



## White Paper

### *OptiRamp*<sup>®</sup> Lean Six Sigma

*Continuous Improvement through Statistical Process Control*

Vadim Shapiro  
Dmitriy Khots, Ph.D.

## Table of Contents

Introduction.....	2
Variable Horizon Assessment.....	2
Key Process Characteristics .....	3
Box plots .....	4
Process Capability Index.....	5
Enterprise Process Control.....	6
Testing for Differences between Two Scenarios .....	7
Testing for Differences between Proposed and Current Scenarios .....	8
About Statistics & Control, Inc.....	10

## Introduction

The *OptiRamp* Lean Six Sigma (LSS) Submodule provides real-time process requirements verification and formal real-time assessment to validate the financial benefit (cost savings, revenue, or capital cost avoidance) generated from improvements in process performance.

LSS continually assesses to evaluate performance results against established goals, redirecting dispatchers as necessary to meet desired objectives. Personnel can use LSS to assist them map and document standardized processes and analyze operation results for further continuous improvement. LSS also provides cumulative financial savings information for each operation, and reports can be automatically scheduled and delivered to dispatchers and executives.

The LSS Submodule performs the following functions:

- Analyzes the actual operating mode for production processes from both operations and financial points of view
- Performs operations and financial analyses for forecasted and suggested operating modes
- Prepares data for visual real-time presentations and reports
- Interfaces with other *OptiRamp* modules and submodules

## Variable Horizon Assessment

The *OptiRamp* LSS Submodule allows for both short- and long-term process performance assessments by tracking key performance metrics within parametrically defined time frames.

Short-term metrics are defined as C1 and C2. C1 metrics are calculated from the subset for the current operating mode (since the last configuration change). C2 metrics are calculated from the subset determined by configuration parameters (e.g., Ts, hours). Long-term metrics are defined as P1 and P2. P1 metrics are calculated from the entire dataset or from a subset determined by configuration parameters (e.g., T11, months). P2 metrics are calculated from the entire dataset or from a subset determined by configuration parameters (e.g., T12, months).

Figure 1 shows the data collection and grouping strategy based on the observation time frame. Parameters  $T_s$ ,  $T_{I1}$ , and  $T_{I2}$  can be specified according to business requirements. The formulas and basic principles used to assess the process during both the short- and long-term time frames are identical.

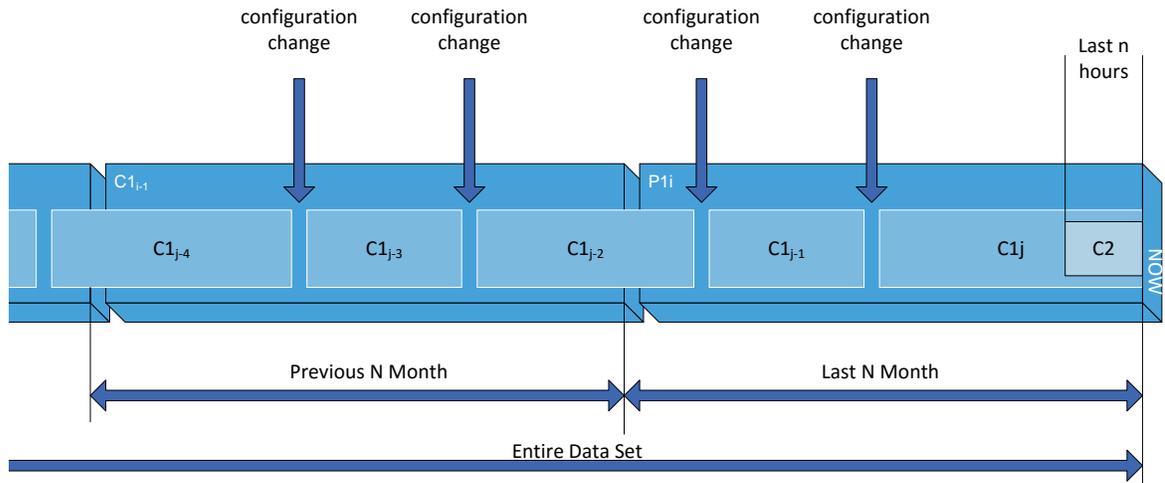


Figure 1. Data collection and grouping strategy

## Key Process Characteristics

The LSS algorithms require calculating the following descriptive statistics for the process metric being assessed: mean, median, interquartile range (IQR), range, and standard deviation. Given the assessed metric  $X$ , the descriptive statistics are calculated as shown in Table 1.

Table 1. Calculating descriptive statistics

Statistic Name	Formula	
	Short-Term (C)	Long-Term (P)
Mean	$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ , $n$ is the number of observations in the short-term dataset	$\mu = \frac{\sum_{i=1}^N X_i}{N}$ , $N$ is the number of observations in the long-term dataset
Median	The median is the midpoint of an ordered set of data. It is the middle value if there are an odd number of data points or the average of the two middle values if there is an even number of data points. The median is denoted $\tilde{X}$ for C metrics and $\tilde{\mu}$ for P metrics.	
Quartile	Quartiles split the dataset into four pieces, each containing 25 percent of data. The 1st Quartile is defined as the median of the segment between the first point and the set median. The 2nd Quartile is the set median. The 3rd Quartile is defined as the median of the segment between the set median and the set max.	

Statistic Name	Formula	
	Short-Term (C)	Long-Term (P)
IQR	The Interquartile Range (IQR) is the distance encompassed by the 1 <sup>st</sup> and 3 <sup>rd</sup> Quartiles.	
Range	$R = X_{\max} - X_{\min}$	
Standard Deviation	$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$	$\sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \mu)^2}{N-1}}$

## Box plots

The characteristics shown in Table 1 are used to create box plots. A box plot gives the user a quick visual assessment of the process metric distribution, e.g., a financial characteristic of a particular scenario. More specifically, the box plot shows the range of data values that comprise 50 percent of the dataset (the 1st and 3rd quartiles) and includes the following characteristics shown in Figure 2. The line that divides the box shows the median. The central black dot is the mean. The box height represents the IQR. Single-line “whiskers” extend from the box, showing the data values at or below the 1st Quartile minus 1.5 IQR and at or above the 3rd Quartile plus 1.5 IQR. The following terminology/abbreviations are used in Figure 2: *Outlier* is a data value that falls above or below the whiskers,  $HL = Q3 + 1.5IQR$ ,  $LL = Q1 - 1.5IQR$ ,  $Q3$  is the 3rd Quartile and  $Q1$  is the 1st Quartile. Such plots allow the users to quickly compare multiple scenarios.

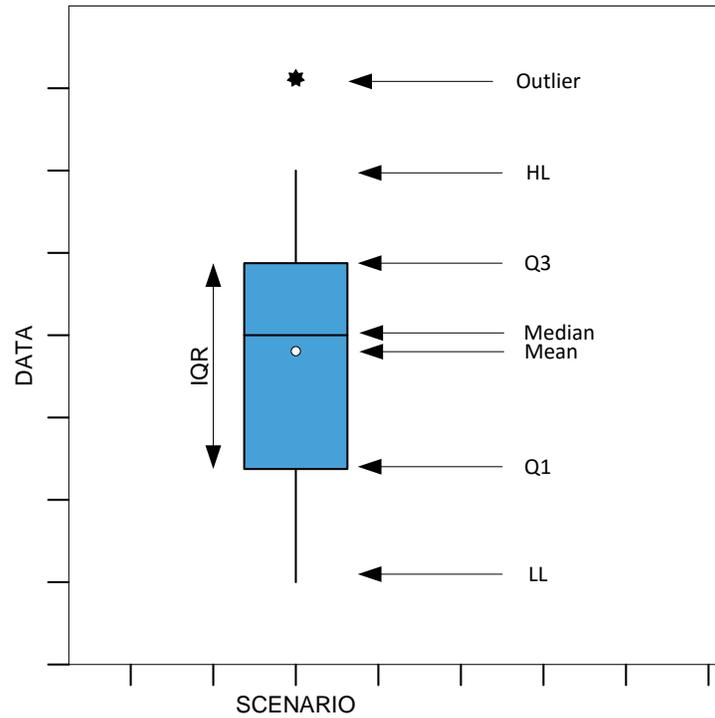


Figure 2. Box plot example

### Process Capability Index

The *OptiRamp* LSS Submodule calculates the process capability index to determine whether the process (process metric values being assessed) is capable of being in control in the short-term or has historically been in control. Table 2 summarizes these calculations. Figure 3 shows a graph of the process capability index.

Table 2. Process capability index calculations

Index	Formula	
	Short Term (C)	Long Term (P)
Lower capability index*	$C_{pl} = \frac{\bar{X} - LS}{3s}$	$P_{pl} = \frac{\mu - LS}{3\sigma}$
Upper capability index*	$C_{pu} = \frac{US - \bar{X}}{3s}$	$P_{pu} = \frac{US - \mu}{3\sigma}$
Capability Index	$C_{pk} = \min \{ C_{pl}, C_{pu} \}$	$P_{pk} = \min \{ P_{pl}, P_{pu} \}$

\* LS is the lower specified limit and US is the upper specified limit, as defined by process requirements.

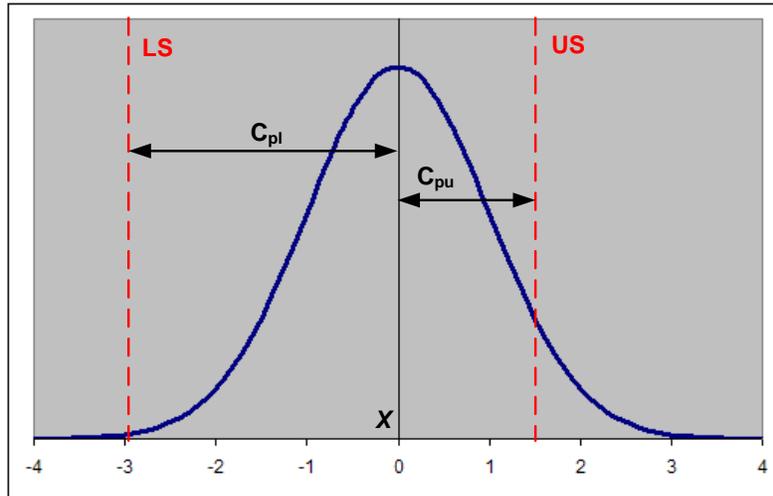


Figure 3. Process capability index

Generally,  $C_{pk}$  values greater than 1.33 indicate that the process is  $\pm 3$  sigma capable, while  $P_{pk}$  values greater than 1.33 indicate that the process has historically been in control.

## Enterprise Process Control

The *OptiRamp* LSS Submodule can be applied at the level of a single unit, a station, an entire network, or at the enterprise level (see Figure 4). The enterprise system has a hierarchical structure.

At the top of the hierarchy is the entire data population. An enterprise box plot provides visual presentation of overall financial characteristics. Further down the hierarchy are operating units (such as the steam distribution and generation network, electrical power generation and distribution network, fuel gas wells and distribution networks, etc.). Box plots visually present entire network financial characteristics. The station level is next. On this level, box plots present financial characteristics for stations and plants. The unit box plots (steam generators, heat recovery steam generators, power generators, gas compressors, etc.) are located at the very bottom of the system structure.

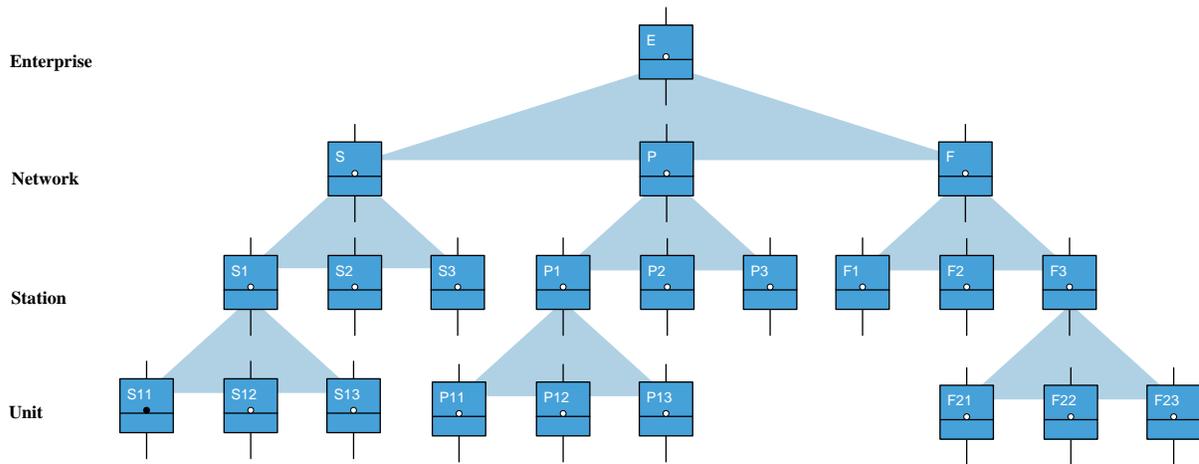


Figure 4. Enterprise system process control

Financial characteristics and box plot scenarios can be presented for each individual unit, station, network, and entire enterprise. The LSS Submodule can test for differences between any two scenarios for the same or different C/P time metrics.

### Testing for Differences between Two Scenarios

The LSS Submodule allows the user to evaluate differences between scenarios by visually examining side-by-side box plots of the scenarios to study their central tendency, variation, and shape. Figure 5 shows how box plots can be compared for differences.

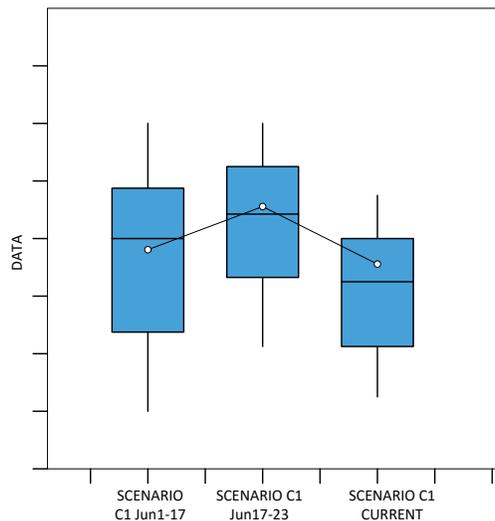


Figure 5. Scenario comparison with box plots

The LSS Submodule also performs statistical tests on the differences in means of two populations. LSS utilizes the t-test for this comparison; the t-test for short- and long-term process metric means is calculated similarly. Table 3 shows the calculations performed.

**Table 3. Statistical tests for population mean differences**

Statistic Name	Formula	
	Short Term (C)	Long Term (P)
t-statistic*	$t_C = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{s_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$	$t_P = \frac{(\mu_1 - \mu_2)}{\sqrt{\sigma_p^2 \left( \frac{1}{N_1} + \frac{1}{N_2} \right)}}$
Pooled variance†	$s_p^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}$	$\sigma_p^2 = \frac{(N_1 - 1) \cdot \sigma_1^2 + (N_2 - 1) \cdot \sigma_2^2}{N_1 + N_2 - 2}$

\* Short-term scenario comparison:  $s_1^2$  is the first scenario variance,  $s_2^2$  is the second scenario variance,  $\bar{X}_1$  is the first scenario mean,  $\bar{X}_2$  is the second scenario mean,  $n_1$  is the first scenario observation count, and  $n_2$  is the second scenario observation count.

† Long-term scenario comparison,  $\sigma_1^2$  is the first scenario variance,  $\sigma_2^2$  is the second scenario variance,  $\mu_1$  is the first scenario mean,  $\mu_2$  is the second scenario mean,  $N_1$  is the first scenario observation count, and  $N_2$  is the second scenario observation count.

The calculated t-statistic is used to test the null hypothesis that the means of two scenarios are the same. For the one-sided t-test, the corresponding p-values are calculated as  $P(t \geq t_C)$  for the short-term and  $P(t \geq t_P)$  for the long-term. For the two-sided t-test, the p-values are  $2P(t \geq t_C)$  and  $2P(t \geq t_P)$ . If the p-value is less than the parametrically defined  $\alpha > 0$ , then the null hypothesis is rejected and LSS alerts the user that the means of two scenarios have statistically significant differences. By default,  $\alpha = 0.05$ ; however, LSS allows the user to define a more or less aggressive confidence interval.

### Testing for Differences between Proposed and Current Scenarios

The *OptiRamp* Dynamic Optimization Submodule finds an array of possible operating scenarios based on the given optimization criteria (proposed scenarios). The submodule then uses a method that locally optimizes the objective function upon algorithm iteration. At the station/enterprise level, the load sharing problem is solved by either unit shutdown or start-up and by redistributing the load between running units. The criteria for optimal load sharing are based on comparing the objective function value at the current point of time with its value computed using the static model after the planned load changes.

The *OptiRamp* LSS Submodule evaluates the gains of proposed scenarios in comparison with the current operating mode. LSS uses a heuristic t-test to determine whether a single-point proposed scenario is expected to gain over the current operating mode. The t-statistic used in the analysis

is  $t = \frac{\bar{X} - \bar{X}_p}{s_1}$ , where  $\bar{X}_p$  is the proposed scenario expected mean. The assumptions of the test are that the proposed scenario mean is the long-term process mean,  $\bar{X}_p = \mu$ , and that the proposed scenario standard deviation is the current process standard deviation,  $s_p = s$ . Figure 6 illustrates a proposed scenario evaluation.

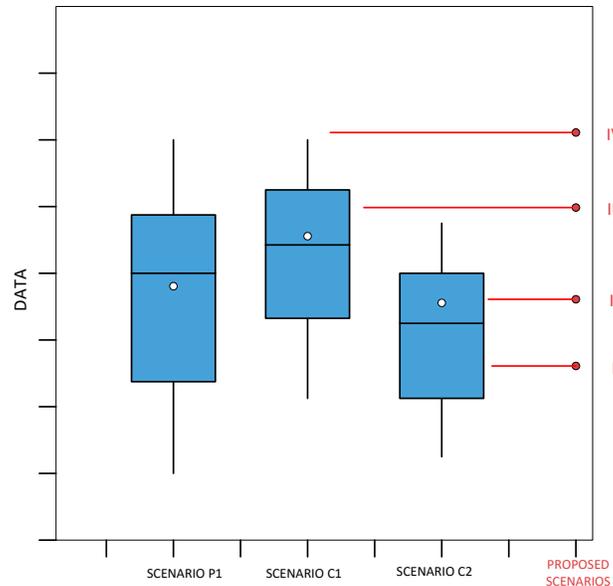


Figure 6. Proposed scenario evaluation

There are four proposed scenarios (I through IV). LSS will make recommendations based on the above assumptions for each of the proposed scenarios. For example, LSS compares proposed scenario I with scenario C2; the calculated t-statistic shows significant differences ( $\bar{X}_1 > \bar{X}_p$ ), thus scenario I is recommended for rejection. Proposed scenario II is also compared with scenario C2. The t-statistic shows significant differences ( $Q_3 > \bar{X}_p$ ), thus scenario II is also recommended for rejection. Proposed scenario III is evaluated against both scenarios C1 and C2. In the case of C1, there is a significant difference ( $Q_3 < \bar{X}_p$ ); in the case of C2, the calculated t-statistic also shows a significant difference ( $Q_3 > \bar{X}_p$ ). The LSS recommends scenario III for acceptance with low priority. Proposed scenario IV is also compared against both scenarios C1 and C2. The t-statistic in this case shows significant differences ( $Q_3 + 1.5IQR < \bar{X}_p$ ) for both existing scenarios; hence, LSS highly recommends scenario IV for acceptance.

The *OptiRamp* LSS Submodule feeds scenario acceptance recommendations to the *OptiRamp* visualization tools as well as the *OptiRamp* Real-Time Optimization Submodule so that proposed scenarios can be recalculated/adjusted based on the recommendations.



## About Statistics & Control, Inc.

S&C—an engineering consulting and technology company headquartered in West Des Moines, IA—solves complex challenges for customers through its unique technology and its highly seasoned team of professionals. The company has a global portfolio spanning the energy, oil and gas, utility, and digital oil field industry sectors. S&C provides clients with turbomachinery control solutions that easily integrate with the existing system as well as *OptiRamp*® solutions, which focus on process and power analytics to optimize processes and, in turn, reduce costs and increase reliability. S&C also provides consulting, dynamic system studies, modeling, automation, training and OTS, and support services.

**Statistics & Control, Inc.**  
4401 Westown Pkwy, Suite 124  
West Des Moines, IA 50266 USA  
Phone: 1.515.267.8700  
Fax: 1.515.267.8701