(50)} + \frac{(.38)^2}{2(50) - 2}} = .38 \pm .12.$$

One can be approximately "95% sure" that C_{pk} for the angles in the EDM drilling process is between .26 and .50. This is a very disappointing situation, if this angle is really a critical one (if these engineering specifications are "real").

Overreliance upon process capability measures like the ones discussed here has come under a fair amount of criticism in the past few years. Critics have correctly noted that

1. 6σ, C_p, and C_{pk} have only dubious relevance when a process distribution is not normal,

2. "one-number summaries" like those discussed here can leave much unsaid about what a process is doing or even the shape of a distribution of measurements it is generating, and

3. the whole business of really going to work tuning a process, monitoring for and removing upsets, and determining what it is really "capable" of doing involves much more than the simple estimation of 6σ or one of the measures C_p or C_{pk}.

Further, the capability ratios C_p and C_{pk} depend upon specifications that are sometimes subject to unannounced change (even arbitrary change). This makes it difficult to know from one reporting period to the next what has happened to process variability if estimates of C_p or C_{pk} are all that are provided. It thus seems that for purposes of comparisons across time, of the measures discussed here, the simple process capability 6σ is most attractive.

Despite the criticism their use has received, the measures discussed here are very popular. Provided one understands their limitations and simply views them as one of many tools for summarizing process behavior, they have their place. But the wise engineer will not assume that computing and reporting one of these figures is in any way the last word in assessing process performance.