SOP No. 9

Recommended Standard Operating Procedure
for
Control Charts for Calibration of Mass Standards

1. Introduction

1.1. This SOP describes procedures for the development of control charts and their use for demonstrating attainment of statistical control of a mass calibration process.

1.2. Prerequisites

The use of this SOP requires that appropriate apparatus, methodology, and standards are available, and that the laboratory thoroughly understands the basic principles of the measurement process used and has had sufficient experience to perform the necessary operations required for the measurements of concern.

2. Summary

An appropriate check standard (or control standard) is incorporated into the measurement process and weighed at established intervals, the results are plotted on an $\bar{x}$ (X-bar) chart. The abscissa (*x*) represents the sequence of measurements and the ordinate (*y*) the measured values. A central line is drawn, indicating the mean ($\bar{x}$) of the measured values and control limits are indicated within which the results of measurements are expected to be randomly distributed, based on statistical considerations. The system is considered to be in statistical control when the individual values are within the designated limits. The system is considered to be out of control if an excessive number of values are present outside established limits, unusual trends are observed, or if the mean exceeds the control limits. The statistical information on which the control limits are based can be used to calculate confidence limits for measurements made while the system is demonstrated to be stable and in a state of statistical control.

3. Procedure

3.1. Definition of Monitored System

The monitored system is considered to consist of the balance, the standard operating procedure, the laboratory environment, the check standard or control standard, the operator, and any other sources that contribute to the variance or bias of the measurement data. Any of the above that can be considered to be constant or negligible contributors to the variance may be consolidated and monitored by a single control chart. Any that cannot be so considered (for example: different standard, different balance, different SOP) may require separate control charts.

The variability of balance precision that is load dependent must be considered. For many balances, precision is a function of load, and a distinct control chart is required (in principle) for every load tested. This is not always feasible, except in the case of SOP 5 or 28 where check standards are incorporated into the measurement process. Hence, control charts used for measurement assurance and evaluation of measurement uncertainty are generally satisfactory if developed using data from check standards at two or three intervals for each balance appropriately spaced within the range of balance use, or at least with one check standard for each decade. On balances where few nominal values (loads) are tested, a control chart should be established for each load.

3.2. Selection, Care, and Calibration of Check Standards

3.2.1. A check standard used in high precision calibration measurement must be stable and normally be comparable to the primary standard or to the typical calibration work, depending on what is being monitored. For lower order calibrations, the check standard should simulate the laboratory's primary standards to the extent feasible. It should be calibrated with an expanded uncertainty equal to or better than the precision of the process being monitored. All check standards should be cared for in the same way as primary standards to prevent their damage or deterioration. Lower order check standards should be recalibrated at regular intervals according to Good Measurement Practice (GMP) 11 on Setting and Adjusting Calibration Intervals.

Table 1. Recommended check standards for typical test situations

Balance	Range of Measurement	Check Standard(s)
Echelon III (Class F) SOP 7, 8	5000 lb to 0.001 lb	2 to 3 values per balance OR 1 chart per load
Echelon II SOP 3, 4, 6, 7	5000 lb to 0.001 lb and 1000 kg to 1 mg	2 to 3 values per balance OR 1 check standard per decade
SOP 5	typically 1 kg to 1 mg	each nominal value incorporates a check standard
SOP 28	typically 1 kg to 1 mg	1 check standard per decade (e.g., 1 kg, 100 g, 10 g, 1 g, 100 mg, 10 mg, 1 mg)

3.3. Establishing Control Chart Parameters

3.3.1. The control chart parameters consist of the central line, the best estimate of the mean of measurements of the check standard, and control (or "action") and warning limits that represent probabilistic limits for the distribution of results around the central line. These parameters are evaluated on the basis of a reasonable number of initial measurements and updated as additional measurement data are accumulated.

3.3.2. Make at least seven (minimum number) and preferably 12 or more, independent measurements of the check standard under the same conditions that will be used to make routine measurements. No two measurements should be made on the same day. This is necessary to estimate the long-term standard deviation to the extent feasible. To make statistically valid decisions or calculate uncertainties based on this data, 25 to 30 points are necessary.

Calculate the mean, $\bar{x}$ and the estimate of the standard deviation, s in the conventional manner.

Establish the control chart parameters as follows:

Central Line	$= \bar{x}$
Upper Control/Action Limit	$= \bar{x} + 3s$
Upper Warning Limit	$= \bar{x} + 2s$
Lower Warning Limit	$= \bar{x} - 2s$
Lower Control/Action Limit	$= \bar{x} - 3s$

Control chart parameters for Echelon III (Class F or other) may be completed as follows to track practical limits:

Central Line	$= \bar{x}$
Upper Control/Action Limit	$= \bar{x}$ + *1/4 tolerance*
Upper Warning Limit	$= \bar{x}$ + *1/10 tolerance*
Lower Warning Limit	$= \bar{x}$ - *1/10 tolerance*
Lower Control/Action Limit	$= \bar{x}$ - *1/4 tolerance*

3.3.3. Upgrading Control Chart Parameters

Upgrade control chart parameters when a significant amount of additional data is available or when the previously determined parameters are no longer pertinent due to changes in the system.

Note: Ordinarily, upgrading is merited when the amount of new data is equal to that already used to establish the parameters in use, or when at least seven additional data points have been recorded.

Calculate $\bar{x}$ and s for the new set of data and examine for significant differences from the former using the t-test and F-test, respectively. If the tests fail and results are significantly different, determine the reason for the difference, if possible, and decide whether corrective action is required. If data does not agree within statistical limits, establish new parameters using the most recent data and note the reasons for not using previous data. If no significant differences between the data sets are found, pool all data and calculate new control chart parameters based on all existing data.

3.4. Frequency of Measurement

The check standard should be measured and plotted with sufficient frequency to minimize the risk of loss of data during the period from last-known-in to first-known-out of control condition. It is good practice to measure the check standard at least once during each period when a set of test measurements is made. For critical calibrations or those of highest accuracy, it is desirable to alternate measurements of test items and check standards, but for real-time evaluation it is preferable to incorporate the check standard in the calibration design as in SOP 5 or SOP 28.

Whenever there has been a long period of inactivity, it is good practice to make a series of measurements of the check standard and to plot the results on a control chart to demonstrate attainment of statistical control prior to resuming measurements with that specific calibration system.

Control charts should be updated as close to real time as feasible to effectively monitor the measurement process to prevent the possible release of questionable data resulting in recall.

4. Use of Control Charts

4.1. Monitoring a Measurement Process

Use the following criteria to interpret control chart results.

4.1.1. If plotted points are stable and randomly distributed within the warning limits, decide that the system is in control.

4.1.2. If a plotted point is outside the warning limits but within the control limits, investigate the presence of calculation errors. If none were made, re-measure the check standard. The re-measured value must be within the warning limits to merit the decision of "in control". If the results are not within limits, consider the measurement process "out of control". Reject all data obtained since last "in-control" measurement and take corrective action (hence “action” limit). Accept no further data until the system is demonstrated to be in-control as indicated by at least two successive measurements of the check standard within the warning limits.

If a plotted point is outside the control limits and arithmetically correct, the system is considered to be out of control. Data are rejected, corrective actions must be taken and re-attainment of statistical control demonstrated, as above, before data may be accepted.

4.2. Transfer of Measurement Statistics

4.2.1. Absence of a significant difference between the central line and the accepted value for the check standard may be considered as evidence of insignificant bias at the level of confidence of the statistical test used. This conclusion is valid, as long as the system remains in control. On occasion, small differences (less than 1 s) from unknown sources will be obvious over time and the value observed for the bias should be incorporated into the uncertainty per SOP 29.

4.2.2. The estimate of the standard deviation of the process, s_p, used to establish the control limits may be used to calculate confidence intervals for all pertinent measurements made while the system is in control. However, see SOP 29 for calculation of measurement uncertainty using the process variability, s_p. The value of the test weight is said to be within the limits

$$\overline{y} \pm U$$

where $\overline{y}$ represents the mean of the measurements on the test weight and $U = k * \sqrt{u_S^2 + s_P^2 + u_O^2}$, with the value of k determined by the confidence level required for the interval.

Note:. For $n \geq 30$, $k \rightarrow z$.