SOP 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. The procedure may be applied to other calibration processes. See also NISTIR 6969, SOP 30 and NISTIR 7383, SOP 17 and 20.

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 $\overline{x}$ (*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 ($\overline{x}$) 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 in statistical control when the individual values are within the designated statistical limits. The system is out of control if values are present outside established limits for which no reasonable and correctable cause have been determined and corrected, 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 Define the Calibration Process

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 show to be constant or negligible variance contributions may be consolidated and monitored by a single control chart. Any that cannot be so considered (for example: different standard, different balance, different SOP) typically require separate control charts.

The variability of balance precision may be load dependent and must be considered. For many balances, precision is a function of load, and a distinct control chart may be required (according to the SOP) for every load tested. An F-test may be used to assess the standard deviations at varying loads to minimize the number of charts that are required in the laboratory when multiple loads have no significant difference in the standard deviation of the measurement process.

In the case of SOP 5 or 28 (NISTIR 5672) check standards are incorporated into the measurement process and designated charts will be required. 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.

In the case of SOP 4, 7, or 8, where check standards are not required at every nominal value (per this SOP, Table 1), see Section 4 in each of these procedures regarding measurement assurance. When a check standard is not used to monitor each reference or working standard, more frequent monitoring or calibration is essential to assure the accuracy and stability of the mass standards and the laboratory measurement results.

3.2 Define Type and Number of Check Standards

3.2.1 A check standard must be stable and is normally comparable to the reference standard or to the typical item submitted for calibration, depending on what is being monitored (standards or process). For lower order calibrations, the check standard should simulate the laboratory's reference or working standards to the extent feasible. It should be calibrated using a better procedure than the one being monitored to ensure that the expanded uncertainty is equal to or better than the uncertainty achievable with the process being monitored. As the uncertainty of the check standard value increases relative to the measurement process, it becomes more difficult to monitor control, drift, and potential bias. All check standards should be cared for in the same way as reference standards to prevent their damage or deterioration. Lower order check standards should be recalibrated at regular intervals according to NISTIR 6969, Good Measurement Practice (GMP) 11 on Setting and Adjusting Calibration Intervals.

3.2.2 Select check standards for mass calibration procedures as shown in in the applicable SOP or as in Table 1 in this SOP. Take care to consider balance load and variability as noted in section 3.1 to ensure that standard deviations of the measurement processes are consistent at different nominal values. Use the next larger standard deviation of the measurement process in uncertainty calculations for intermediate nominal values. When a check standard is not used at each nominal value, additional monitoring of the working standards is required and more frequent calibrations may be required.

Table 1. Recommended check standards for typical calibrations.

Procedure	Range of Measurement	Check Standard(s)
Echelon III (Class F) SOP 7, 8	5000 lb to 0.001 lb or 2268 kg to 1 mg	SOP: Use 1 chart per nominal value Optional: 2 to 3 values per balance
Echelon II SOP 3, 4, 6, 7	5000 lb to 0.001 lb and 1000 kg to 1 mg	SOP: Use 1 chart per nominal value Optional: 2 to 3 values per balance
SOP 5	Typically, 1 kg to 1 mg	Each nominal value incorporates a check standard
SOP 28	Typically, 1 kg to 1 mg	Use 1 check standard per decade (e.g., 1 kg, 100 g, 10 g, 1 g, 100 mg, 10 mg, 1 mg)

3.3 Measure – Perform Initial Measurements to Establish Control Charts

3.3.1 To establish a new control chart, make at least seven (minimum number) and preferably 12 or more, independent measurements of the check standard using the same standards, equipment, procedure, and under the same conditions that will be used to make routine measurements. No two measurements may be made on the same day. The time of day should be varied as would be typical during routine laboratory operations. 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.

3.4 Analyze Data – Perform Statistical Analysis

3.4.1 Create a control chart using the data and the statistics associated with the initial measurement results. The control chart parameters consist of the central line, and control (or "action") and warning limits that represent probabilistic limits for the distribution of results around the central line. These parameters are evaluated based on a reasonable number of initial measurements and updated as additional measurement data are accumulated. A known value is based on a higher-level calibration of the check standard that is preferably independent of the measurement system being monitored. The central line is established by the mean of measurements. If no higher-level calibration is available, the central line may be used as the known value, but this is not recommended since it will allow no evaluation of measurement accuracy or bias. Upper and lower control limits should be fixed, and adjusted after periodic evaluation when/if appropriate.

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

3.4.3 Establish the control chart parameters as follows:

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

3.4.4 Control chart parameters for Echelon III (Class F or other) may be completed as follows to track practical limits, although this approach may not show potential trends as well as statistical limits show:

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

3.5 Analyze a Measurement Process on a Control Chart

Use the following criteria to analyze control charts. Limits on control charts should be fixed at this time to avoid shifting limits as data continues to be collected and entered. Moving limits may indicate in control status of a process and not show potential trends as well. If uncertainty values are determined at fixed intervals, the evaluation and fixing of limits in control charts should be completed at the same time. During the control phase, after analysis (e.g., F-test and t-test) determines data consistency, data may be pooled and all data included in setting new limits.

3.5.1 If plotted points are stable and randomly distributed within the warning limits, the system is in control. If the process is in control (and the process statistic is normal), 99.73 % of the points will fall within the control/action limits. Observations outside the limits, or systematic patterns within, should be investigated and corrected as appropriate or possible.

3.5.2 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 re-measured results are not within limits, consider the measurement process "out of control" (See 3.5.3.2). Reject all data obtained since the last "in-control" measurement and take corrective action (hence “action” limit). Keep the “out of control” data in the chart and flag it as problematic. 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 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.

3.5.3 Additional guidelines for the evaluation of control charts based on probability statistics should be used to evaluate the presence of drift, shifts, and possible bias. Examples for further evaluation include:

3.5.3.1 Any single point or series of points outside of three standard deviations (keeping in mind the probability that a three standard deviation limit could reasonably expect to allow three points out of one thousand to be outside these limits);

3.5.3.2 Two of the last three points are above (or below) two standard deviations;

3.5.3.3 Four of the last five points are above (or below) one standard deviation;

3.5.3.4 Eight consecutive points are on one side of the mean or reference value;

3.5.3.5 Six points in a row are trending up (or down); and

3.5.3.6 Fourteen points are alternating up and down (sawtooth pattern) about the mean or reference value.

3.6 Analyze Reference Values

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

Note: this assessment should be conducted during the initial creation of the control chart and periodically thereafter during monitoring and control of the process.

3.6.2 When a Reference Value for the check standard is more than one standard deviation from the mean value, it may necessitate obtaining an updated calibration or evaluating the bias or deviation further to determine the cause and correct it. The deviation or offset must be considered with respect to the reported uncertainty for the reference value as well as the measurement process being used to evaluate the value. In some cases, such as with very large tolerances, a measurement process offset might be quite small compared to the tolerance, in which case the offset can be used as an uncorrected systematic error in the uncertainty calculations. Where tolerances are small or uncertainty requirements stringent, updated calibrations may be required.

3.6.3 The Normalized Error, E_n, may be used to compare the mean mass value of the check standard, S_c to the calibrated reference value, $S_{c(cal)}$ taking care to ensure adequate metrological traceability for the reference value. The expanded uncertainty of the reference value is taken from the calibration certificate and the uncertainty of the mean value is determined using the following equation, based on the standard deviation of the process from the control chart, s_p, where n is the number of relevant data points; other components are the uncertainty for the standards, u_s, and any other critical components to be considered, u_o. The coverage factor, k, is determined based on the desired level of confidence and the associated effective degrees of freedom.

$$u_c = \sqrt{\frac{s_p^2}{n} + u_s^2 + u_o^2 + u_o^2}$$
$$U_{\bar{S}_c} = u_c * k \qquad \text{Eqn. (1)}$$

E_n is calculated using the following equation:

$$E_n = \frac{\left|\bar{S}_c - S_{c(cal)}\right|}{\sqrt{U_{\bar{S}_c}^2 + U_{S_{c(cal)}}^2}} \qquad \text{Eqn. (2)}$$

The E_n value must be less than one to pass. If the E_n is greater than one, corrective action is required.

Note: the reference value may exceed acceptable limits even when a measurement process appears statistically in control. When tolerances are large compared to process variability, reference value offsets may simply be noted.

3.7 Improve the Process

During analysis of the measurement process or the reference and check standards, opportunities for corrective action and improvement action may be identified. Control charts or values that do not have normal distributions or which have significant differences between observed measurement results and reference values are cause for action of some type. During review of the charts, the first step is to identify the source/cause of the concern, and the associated actions will generally follow. The following table can provide some ideas.

Table 2. Assignable causes and example action items.

Assignable Cause – Source of Problem	Example Action Item
Standard deviation has increased after new staff member is hired	Staff might need training, instruction, or oversight; uncertainties may need to be increased
Standard deviation has increased over the past year	Balances may need service; uncertainties may need to be increased
Standard deviation has gotten smaller over the past few months (due to staff training or balance service)	Uncertainties may benefit from being reduced (may also not matter if tolerances are large enough)
Observed values of the check standards suddenly changed. Possibilities might be standards were switched, standards were damaged, values of standards after calibration were not updated properly, new software was implemented without proper validation, standards recently returned from calibration and the provider cleaned them.	Identify root cause – then take appropriate action. For example, replace standards, update values of standards, validate software, contact calibration provider. Evaluate whether the shift corresponds in direction and magnitude to changes in the calibration value of standards used.
Standard values are demonstrating a drift over time. Possibilities may be that standards were not equilibrated long enough before being placed into service; standard or check standard type and design might be inherently unstable.	Allow standards to equilibrate longer; replace unstable standards or check standards.

3.8 Control – Updating Control Chart Parameters

Update 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, updating 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. Periodic review of control charts on a quarterly basis, and no longer than an annual basis, is common with a workload that supports ongoing staff proficiency. Performing fewer than seven control measurements in any given year in related measurement parameters raises the question of whether there are adequate measurements being made to maintain staff proficiency. Pooling of related processes and quantities may be considered when determining the best frequency for updating control chart parameters.

Calculate $\overline{x}$ 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 do not agree within statistical limits, establish new parameters and limits using the most recent data and note the reasons for not using previous data or correct the causes of variation. If portions of the process or standard variations pass, be sure to note the degrees of freedom to support uncertainty analyses and coverage factors. If no significant differences between the data sets are found, pool all data and calculate new control chart parameters based on all existing data.

3.9 Control the Process – 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.

Check standard measurements should be plotted in control charts as close to real time as feasible to effectively monitor the measurement process and to prevent the possible release of questionable data that may result in the recall of laboratory work.

4 Transfer of Measurement Statistics

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 each SOP for calculation of measurement uncertainty using the process variability, s_p. The value of the weight being calibrated is said to be within the limits if the combined mean of the measurements on the test weights and the uncertainty, which is expanded following SOP 29, are within limits.

Appendix A - Checklist for Creation or Evaluation of Control Charts

This checklist may be used as design criteria or to assess the quality of control chart construction during an internal audit. It does not evaluate the control status of control charts. Additional evaluations are required for evaluation of stability and trends.

Control Charts cover entire scope (are available for each measurement parameter).	
Control charts have titles (or otherwise include)	
	Laboratory name or other identifying information
	SOP(s) used to generate measurement result
	Balance/Other equipment
	Standard Identification, Check standard Identification
	Nominal load/value
	Dates/chronology/time periods if not used on the x axis
	Legends if multiple series and extra information are plotted; to avoid confusion
Control charts have x and y axis with labels and	
	All measurement values have units of measure associated with them
Control charts have	
	Mean value (and units)
	Standard deviation (and units)
	Degrees of freedom or number of points noted (if not obvious or if small number)
	Alternative summaries of this information and suitable references (e.g., tables)
	Printed summary reports of control data with the charts
	Reference values and source and bias if appropriate/available
Control charts have limits that are based on	
	Statistical controls of:
	Warning limits (i.e., two standard deviations) and
	Action/control limits (i.e., three standard deviations)
	Or, specification limits (e.g., tolerances or smaller ratios of tolerances)
Good Items (on chart or in spreadsheet or database table summaries). Control charts have (when applicable and meaningful if not otherwise noted, e.g., in a table)	
	Tolerances: when applicable
	Uncertainties: for the reference value, check standard, and the process output
	Equipment information: device readability, configuration (stability settings/timing)
	Standard information: calibration date and interval information
	Responsible staff: need on chart or in database
	Status of control: in control, out of control with latest date of review
	History: previous limits and history of the chart/data with F-test and/or t-test results

This page is intentionally blank.