MINI PAPER
Process Capability Indices

By Robert H. Mitchell
3M Company

Process capability indices have been popular for over 20 years, since Joseph Juran[12] popularized the Capability Ratio (Cr) in his Quality Control Handbook. Eagerness to establish a single index to measure process capability has resulted in the proliferation of indices: Cp/Cpk, Vp/Vpk, Pp/Ppk, Cpm, Tz, %OOL, PPM, and more recently – "Generalized" Cp/Cpk. Abuse of these indices is well documented (Gunter)[8]. Though simple to compute, they can lead to scorekeeping by management, incorrect interpretations and tampering, with little or no product or process improvement. This article reviews the basics of estimating process capability and introduces new process capability indices.

Cp, Cpk

The original Capability Index published in Juran's Quality Control Handbook is defined as Tolerance Width divided by Process Capability.

Cp = Process Potential
= (Tolerance Width) / (Process Capability).

Juran defines Process Capability as six standard deviations for a process in statistical control. **Process Capability** = 6σ where σ is the in-control process standard deviation. The easiest method to determine process standard deviation is from the control chart of a stable process: Process Std. Dev. = R-bar/d_2, where the appropriate value of d_2 is read from a table for the subgroup size.

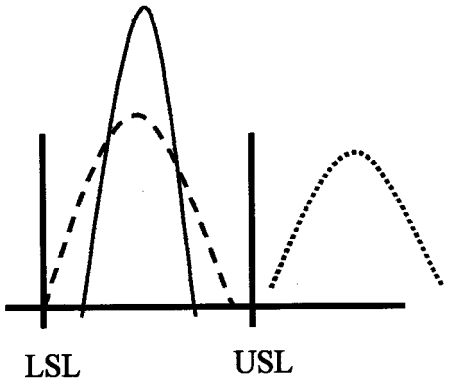
For the Cp/Cpk index to be valuable, the tolerance width, or specification range, should have real meaning, i.e. be based on functional limits per known end-user requirements ("Voice of the Customer"). Occasionally, for lack of end-user requirements, one is asked to set specifications (specs) based solely on the producer's process capability. The danger is that a process' aim and natural variation can have little relation to customer-perceived quality. How does one choose specs: use arbitrarily wide limits (+/- 6s) to ensure good Cpk values? Use arbitrarily narrow limits in an effort to lock out competition (and unwittingly, perhaps yourself)? Neither strategy focuses on the customer. Furthermore, focusing on specs alone asks, "What is the worst we can get away with?" while emphasis on target alignment and variability reduction asks, "What is the best we can do?"

According to Quality Assurance for the Chemical and Process Industries – A Manual of Good Practices 2nd Edition[1]:

> Values of Cp exceeding 1.33 indicate that the process is adequate to meet the specifications. Values of Cp between 1.33 and 1.00 indicate that the process, while adequate to meet specifications, will require close control. Values of Cp below 1.00 indicate the process is not capable of meeting specifications.

Cp is called the "Process Potential" - it simply relates the Process Capability (6σ) to the Spec Range; it does not relate the location of the process with respect to the specs. Consider the three distributions in fig. 1, all from processes having a Cp of 1.00 or better.

Figure 1



If the process is centered within the specs, and is approximately "normal" then Cp = 1.00 results in a fraction non-conforming (f.n.c.) of 0.27%. Note that Cp=1.00 doesn't guarantee that there will be only 0.27% non-conforming product. What it does guarantee is that, assuming normality, a stable process centered with respect to the specs, and the correct value of σ, there will not be less than 0.27% of non-conforming product.

Cpk = Process Capability Index.

CPK is used to summarize how a process is running relative to its spec limits. As with Cp, this measure is appropriate only when the process is stable (in-control).

$$\text{Cpk} = \text{minimum of} \left\{ \frac{\text{X bar - LSL, USL - X bar}}{3\sigma} \right\}$$

where Xbar and σ are the mean and standard deviation of the process. So, Cpk measures how far the process mean is from the nearer spec limit in terms of 3σ distances. For processes with a one-sided spec the term corresponding to the 'missing' limit is omitted.

MINI-PAPER

Continued from page 16

Special consideration must be made for non-normal distributions. Cpk works well only for the bell-shaped "normal" (Gaussian) distribution. For others it is an approximation. Unlike Xbar control charts there is no Central Limit Theorem effect in estimating process capability because Cp and Cpk relate to the distribution of **individual** items. We need to be concerned about non-normal distributions. Table 1 gives the fraction non-conforming (PPM) for different distributions, all having Cpk=1.

Table

Distribution	f.n.c.
Chi-Square (4.5df)	14,000
Heavy-tailed (B_2>3)	4,000
Uniform	0
Normal	2700

One approach to dealing with non-normal data is to transform the data. Typical transformations include taking the reciprocal, reciprocal square root, natural log, or square root of the raw data. The corresponding spec limits must be similarly transformed. Box, Hunter, and Hunter[5] offer a relatively simple method for determining a transformation to give constant variance.

Theoretical Cp, Cpk

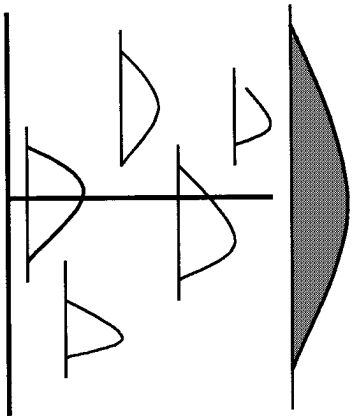
Cp* and **Cpk*** are calculated when the process is not stable, yet one desires to estimate how good the process might be if no Special Causes existed. **Theoretical Cp, Cpk** use a "best estimate" of the true process standard deviation (sigma-hat). Special Causes are excluded from the data when appropriate, to estimate the "potential" natural process variation. A theoretical process sigma-hat is calculated and Cp*/Cpk* estimated.

In general, I advise against this methodology because of potential confusion. Although meant as a measure of how good the process might be, one cannot predict since the process is not stable. These are not true Cp, Cpk values.

Pp, Ppk

Pp and Ppk are measures of process performance from a customer perspective. Process capability measures short term ability of a process to meet specs. Process performance measures long term ability of a process to meet specs. Process Performance should be distinguished from Process Capability. Pp is similar in definition to Cp, and Ppk is similar in definition to Cpk but in each we use the overall standard deviation of the data, including any Special cause variation, rather than the short term estimate of standard deviation. The overall standard deviation is a weighted average of both within-group and between-group variation. As a customer, one wants to measure performance based on total incoming variability. This is illustrated in figure 2.

Figure 2. Within subgroup Capability versus weighted average of both the within and between spread (Performance).



Cpm (Cp-Taguchi)

The **Cpm** index was introduced in 1988 (Chan[6]). The principal difference between Cpm and Cpk is the relative importance of Conformance to Specs versus Run to Target. Cpk measures how well the process mean is centered within the spec limits, and what percentage of product will be within spec. Instead of focusing on spec limits Cpm focuses on how well the process mean corresponds to the process target, which may or may not be midway between the spec limits. Cpm is motivated by Taguchi's "Loss Function". The denominator of Cpm includes the Root Mean Square deviation from the target.

$$\mathbf{Cpm} = \frac{USL - LSL}{6\{\sigma^2 + (Xbar - Tgt)^2\}^{1/2}}$$

Cpk is preferred to Cp because it measures both process location and process standard deviation. Cpm is often preferred to Cpk because the variability term used in the index is more consistent with Run to Target philosophy.

Cr

Capability ratio (Cr)[12, 13] is the inverse of Cp. If Cp = 1.33 or more is considered a capable process, then a value 0.75 or less is desired for Cr.

$$\mathbf{Cr} = \frac{6\sigma}{USL \quad LSL}$$

Tz

Target Z is a measure of targeting and is similar to, but simpler to estimate than Cpm.

$$\mathbf{Tz} = \frac{(Xbar - Target)}{\sigma}$$

Values of Tz between –0.5 and 0.5 are considered good.

MINI-PAPER

Continued from page 17

%OOL

Percent Outside of Limit uses the z-statistics to estimate the porportion of a population that lies beyond the spec limits. This measure assumes a Normal distribution. For a two-sided specification, the Z_{Lower} and Z_{Upper} proportions are calcuated from the process measn, standard deviation, and spec limits.

$$Z_{Lower} = \frac{Xbar - LSL}{\sigma}$$

$$Z_{Upper} = \frac{USL - Xbar}{\sigma}$$

Then look up the proportions from any Z-table. One-sided or nonsymmetrical specifications usually correspond to heavily skewed distributions and can yield exaggerated %OOL values.

PPM

Parts Per Million (defective) is similar to %OOL. One multiplies the Z_L and Z_U proportions by 1,000,000 each, then sum together. Again, these are theoretical estimates based on the assumption of a normal distribution. Table 2 illustrates the relation between Cpk and PPM. If receiving inspection is performed, then it is possible to compare PPM-Observed (from receiving inspection) and PPM-Calculated (from the Normal Distribution).

Table 2 PPM and Cpk fraction non-conforming (f.n.c.)

Cpk	PPM
.333	317,400
.667	45,500
1.0	2,700
1.33	63
1.50	7
1.67	0.6
2.00	2.0 PPB

*Note: This comparison of Cpk and PPM does not include the 1.5σ shift included in the "Motorola Six Sigma" program.

GCpk, GPpk

Traditional Cpk can only be estimated from a stable process; i.e., no Special Causes. Joe Voelkel[19], at the 1998 Fall Technical Conference, introduced Generalized Cp/Cpk ("GCp/GCpk"). Joe noted that there are two distinct types of Assignable (Special) cause variation, as discussed by Brian Joiner[11] - ***Erratic cause*** and ***Structural cause***. Examples of erratic special cause are an untrained operator, raw material variability, or an unknown process shift. Tool wear, and multiple cavity tools are examples of structural variation. GCpk is calculated from the fraction non-conforming (f.n.c.) of a given process. Traditional Cpk should only be estimated for stable, predictable processes; Generalized Cpk is promoted for use in cases where structural type of Special Causes is present, and Ppk should be calculated for processes affected by erratic Special Cause variation. Again, without stability there is no prediction. Software to estimate Generalized Cp/Cpk is not yet commercially available.

Additional Considerations

Structural Variation

Structure occurs when there are consistent, repeating patterns in the data. The patterns can occur over time (e.g. cycles, process deterioration), or within a subgroup (e.g. multiple cavity tooling, fixed crossweb differences, etc.). The effect of structural variation is exaggerated (wide) control limits. Structural variation is usually fairly easy to identify; points plotted on the X-bar or I chart cluster around the centerline. The ideal solution to structural variation is to eliminate the structure; however, this requires a process change, often difficult to achieve. If the structure occurs within the subgroups the 3-Chart method (Wheeler[20]) will limit the effects to the within-subgroup Range chart. The preceding paragraph introduced erratic and structural types of Special Cause variation. Stu Janis[10] elaborated on the topic of structural variation for Moving Web, Injection Molding, and batch processes. Stu explained that the standard deviation used to determine control limits should be based on the random portion of variability. It should not include biases such as fixed differences between cavities in mold tooling or crossweb differences in a web. The fixed bias only comes to play in determining the central line of a chart to control variability across cavities or across a web. Fixed differences often result in exaggerated control limits.

Multiple Sources of Random Variation

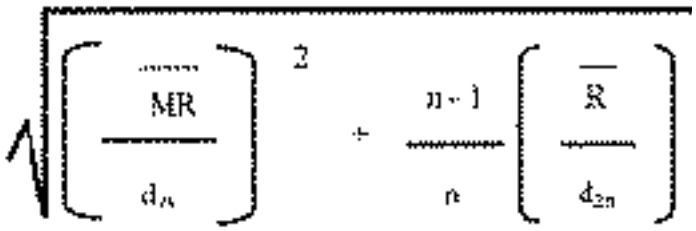
In the same paper Janis also explained the "Space vs. Time" concept of random variability. The sources of variability affecting within-subgroups (i.e. cavity-to-cavity or crossweb) are different than those affecting between subgroups (shot-to-shot or jumbo). Attempts to use within-subgroup variation (space) to set control limits for between subgroup averages (time) often result in limits that are too narrow. "3-Way Control Charts" (Don Wheeler[20]) monitor within-subgroup variability (space) using a Range chart, use a Moving Range chart to monitor short-term between-subgroup variability (time), and an Individuals charts to monitor differences between subgroup Means (time - longterm).

These same considerations may apply when estimating the standard deviation of individuals for Cp, Cpk indices of 3-Chart processes. If no bias exists, then the σ for

Continued on page 19

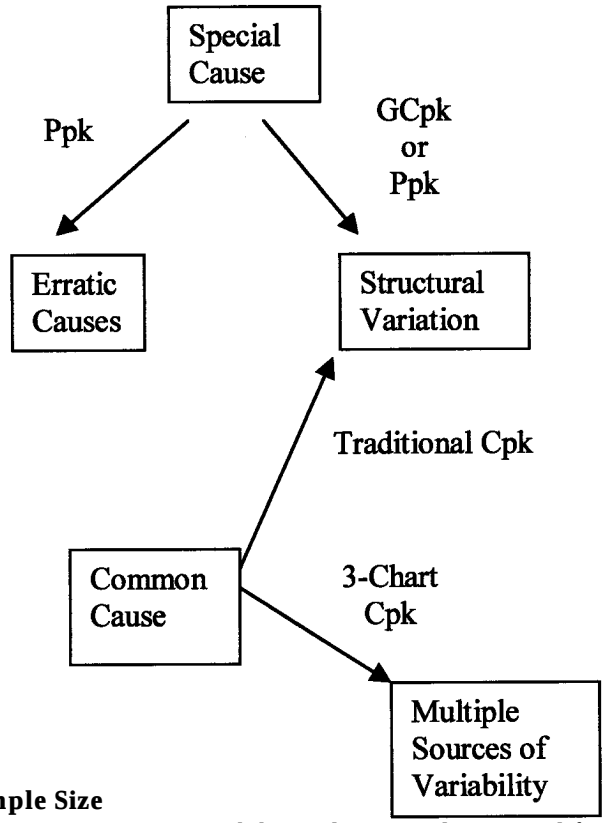
Continued from page 18

Individuals can be estimated using:
where d_{2k} = the value for d_2 using k number of sub groups



in the Moving Range, and
d_{2n} = the value for d_2 using n number of individuals within a subgroup

Roadmap



Sample Size

Because process capability indices are determined from estimates of standard deviation, they are affected by sample size (degrees of freedom). As expected, the stability of estimates of the standard deviation increases with sample size. We can show, using Chi-square tables and bootstrap techniques[7] , that a sample size (n) of 10 does not provide a very stable estimate of process capability. Even when n is 40 there is still substantial uncertainty in the estimator of Cpk. Tables 3 and 4 provide estimates of 95% Confidence Bounds for Cpk (lower bound) and Ppk (two-sided interval), assuming normality:

Table 3[7] Approximate 95% lower bound for Cpk.

Cpk	n=30	n=50	n=75
1.00	0.72	0.79	0.83
1.10	0.80	0.87	0.91
1.50	1.12	1.21	1.26
1.667	1.25	1.35	1.40

Table 4[15] 95% Confidence Interval for Ppk

Ppk	n=30	n=60	n=120
1.00	.76-1.31	.83-1.21	.88-1.14
1.33	1.02-1.76	1.11-1.61	1.17-1.52
1.67	1.29-2.19	1.49-2.01	1.47-1.90

Formulas (and further discussion) to compute the 95% Confidence Intervals for process capability indices are shown in Montgomery's[16] Introduction to Statistical Quality Control.

Practitioners often forget that process capability indices are merely **point estimates**. To avoid the pitfalls of making decisions using point estimates the concepts of Statistical Thinking should be employed (variability exists), and Cpk values plotted on control charts. A plot of Cpk values assists in the detection of process deterioration (or improvement).

Capability Studies

Process capability refers to the uniformity of the process. Montgomery[16] defines process capability analysis as an engineering study to estimate process capability. The AT&T Statistical Quality Control Handbook[3] defines the process capability study as a "Scientific systematic procedure for determining the capability of a process"… and defines capability as "the predictable series of effects produced by a process when allowed to operate without interference from outside causes…". The estimate of process capability may be in the form of a probability distribution having a specified shape, center, and spread. For this definition a process capability analysis may be performed without specs. i.e. Process Capability = 6σ.

Or, process capability may be expressed as a percent of product outside spec limits. This type of capability study ususally measures product functional performance, not the process itself. When the engineer can directly observe the process and can control the data collection methods this study is a "true process capability study" (Montgomery). When historical data is used and direct observation of the process is not possible, Montgomery refers to this as a product characterization study. "In a product characterization study we can only estimate the distribution of the product quality characteristics; we can say nothing about the statistical stability of the process."

Continued on page 20

MINI-PAPER

Continued from page 19

There are three primary techniques used to estimate process capability: histograms and probability plots, control charts, and hierarchical (nested) designed experiments.

Histograms (or stem-and-leaf plots) require at least 100 observations. If the data sequence is preserved, Mean Square of Successive Differences (MSSD) can be used to estimate the Short Term Standard Deviation (STSD). Or, an estimate of process standard deviation can be obtained from σ-hat=Rbar/d_2.

The probability plot has an advantage over histograms since it produces reasonable results for small sample sizes (Montgomery). However, other statistical methods are often needed to supplement the probability plots.

The control chart method is a simple, effective tool for process capability analysis. The control chart is the preferred technique for process capability analysis because it displays the potential capability of a process: patterns, trends, and other Special Cause signals.

An important consideration with use of the control chart method is selection of the proper rational subgroup to estimate the common cause process variability, against which special cause variability estimates are compared.

The hierarchical experimental design is a systematic approach to document and quantify the sources of variability in a process, and aids in identifying variation reduction opportunities. The fully balance, nested model is generally preferred due to its ease of statistical analysis, though unbalanced, staggered, and mixed models can be more accurate.

Final Thoughts

I have just spent several pages introducing various process capability indices. Nonetheless, I prefer to monitor progress of continuous improvement efforts with control charts rather than columns of process capability indices.

Bert Gunter[8] eloquently listed problems associated with focus on Cpk values:

1. Cpk cannot be used with one-sided specs or when the process is not normal.
2. Because the sampling distribution of the Cpk statistic is so variable it should not be used unless relatively large sample sizes are obtained (100-200).
3. Cpk goals can be impossible to meet when measurement error is large. Recall that
$$\sigma^2_{\text{Total}} = \sigma^2_{\text{Product}} + \sigma^2_{\text{Test error.}}$$
Reducing test error improves the Cpk value but does not really improve the product. Conversely, any process variability reduction without test method improvements may not result in much larger Cpk values.
4. Widening the product specs will result in a better (bigger) Cpk value but do nothing to improve the product or satisfy the customer.
5. Most importantly, Cpk is a meaningless measure of process capability unless your process is in a state of statistical control. Without statistical control a process is not predictable.

Unfortunately, a simple to understand, easy to calculate alternative to Cpk does not exist. Paraphrasing Bert Gunter, we must exercise caution to prevent process capability index scorekeeping from being confused with real improvement. Δ

Bibliography

1. ASQ Chemical & Process Industries Division, Specifications for the Chemical and process Industries: A manual for development and use, (ASQ Quality Press), 1996.
2. ASQ Chemical & Process Industries Division, Quality Assurance for the Chemical and Process Industries, (ASQ Quality Press), 1997
3. AT&T Statistical Quality Control Handbook, (Delmar Printing Co.), 1956.
4. Bailey, S. P., "How to Deal with the process Capability Catch-22", ASQ AQC Transactions, 1992.
5. Box, G. P., Hunter J. S., Hunter W. G., Statistics for Experimenters, (Wiley) 1978.
6. Chan, L. K., Cheng, S.W. and Spiring, F.A. "A New Measure of Process Capability – Cpm", JQT Vol. 20 No. 3., 1988.
7. Franklin, L. A. and Wasserman, G. S., "Bootstrap Lower Confidence Limits for Capability Indices", JQT Vol. 24 No. 4, October 1992.
8. Gunter, Bert, "The Use and Abuse of Cpk", Quality Progress magazine, Jan., Mar., May, July 1989.
9. Herman, J. T. "Capability Index – Enough for Process Industries?" 43rd ASQ AQC Transactions, 1988.
10. Janis, Stuart J., "Is Your Process Too Good for Its Control Limits?", 43rd ASQ AQC, 1988.
11. Joiner, B. L., Fourth Generation Management, (McGraw-Hill) 1994.
12. Juran, J. M., Quality Control Handbook 4th Edition, (McGraw-Hill) 1988.
13. Juran, J. M. and Gryna, F. M., Quality Planning and Analysis 3rd Edition (McGraw-Hill) 1998

Continued on page 21

NEW FTC AGREEMENT

The three sponsors of the Fall Technical Conference (Stat and CPI Divisions of ASQ and SPES Section of ASA) recently signed a new sponsorship agreement covering the next five FTC's. There are a number of changes to the agreement, changes intended to improve the FTC, and thus provide a better experience for members. Changes have been made in site selection procedures, vendor displays, Youden speaker selection, and short course availability.

Some of the changes:

Site selection; CPID will select the site, but will now do so following input from SPES and STAT.

Vendors: A Vendor Chair will be appointed, and vendors will be allowed to exhibit at future FTC's.

The Technical Program Committee will continue to have complete authority to fill the entire technical program, with three tracks: 1) Statistics, 2) Quality Control, and 3) Tutorials and Case Studies.

A committee of past-chairs of the three sponsors will select the Youden Speaker. The STAT past-chair will chair this committee, and the STAT Division will be responsible for administrative detail for this speaker.

The Conference Registration fee will be kept as low as possible. This fee will be waived for one speaker for each technical talk, and for the two luncheon speakers, the Youden Speaker, and the Hunter Award recipient.

Short Courses; There will be as many as four short courses, and these will be scheduled either on Tuesday/Wednesday preceding the conference or on Saturday following the conference. STAT will "own" two courses, and CPID and SPES one each. STAT will be fully responsible for planning and executing the entire short course program. Short course fees will be kept to a minimum, and be consistent across sponsors. Δ

MINI-PAPER

Continued from page 20

14. Kane, V. E., "Process Capability Indices", JQT Vol. 18 No. 1, 1986.
15. Kotz, Samuel and Johnson, Norman, Process Capability Indices, (Chapman & Hall) 1993.
16. Montgomery, Douglas C., Introduction to Statistical Quality Control 3rd Edition, (Wiley) 1997.
17. Pignatiello, J. J., and Ramberg J. S., "Process Capability Indices: Just Say No!", ASQ AQC Transactions, 1993.
18. Rodriguez, R. N., "Recent Developments in Process Capability Analysis", JQT Vol. 24 No. 4, October 1992.
19. Voelkel, J. O., "Process Indices for Certain Non-Stable Processes", ASQ 43rd Fall Technical Conference, Corning, NY, 1998.
20. Wheeler, D. J. and Chambers, D. S., Understanding Statistical Process Control 2nd edition, (SPC Press) 1992.

Bob Mitchell is a Sr. Quality Specialist with 3M Personal Care & Related Products Division. Bob is a senior member of the American Society for Quality, a Certified Quality Engineer, a Certified Quality Manager, and Chair-elect of the ASQ Statistics Division.

WEB SITES

The Statistics Division is responsible for developing and maintaining three web sites. The URL for each of these sites, and a brief discussion of the purpose/content of each site follows.

1. Statistics Division web page (www.asq.org/statdiv) is the official division electronic home. The viewer will find a list of officers, committee chairs and other volunteers; information on upcoming conferences (such as AQC, FTC and Applied Statistics); a list of division products and how to obtain them; the Virtual Academy, a page of links to statistics on-line tutorials; minutes of Stat Division meetings, etc.

2. The Statistical Clearinghouse. This page will primarily provide links to various Statistical resources on the web. For example, there will be links to major software vendors, major publishers of statistics materials, major statistics journals, and other statistics societies. There will also be reviews of statistics texts, software, etc. We expect this to be one of the first places someone would look in trying to answer the question – I wonder if… statistics…. – Temporarily, the URL for this page is internet.roadrunner. com/~webstar/

3. Previous copies of the Statistics Division Newsletter, beginning with the 1980 issues, are being posted at www.cba.bgsu.edu/asor/asqnews/letter.html. We hope to eventually have all previous newsletters posted, with separate pages for past Youden Addresses, past minipapers and tutorials, past lists of officers and committee people, etc. Δ