Statistical Process Control of Project Performance

Walt Lipke
Oklahoma City Air Logistics Center

With today's increased emphasis on statistical process control (SPC) as a management technique for software development, software organizations are attempting to employ the method for quality and project control. The focus of these efforts has primarily been with organizations having a Software Engineering Institute Capability Maturity Model® (CMM®) Level 4 or 5 rating. A few CMM Level 4 and 5 organizations have experimented with applying SPC control charts to another management technique – Earned Value Management (EVM). This article discusses the application of SPC control charts to the EVM indicators, schedule, and cost performance indexes.

The software division at the Oklahoma City Air Logistics Center was assessed as Software Engineering Institute (SEI) Capability Maturity Model® (CMM®) Level 4 in 1996, and became registered under the ISO 9001 standard for Quality Systems in 1998. The ISO registration was under the software implementation of the ISO standard known as "TickIT." For these accomplishments and several others, the software division was the recipient of the Institute of Electrical and Electronics Engineers' Software Process Achievement Award for 1999, a truly significant award for the division's efforts.

A large portion of the division's success has been due to embracing the Earned Value Management (EVM) methodology. EVM provided the needed structure to achieve many of the CMM Level 2 and 3 Key Process Areas (KPA) of the SEI's CMM. And, due to its numerical basis, EVM facilitated the achievement of the CMM Level 4 KPA, Quantitative Process Management (QPM), at that time.

However, today the updated QPM KPA strongly urges using control charts for statistical process control (SPC) with the new goal: "Statistically Manage the Sub-Processes [1]." CMM evaluators are now looking for SPC control charts as evidence of satisfying this KPA. Along with the rest of the software industry, we have struggled to meaningfully apply SPC control charts.

Although there is growing evidence of organizations following the CMM goal by implementing SPC with the defect data obtained from peer reviews, only a handful of organizations are employing the technique for controlling and improving software development process performance. The performance application is more difficult, but we believe it has more far-reaching results [2, 3].

Furthermore, we believe the application to performance management is more

SM CMMI and CMM Integration are service marks of Carnegie Mellon University.

in line with the intent of SPC, i.e., SPC is intended to optimize the performance of a system, not a component subsystem. The quality guru of the 1980s, Dr. Edward Deming, warned against applying SPC to sub-processes by themselves; he believed these actions could lead to optimizing the sub-process, possibly at the expense of the system. Thus, the following discussion concerns the application of SPC to managing the performance of software projects.

> **"Statistical process control charts, if successfully applied, can be a significant impetus for software process improvements."**

Statistical Process Control

There are several methods for performing SPC: scatter diagrams, run charts, cause and effect diagrams, histograms, bar charts, Pareto charts, and control charts [4, 5]. Although all of these methods are useful, we will focus this article on control charts.

SPC control charts, if successfully applied, can be a significant impetus for software process improvement. The method provides distinction between normal and anomalous process data; it is, in effect, a filter [6]. By knowing our normal process, we can reengineer it to obtain improvement in some performance aspect. And, by identifying anomalous behavior, we can seek the special cause (an influence from outside the system) and take action to prevent it from affecting future performance.

The fundamental idea of process

improvement is that as the system is observed over time, the process decreases its variation and, increasingly, gets closer to achieving its planned performance objective because of the introduction of improvements. SPC control charts facilitate this process improvement concept. Thus, you have the reason why the recently issued Software CMM Integration^SM (CMMI^SM) [1] has specifically used the words "statistically manage" in its CMM Level 4 Process Area, "Quantitative Project Management."

There are seven SPC control chart types, each having a specific application [4, 5]. The control chart required for our application is termed "Individuals and Moving Range." Symbolically, it is shown as XmR, where X represents the individual observations, and mR represents the moving range, the difference between successive observations. The XmR control chart is used when there is only one measurement of the variable in an observation period.

For all types of control charts, the control limits establish the filtering mentioned earlier. The high limit is plus three sigma from the average of the observations, whereas the low limit is the average minus three sigma. Sigma is a standard statistical measure of the variation in the process. An estimate of sigma is determined from the moving range. Measured values outside of the control limits have an extremely low probability of occurrence, only 0.27 percent if the process follows a normal distribution. Thus, any measured value beyond the control limits is deemed an anomaly, or in SPC terminology, a "signal," and should be investigated by management.

SPC is a much more involved subject than has been discussed here. Significantly more complete information is available in the references [4, 5, 6].
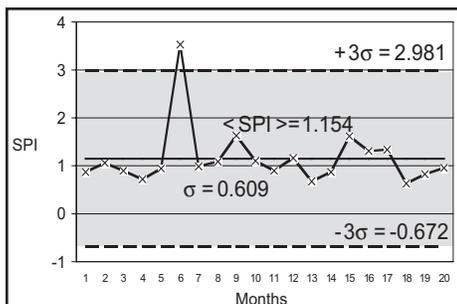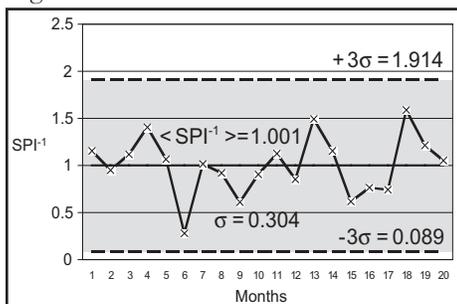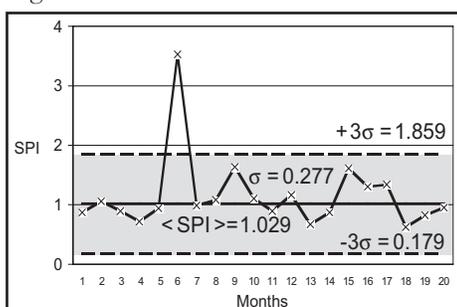
Figure 1: *SPI Control Chart*

Figure 2: *SPI⁻¹ Control Chart*

Figure 3: *SPI Control Chart (n=6 Removed)*

Earned Value Management

An excellent reference for EVM is a book by Quentin Fleming, *Cost/Schedule Control Systems Criteria, The Management Guide to C/SCSC* [7]. Just as with SPC, EVM is much more involved than the discussion in this paper. Here, we will only introduce the EVM indicators "cost performance index" (CPI) and "schedule performance index" (SPI).

EVM is based upon establishing a project baseline to achieve the "budget at completion" (BAC); BAC identifies the cost and completion points for the project manager. The baseline performance is a S-curve termed Budgeted Cost for Work Scheduled (BCWS); it is a graph of expected cost versus time. The in-process performance tracking is facilitated by two other curves, Actual Cost for Work Performed (ACWP) and Budgeted Cost for Work Performed (BCWP). BCWP is the earned value; it represents the completion of project tasks and is traceable to the values of cost and time duration allocated to those tasks during the project planning.

During project execution, the CPI and SPI indexes provide information about performance efficiency. The indexes are ratios. SPI is the efficiency of achieving earned value with respect to the performance baseline (SPI=BCWP/BCWS). Similarly, CPI is the efficiency of achieving earned value with respect to the actual costs (CPI=BCWP/ACWP).

Application/Data Analysis

Approximately three years ago, we began applying SPC to the EVM indicators SPI and CPI. We believed the merging of the two powerful management techniques held a considerable amount of promise. Our concept was that the application of SPC control charts to the monthly SPI and CPI values could be used in the following ways:

1. As a predictor of performance for the remainder of a project in work.
2. To improve the planning of new projects by using historical data from completed projects.
3. To effect process improvement, i.e., improve both execution and planning by using the measures of variation (sigma) in monthly performance and variance from the project plan (planned cost and completion date).

As stated, we have been using the method for some time. We have shared the ideas and results in two previous articles [2, 3]. Our results thus far indicate the method will fulfill its promise. However, its employment does require some additional understanding.

When we began preparing the control charts, we observed that the representation of the data affected the analysis and calculated results. To illustrate, we will use a small sample of actual data represented as both SPI and inverse SPI. Control charts for each data representation are shown in Figures 1 and 2. For the SPI chart, a signal is indicated at data point six. By removing the statistically anomalous data point six, the true process performance can be obtained. The control chart for SPI with data point six omitted from the calculations is shown in Figure 3. The true process has an average value of SPI (symbolically, <SPI>) and an estimate of sigma (σ) equal to 1.029 and 0.277, respectively. The inverse SPI chart (Figure 2), however, indicates there are no signals. Therefore, the true process for this data representation has an average value of SPI⁻¹ equal to 1.001, while sigma is estimated to be 0.304. As you can clearly see, the analysis results for SPI and SPI⁻¹ are not equivalent.

Problem/Proposed Solution

Of course, we should not expect the average values to be equal for the SPI and SPI⁻¹ analysis. However, if the signals found and the estimates of sigma are not identical for the two data representations, then we must ask the question, "Which result is correct, or is neither?" If we do not have a basis for choosing a way to represent the data and perform the analysis, then none of the three desired outcomes expressed in the Application/Data Analysis section are achievable.

Another problem can be seen from the histograms of CPI and CPI⁻¹ shown in Figure 4. The histograms were created from nearly six years of monthly data from one of our software development projects. By visual inspection, these histograms indicate that the data distributions are probably not "normal." Thus, predictions made by applying a normal distribution to the population would likely be inaccurate [5]. Therefore, similarly to the discussion in the preceding paragraph, unless there is a way to correct the behavior of the data, we cannot use the SPC information derived from the CPI and SPI data for the performance prediction, project planning, and process improvement applications cited earlier.

There are several recognized correction methods that can be used when the distribution of the data is not normal [5]. However, the most appealing is to transform the data in a mathematical way to approximate a normal curve. This is the solution approach discussed in the remainder of the article.

As we became more curious about the differences in the results from the control charts of SPI versus SPI⁻¹ and CPI versus CPI⁻¹, we noticed a general bias. The average of the monthly values for either representation is generally larger than its corresponding cumulative value (e.g., <SPI> > SPI_{cum}) and the signals found using XmR control charts are predominantly the observations having values greater than 1.0. Our analysis indicates the problem occurs because the performance indicator (PI) values below 1.0 cannot be less than zero. It is impossible to have a negative value for the PI because it is, simply, a ratio of two positive numbers. However, the values of the PI above 1.0 are unlimited[1].

This behavior of the PI was deduced to be incongruent with the three sigma process limits computed for the individual control chart. The process limits themselves are unbounded; conceivably, they can have values ranging from plus to minus infinity. The process limits are equally spaced above and below the PI average value. However, equivalent good and poor observed values for the PI are not spaced equally above and below the nominal value of 1.0. The PI values less

than 1.0 are virtually ignored; observed performance values below the lower process limit are rare occurrences. For example, signals identified as values higher than the upper process limit for non-inverted data are not detected when the data is inverted. Both Figures 1 and 2 illustrate this inconsistency. Data point six is identified as a signal by the SPI control chart (Figure 1). However, when the data are inverted, the control chart for SPI^{-1} (Figure 2) indicates data point six as nothing unusual; it is within the lower process limit, and thus, is considered to be part of the process. Another significant observation from Figure 1 is that the lower process limit value is below zero; therefore, any SPI value less than 1.0 cannot be detected as a signal.

To make this point a little clearer, let us use SPI (BCWP divided by BCWS). Suppose BCWP is five units and BCWS is one unit. Now suppose the performance is reversed; BCWP is one unit, and BCWS is five units. For the first instance, $SPI_a = 5.0$, and for the second, $SPI_b = 0.2$. The two instances are, numerically, the reciprocal of the other and represent equivalent anomalous performance. $SPI_a = 5.0$ is excessively good schedule efficiency, whereas $SPI_b = 0.2$ is excessively poor schedule efficiency. The two values are not equidistant from 1.0, the nominal value; SPI_a is four units away from 1.0, whereas SPI_b has a separation of only 0.8 units.

If we constructed a control chart with these data points included, SPI_a would likely be detected as a signal. And if the lower process limit happened to be a negative number, SPI_b would be seen as part of the process. However, for the inverse data representation the signal detected becomes $SPI_b^{-1} = 5.0$, while $SPI_a^{-1} = 0.2$ is ignored. The signals identified are switched with the change in data representation. Thus, it should now be understood that the SPC control limits generally detect EVM performance index indicator signals greater than 1.0 and ignore poor performance values less than 1.0, regardless of whether the data are represented as inverted or non-inverted.

Due to the incongruence between the PI data and the XmR process limits, a method was sought to transform those PI (or PI^{-1}) values less than 1.0 from being bounded by zero to unbounded. It was hoped that if the data values less than 1.0 could be made to resemble the characteristics of those values greater than one, then consistent results could be obtained from data represented as either inverted or non-inverted.

The method for representing all of the data as unbounded is extremely simple. The PI data are transformed for the SPC analysis by using the natural logarithm function. Applying the natural logarithm function causes PI values less than 1.0 to be represented by negative numbers. The transformed values are a data set that is congruent with its corresponding three sigma process limits. Thus, when the transformed data are used for creating the SPC control chart, a negative value lower process limit will not necessarily mean that a signal whose value is less than 1.0 will be ignored. The transformed data provides the possibility of identifying signals for both high and low PI values.

Solution Criteria/Testing/Results

With the mathematical method defined for transforming the data, tests can be performed to determine whether or not the application of the logarithm function meets a set of desirable behavior characteristics. Fundamentally, if the solution is the correct one, it should not matter which data representation is used for the SPC analysis, inverted or non-inverted.

Specifically, the un-transformed average value of the inverted data should be the reciprocal of the un-transformed average value of the non-inverted data[2]. The value of sigma representing the process variation for the non-inverted data should equal the value of sigma determined from the inverted data. The same data points should be identified as signals in either data representation. Lastly, the transformed data should show improved agreement to the normal distribution. If the solution meets all of the criteria, we can feel confident in its use.

To test the transformation, it is applied to the SPI data previously analyzed (see Figures 1 and 2). The transformed SPI data, both the non-inverted and inverted representations, are shown in the SPC control charts, Figures 5 and 6, respectively. Reviewing and comparing these control charts, it can be said that the results satisfy three of the four solution criteria:

1. The un-transformed average values, i.e., $<SPI>_u$ and $<SPI^{-1}>_u$, are 0.994 and 1.006, respectively; $<SPI^{-1}>_u$ is the reciprocal of $<SPI>_u$.
2. The signal identified in Figure 5 is identical to the signal found in Figure 6, i.e., data point six. Figure 6 is the mirror image of Figure 5 with respect to the ordinate value 0.0.
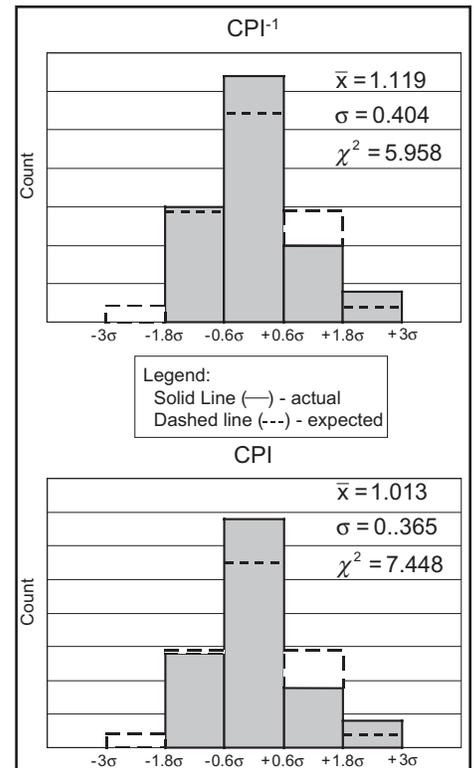

Figure 4: *Histograms – CPI and CPI^{-1}*

3. The estimate of sigma is the same value for both the inverted and non-inverted representations of PI; i.e., the value of σ is 0.259.

One test using a small data sample certainly is not proof that the transformation meets the criteria in every instance. However, we have run the analysis using the natural logarithm transformation on data from several projects (more than 400 data points) and have seen consistent results. Furthermore, although it is beyond the scope of this paper, through mathe-
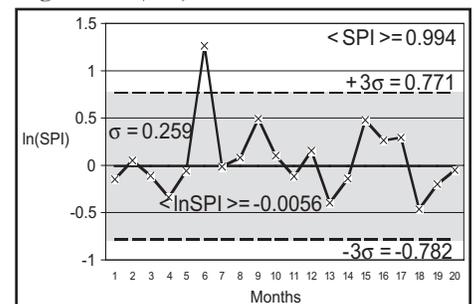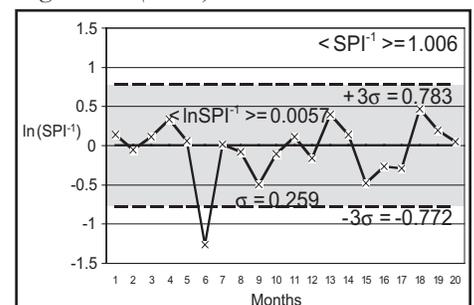
Figure 5: *ln(SPI) Control Chart*

Figure 6: *ln(SPI-1) Control Chart*

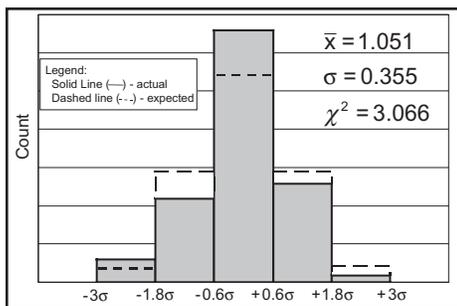$$\bar{x} = 1.051$$
$$\sigma = 0.355$$
$$\chi^2 = 3.066$$

Legend:
Solid Line (—) - actual
Dashed line (---) - expected

Count

-3σ -1.8σ -0.6σ +0.6σ +1.8σ +3σ

Figure 7: *Histogram − lnCPI[1]*

matical analysis it can be shown that each of these three criteria for data representation equivalence is achieved by applying the natural logarithm function.

To illustrate that the transformation improves the normality, visually compare the histogram shown in Figure 7 to its counterpart in Figure 4. The same data are used for both histograms; however, Figure 7 uses the transformed values. The histogram of the logarithm values has a more normal appearance. To further explain the improvement, the "goodness of fit" Chi-Square (χ^2) statistical test [5, 8] is applied to the histograms[3]. Smaller values of Chi-Square indicate more conformity to a normal distribution. The Chi-Square value for the raw CPI[-1] data is 5.958 (see Figure 4), whereas the value for the transformed data is 3.066 (see Figure 7). Thus, there is 48.5 percent improvement in Chi-Square by the use of the natural logarithm function.

In addition to high and low signals being detected equally, another significant outcome from using the natural logarithms of the data is that the resulting average value of the performance indicator ($<PI>_u$) is unbiased. It is not generally greater than the cumulative value of PI (PI_{cum}). Thus, the average performance indicator values, $<SPI>_u$ and $<CPI>_u$, and their reciprocals can be used with confidence in EVM calculations.

Summary

SPC applied to the EVM indicators, CPI and SPI indexes, can be a very powerful management tool. It has immense potential to improve both the planning and execution of software development projects and to provide a measure of that improvement. We have shown, however, that the application has inconsistencies that detract from its employment. A solution is proposed for resolving the problem. A test for the solution is described and performed using a small set of SPI data. The results of more extensive testing and further mathematical analysis indicate the recommended solution has merit. Employing the data transform technique significantly reduces the SPC-EVM application inconsistencies, thereby yielding much improved results.◆

References

1. CMMI-SE/SW, Version 1.02, CMU/SEI-TR-018, CMMI Product Development Team, Nov. 2000.
2. Lipke, Walt, and Jeff Vaughn, eds. "Statistical Process Control Meets Earned Value." CROSSTALK June 2000: 16-20.
3. Lipke, Walt, and Mike Jennings, eds. "Software Project Planning, Statistics, and Earned Value." CROSSTALK Dec. 2000: 10-14.
4. Florac, William A., and Anita D. Carleton, eds. Measuring the Software Process. Reading, Mass.: Addison-Wesley, 1999.
5. Pitt, Hy. SPC for the Rest of Us. Reading, Mass.: Addison-Wesley, 1995.
6. Wheeler, Donald J. Understanding Variation, The Key to Managing Chaos. Knoxville, Tenn.: SPC Press, 2000.
7. Fleming, Quentin. Cost/Schedule Control Systems Criteria, The Management Guide to C/SCSC. Chicago: Probus, 1988.
8. Crow, Edwin L., Francis A. Davis, and Margaret W. Maxfield, eds. Statistics Manual. New York: Dover Publications, 1960.

Notes

1. The observations discussed in this paragraph are made for the non-inverted data representation. The statements apply equally as well for the inverted data representation (PI^{-1}).
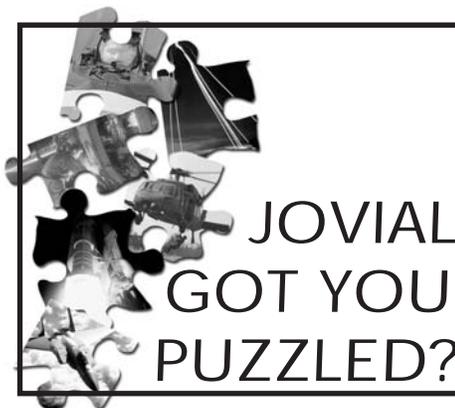
2. The reciprocal relationship of the untransformed average values is also needed for EVM calculations. The cumulative values of the EVM performance indicators, represented as inverted and non-inverted data, possess this characteristic, i.e., $(PI_{cum})^{-1} = (PI^{-1})_{cum}$. Thus, to be confident using the numbers from the SPC analysis in EVM calculations, they must behave in accordance with the cumulative values.

3. The definition of Chi-Square for this application is: $\chi^2 = \Sigma_i$ [(expected count$_i$ − observed count$_i)^2$ / expected count$_i$], where "i" designates one of the five histogram areas (e.g., 0.6σ to 1.8σ).

About the Author

Walt Lipke is the deputy chief of the software division at the Oklahoma City Air Logistics Center. He has 30 years of experience in the development, maintenance, and management of software for avionics automated testing. In 1993 with his guidance, the Test Program Set and Industrial Automation (TPS and IA) functions of the division became the first Air Force activity to achieve the Software Engineering Institute's Capability Maturity Model® (CMM®) Level 2. In 1996, these functions became the first software activity in federal service to achieve CMM Level 4 distinction. The TPS and IA functions, under his direction, became ISO 9001/TickIT registered in 1998. Lipke is a professional engineer with a master's degree in physics.

OC-ALC/LAS
Tinker AFB, OK 73145-9144
Phone: (405) 736-3341
Fax: (405) 736-3345
E-mail: walter.lipke@tinker.af.mil