NISTIR 5672

Advanced Mass Calibration and Measurement Assurance Program for State Calibration Laboratories (2019 Ed)

Kenneth L. Fraley
Georgia L. Harris

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.5672-2019

National Institute of Standards and Technology
U.S. Department of Commerce

NISTIR 5672

Advanced Mass Calibration and Measurement Assurance Program for State Calibration Laboratories (2019 Ed)

Kenneth L. Fraley
retired, Oklahoma Bureau of Standards

Georgia L. Harris
NIST Office of Weights and Measures

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.5672-2019

May 2019

U.S. Department of Commerce
Wilbur L. Ross, Jr., Secretary

National Institute of Standards and Technology
Walter Copan, NIST Director and Undersecretary of Commerce for Standards and Technology

National Institute of Standards and Technology Interagency or Internal Report 5672
Natl. Inst. Stand. Technol. Interag. Intern. Rep. 5672, 58 pages (May 2019)

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR. 5672-2019

Preface

This publication was originally written by Ken Fraley, metrologist with the State of Oklahoma, and Georgia Harris, physical scientist with the NIST Office of Weights and Measures. Ideas from final users regarding publication content were sought at the 1994 National Conference on Weights and Measures meeting held in San Diego, CA. The publication was written to provide guidance for calibration laboratories in their desire to provide improved precision mass calibrations and for the NIST Office of Weights and Measures to ensure uniform evaluation of laboratories seeking to make mass measurements and be accredited at the advanced level of mass calibration.

Since the earlier editions (1995, 2005, 2012, 2014, 2018) numerous practical questions have been raised and additional input has been sought from mass calibration experts. This edition seeks to enhance prior editions and provide additional guidance. Additional weighing designs, equations for between-time standard deviations, and updates for uncertainty analysis are included as well. The 2012 edition was updated to address updated definitions in metrological traceability. The 2014 update was primarily editorial with minor clarifications. SOP 5 and SOP 28 underwent additional updates to improve consistency in mass calibration uncertainties and reporting. In 2018 and 2019, additional clarifications were provided.

Beginning with the 2012 revision, updated copies of SOP 5, SOP 28 (3-1 Weighing Designs and Advanced Weighing Designs respectively) were included along with ideas and content for spreadsheet analysis. Enhancements in response to the many technical questions from laboratories already working at this level and based on the review of numerous annual submissions for laboratories seeking formal Recognition at this level are also provided. The Office of Weights and Measures is responsible for providing technical support and guidance to the State legal metrology laboratories to ensure uniformity in the legal metrology measurement infrastructure; this publication is intended to provide support and guidance not only for State weights and measures laboratories, but also for other calibration laboratories seeking to implement advanced weighing designs.

Two supplemental spreadsheets are available on the NIST Office of Weights and Measures website. One provides detailed examples and calculations for uncertainties starting with restraint and calibration information from the NIST Mass Code report that was created by Val Miller. The additional NISTIR 5672 Equations spreadsheet consists of hundreds of additional calibration designs and variations of restraint and check standards positions. Both are posted with this publication, and with SOP 5 and SOP 28. This document and data reduction methodology described in this document are based on the use of NIST weighing designs and the NIST Mass Code software.

Acknowledgements
In the 1995 edition, the authors thanked the late M. Carroll Croarkin (NIST) for providing between-time standard deviation formulae and assistance regarding updating mass calibration uncertainties to meet the ISO Guide to the Expression of Uncertainty in Measurement, to Jerry L. Everhart (JTI Systems, formerly with EG&G Mound) for providing guidance in Process Measurement Assurance Programs, and to all of the metrologists who regularly participate in OWM training and regional meetings for their questions, comments, and desire to make precision mass measurements to the best of their capabilities. For the 2005 update, the authors thanked Hung-kung Liu of the NIST Statistical Engineering Division in supplying missing factors for the

between-time standard deviation equations and for his critical review of the document and supplemental spreadsheet calculations. The 2018 edition included additional reviews by Aaron Aydelotte (Oregon), L.F. Eason (retired North Carolina), and Carol Hockert (retired NIST), in addition to Val Miller and Tim Osmer both of the NIST Office of Weights and Measures.

About the Authors

Ken Fraley is a retired metrologist with the State of Oklahoma. He has over 20 years of experience in making precision mass measurements using weighing designs as described in this publication. He has carried out extensive experimentation and implemented measurement assurance programs to ensure that measurements made in the Oklahoma laboratory are consistent and uniform with those at the national level. As background experience, Ken developed the first draft of this document based on discussions with NIST and early users of weighing designs in the State laboratories and provided a critical review of Advanced Laboratory Auditing Program (LAP) problems submitted after the first Advanced Mass Metrology seminar in 1993. Mr. Fraley coordinated and analyzed several interlaboratory comparisons among laboratories working at the advanced level described in this publication. He also analyzed data, prepared preliminary and final analyses, and presented results of many interlaboratory comparisons conducted at basic, intermediate and advanced levels. In this edition, Ken helped clarify ideas that now have over twenty years of refinement, introduced extensive spreadsheet usage, and provided additional graphic content.

Georgia Harris is a Program Leader in the NIST Office of Weights and Measures. She provided direction and encouragement to Ken in developing the ideas and provided editorial support for the initial publication. Beginning with the 2012 revision, Ms. Harris has included updated copies of SOP 5, SOP 28 and managed the ongoing development of this document. The Office of Weights and Measures is responsible for providing technical support and guidance to the State legal metrology laboratories to ensure uniformity in the legal metrology measurement infrastructure; this publication is intended to provide support and guidance not only for State weights and measures laboratories, but also for other calibration laboratories seeking to implement advanced weighing designs.

Contents

Advanced Mass Calibration and Measurement Assurance Program for State Calibration Laboratories

Abstract

This publication provides guidelines for evaluating data from advanced mass calibrations and for establishing measurement assurance programs in precision mass calibration laboratories. The NIST Office of Weights and Measures (OWM) uses these guidelines when evaluating advanced mass calibration data for State laboratories that request technical support, recognition, and/or NVLAP accreditation.

Key words

Calibration procedures; mass calibration; metrological traceability; weighing designs.

Program Objective

This publication provides guidelines for evaluating data from advanced mass calibrations and for establishing measurement assurance programs in precision mass calibration laboratories. The NIST Office of Weights and Measures (OWM) uses these guidelines when evaluating advanced mass calibration data for State laboratories that request technical support, Recognition, and/or NVLAP accreditation to NIST Handbook 143, Program Handbook, based primarily on ISO/IEC 17025[1]. The guidance and requirements provided in this publication are supplemental to requirements in NIST Handbook 143 and the ISO/IEC 17025 standard.

Advanced mass calibrations use weighing designs, such as those originally found in NBS Handbook 145[2] (SOP 2, 4, and 5), NISTIR 6969, Selected Publications[3], NBS Technical Note 952[4], and the NIST/SEMATECH e-Handbook of Statistical Methods[5] that require the use of computer software (Mass Code) for the data reduction. These weighing designs are normally used when high precision (low uncertainty) mass measurement results are sought, although weighing designs can be used at any uncertainty level. The uncertainty reported using advanced weighing designs is based on the historically observed process of similar measurements and is very dependent upon correct procedures for defining these processes.

---

[1] ISO/IEC 17025:2017, General requirements for the competence of testing and calibration laboratories.

[2] Taylor, John K. And Henry V. Oppermann, NIST [NBS] Handbook 145, Handbook for the Quality Assurance of Metrological Measurements, November 1986.

[3] NISTIR 6969, Selected Laboratory and Measurement Practices, and Procedures, to Support Basic Mass Calibrations, 2019 Edition. Contains Good Measurement Practices (GMP) 11 on Assignment and Adjustment of Calibration Intervals for Laboratory Standards and 13 on Ensuring Metrological Traceability that are referenced in this publication.

[4] Cameron, J. M., M. C. Croarkin, and R. C. Raybold, Technical Note 952, Designs for the Calibration of Standards of Mass, June 1977.

[5] NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/, 2019-04-10.

Reference calibrations (such as those provided by NIST) provide traceable measurement results for standards at one point in time and periodic calibrations are essential for ongoing assurance of metrological traceability. The major advantage in this Advanced Mass program is the ability to evaluate reference/working standards and the measurement process over time, providing ongoing assurance regarding accuracy and metrological traceability of the mass standards for both the laboratory and its customers. Ongoing evaluation of the measurement process provides the laboratory with data that can be used to establish or adjust calibration intervals for reference standards, check standards, and working standards. The measurement assurance program is also critical for defining and reporting realistic uncertainties.

Program Prerequisites

The following items are listed as general guidelines for a laboratory conducting an internal technical audit (readiness assessment) of its program for suitability in the advanced mass calibration program for States. The forms and structure of the pre-work associated with the Advanced Mass Seminar may be used to document the audit/assessment. These guidelines have been established based on good measurement practices, good laboratory practices, and a similar fee-funded program (NIST Mass MAP) operated by the NIST Mass and Force Group. Many specific technical recommendations were originally taken from NIST Handbook 143, State Weights and Measures Laboratory Program Handbook (2007 edition) and NIST Handbook 150-2G: NVLAP Calibration Laboratories, Technical Guide for Mechanical Measurements (2004) both of which are being updated as of 2018. Deviations from recommendations are occasionally made when data is available to support it; however, judgments should be made carefully when evaluating data, since some deviations from these practices will inadvertently increase measurement uncertainties and may contribute to measurement errors.

Staffing/Training (Supplemental to requirements in Section 6.2, ISO/IEC 17025)

< Satisfactory completion of a Fundamentals of Metrology course and the OWM Mass Metrology Seminar (or Basic Mass and Intermediate Metrology Training courses if completing seminars prior to 2012) *and* Laboratory Auditing Program (LAP) problems within the last five years are expected before attendance at the Advanced Mass Training course which is traditionally taught in odd-numbered years.

< Although required for NIST Recognition, attendance at the Regional Measurement Assurance Program (RMAP) meetings is not required to perform advanced mass calibrations. However, regular updates on precision mass procedures and issues are often provided at the regional meetings and proficiency tests are often coordinated through the RMAPs and will be required for Recognition or Accreditation.

< Satisfactory completion of proficiency tests at Mass Echelon II, using SOP 5 as a minimum procedure.

< Satisfactory completion of the Advanced Mass Training course and Advanced LAP problems and successful application of these guidelines are expected before Recognition or accreditation at the Echelon I level as described in NIST Handbooks 143:2019 Program Handbook and NIST Handbook 150-2, NVLAP Calibration Laboratories, 2016, and Associated Annexes. In addition to successful completion of proficiency tests. The problems ensure that suitable measurement assurance and uncertainty analysis data is in

place in the laboratory to provide evidence of valid measurement results; the problems also provide additional demonstration of competency.

Facilities and Accommodation (Supplemental to requirements in Section 6.3, ISO/IEC 17025)

< Environment – The temperature for the laboratory where mass measurements are made should be selected at a point between 18 °C and 23 °C, with maximum changes of ± 0.3 °C per hour during a calibration series and a maximum change of ± 0.5 °C per 12 hours. Air flow must be low enough so that it does not interfere with balance or mass comparator operation. Humidity should be set between 40 % and 60 % relative humidity (e.g., 45 % ± 5 %). Environmental conditions must be monitored according to technical criteria and measurements should not be made when prescribed conditions are not maintained. Deviation from stated environmental parameters requires a thorough evaluation of the impact of the deviation on measurement errors and uncertainties. Additional details are documented in SOP 5 and SOP 28. Air flow for mass laboratories should be less than 2 m/s to minimize impact on the balances and weighing processes.

< Vibration – The laboratory location and design should be such as to avoid or minimize potential sources of vibration that will interfere with precision mass calibration.

< Cleanliness – Good housekeeping practices and cleanliness is essential for working at this level of mass calibration. Contamination from dust, hair, paper, shipping/packaging/storage materials like felt and velvet, and other air contaminants has been found to be a critical concern and will cause statistical failures in the measurement process or an inability to attain the expected levels of uncertainty. Office spaces and desks do not belong in the laboratory.

Equipment (Supplemental to requirements in Section 6.4, ISO/IEC 17025)

< Balances – Balances must be evaluated for suitable capacity, resolution, and repeatability and their ability to produce measurement results with standard deviations that meet the applicable uncertainty and tolerance requirements. A list of the laboratory's balances and process control chart data may be submitted to OWM for evaluation (required for State laboratories seeking Echelon I recognition). Control data for each balance should be available for SOP 5 calibrations, and consist of the list of items shown in NISTIR 6969, SOP 9 for measurement assurance in mass calibrations, with at least the following:

- Balance manufacturer, model number and serial number;
- Capacity and resolution;
- Balance configuration settings; and
- Pooled standard deviations (accepted within-process standard deviations, s_w, and accepted long term standard deviation of the check standard, s_t), showing adequate number of degrees of freedom, loads, and specific weighing designs.

Laboratory balances will be evaluated to determine their suitability for this program. Balances and mass comparators should be kept clean and be well maintained. (See NISTIR 6969, GMP 11 for recommended service/calibration intervals). Minimum balance performance specifications recommended for this program are as follows:

Table 1. Balance Performance Limits.

Balance Nominal Design Loads	[a]Standard Deviation Should be Less Than:
20 kg	1.5 mg
10 kg	0.20 mg
1 kg	0.050 mg
600 g, 500 g, 400 g, 200 g	0.025 mg
60 g, 50 g, 40 g, 20 g	0.0050 mg
6 g, 5 g, 4 g, 2 g	0.0025 mg
600/500/400/200 mg	0.0013 mg
60 mg, 50 mg, 40 mg, 20 mg	0.00050 mg
6 mg, 5 mg, 4 mg, 2 mg	0.00030 mg

[a]The standard deviations in this table were developed by calculating 1/3 of OIML Class E_1 tolerances, accounting for typical uncertainties associated with the standard and other factors and developed by working backwards from a target expanded uncertainty. Many laboratories will not have balances or processes that can achieve these results on a routine basis.

- Barometer – A barometer having documented accuracy of ± 65 Pa (0.5 mmHg) with evidence of metrological traceability (typically from an accredited laboratory) must be available.
- Thermometers – Thermometers to measure air temperature having accuracy of ± 0.1 °C with evidence of metrological traceability (typically from an accredited laboratory) are required. Temperature measurements made at this level are generally made within the balance chambers.
- Hygrometer – Percent relative humidity should be measured with an accuracy of ± 5 % and have evidence of traceable measurement results (typically from an accredited laboratory).
- An environmental recording device is required for monitoring laboratory conditions. Even though environmental corrections are made during the mass measurement process, ensuring environmental stability during the 24-hour period preceding a calibration (particularly for temperature) is important to ensure proper thermal and environmental equilibration of mass standards. Digital monitoring that can be downloaded and statistically evaluated is preferred to manually maintaining chart recorders with paper graphs.
- Computer and printer – A computer of enough memory and processing ability is essential. In the laboratory, every effort must be made to minimize the impact of introducing temperature gradients in the measuring areas. Laptops or other wireless or panel monitors are preferred to large heat-producing systems however, even when a laptop computer is used, exhausted air must be considered when situating equipment for these measurements. If the laboratory will print reports, the printer should never be in a precision mass laboratory.

Standards and Metrological Traceability (Supplemental to requirements in Section 6.5, ISO/IEC 17025)

< Design – All reference and check standards used at this level should be Type I (one-piece design) to enable actual density determinations, to provide the necessary stability, and to act as surrogates to the reference standards). Reference standards should have measured densities, be highly polished, and have been assessed to comply with limits on magnetic susceptibility as required in ASTM and OIML documentary standards, in the following decade denominations: 1 kg, 100 g, 10 g, 1 g, 100 mg, 10 mg, 1 mg. "ASTM Class 1, Type 1, Grade S" or "OIML Class E_2" verbiage may be used if specifying weights for purchase. Additional check standards above 1 kilogram are needed to handle the entire range of calibration services, e.g., 10 kg.

< Reference Standards (formerly called Primary Standards) – A minimum of two 1 kg reference standards (four 1 kg reference standards are recommended) with calibration and density determination needed and metrological traceability to the International System of Units (SI) usually through a National Metrology Institute, such as the U.S. National Institute of Standards and Technology. Two reference standards should be calibrated at least every four years. The second set of two 1 kilogram standards should then be calibrated two years after the reference standards. In this way, calibration values used in the laboratory will be less than two years old. Even though a measurement assurance program that monitors the reference kilogram standards will be in place to demonstrate ongoing stability and validity of the mass values, mass calibrations are performed through comparison methods and all standards may be changing or drifting and fail to identify systemic changes. If only two reference standards are available, and if they are used with equal frequency, the measurement assurance program will not be considered adequate without some type of verification using standards from outside the laboratory that have recent suitable calibrations.

When most of the work in the laboratory will be the calibration of 100 gram weight sets, it is also desirable to have two 100 g starting restraints, and two 100 g check standards that are calibrated and cycled in a similar fashion as the 1 kg standards.

< NISTIR 6969, GMP 11, Good Measurement Practice for Assignment and Adjustment of Calibration Intervals for Laboratory Standards, and GMP 13, Good Measurement Practice for Ensuring Traceability (or other equivalent procedure), must be implemented in the laboratory to document metrological traceability and calibration intervals that the laboratory will follow. (See discussions on Proficiency Tests and Graphs and Control Charts.)

< Check standards – Check standards (sometimes called control standards) must have suitable calibrations with evidence of metrological traceability. Having the check standards calibrated by a National Metrology Institute or an Accredited laboratory provides an effective mechanism to identify and evaluate biases that may be occurring in the measurement processes that would otherwise go undetected; having an external calibration is essential for evaluating bias in addition to monitoring the stable output of the measurement process. Appropriate supplier evaluation must be completed before obtaining calibrations from an accredited laboratory to ensure that the measurements of interest are

on the accreditation Scope and to ensure that sufficiently small and validated uncertainties are available. It is also a good idea to request information about the provider laboratory's latest participation in a proficiency test at the Echelon I level of work.

Measurement Assurance (Control) (Supplemental to requirements in Section 7.7, ISO/IEC 17025)

< The laboratory needs to have a measurement assurance (control) system already in place, at least for SOP 5, before trying to perform advanced mass calibrations. Practical, hands-on experience in the laboratory is essential to making good mass measurements. A current measurement assurance system and data are essential for initial statistical analysis, for demonstrating measurement proficiency, and/or justifying any deviations from these recommendations which also require documentation and validation.

Advanced LAP Problems[6]

See Figure 1 for a graphic view of the components required in a complete analysis of the Advanced LAP problems.

Laboratory Auditing Program (LAP) problems are used for establishing a baseline for the initialization of the check standards, to provide initial data to assess the between-time component of the measurement process, for providing validation on uncertainty statements, and for evaluating the proficiency level at which the laboratory uses the Mass Code. The data collected in the LAP problems is reduced by each laboratory using the Mass Code. Each qualified metrologist must complete training and the Advanced LAP problems and be able to reduce and analyze their own data. Each laboratory is responsible for graphing and analyzing data when determining the "in control" or "out of control" condition of their standards (see sections on Establishing Measurement Control and Graphs and Control Charts). Observed surveillance values must be compared to the reported reference values to determine the level of control. A copy of all data, data files, reports, graphs, and final analysis for State laboratories are to be sent to OWM for evaluation along with the most recent reference calibration reports for the standards used. The final written analysis, demonstrating a thorough understanding of the measurement assurance system and uncertainty analysis at this level of work is considered the most critical component of completing the Advanced LAP problems.

NOTE: Standards should not be cleaned using solvents during the initial data-collection period as these tests will provide data for determining the total and between-time standard deviations for each series. Cleaning standards changes their mass values and may invalidate the calibration. Cleaning plans and procedures must be documented as a part of the laboratory procedures. At least 7 to 10 days are required for environmental equilibration on standards that have been cleaned with solvents prior to recalibration; however, several months are often required. New mass standards may require several years before suitable stability has been attained.

[6] Laboratory Auditing Problems are assigned to each State metrologist upon completion of the Advanced Mass seminar. Successful completion of the LAP problems and applicable proficiency tests are required by State laboratories to demonstrate competency prior to receiving Recognition or Accreditation.

LAP Problem 1: Ten (10) complete runs on reference/working State standards from 1 kg to 1 mg. (Initial data from the first one or two runs may be submitted to be sure the laboratory is on the right track.)

LAP Problem 2: Two (2) complete runs on standards from 30 kg to 2 kg.

The series selected when ascending from 1 kg may be the same as those used when descending; however, a single restraint is usually used, and the between-time standard deviation formulae must be derived or obtained from the referenced design job-aid spreadsheet.

If the laboratory maintains both metric and avoirdupois standards with metrology traceability to the SI, one additional LAP problem should be conducted. The avoirdupois standards are calibrated through a crossover method at the 1 kg level against 2 lb and a 2 lb summation made of the reference 1 lb standards. The 1 lb values are assigned reference values through additional measurements and the entire avoirdupois system is calibrated with the 1 lb standards as the starting restraint. Special tare weights (92.815 g) can be obtained to facilitate this process and may be calibrated using SOP 5.

LAP Problem 3: Two (2) complete runs on standards from 50 lb to 1 lb.

Follow up: Note measurement assurance guidelines and traceability for using the initial LAP problems in continuing measurement assurance.

Note: Two complete runs are not adequate to provide data for establishing initial limits nor for establishing a baseline for acceptable measurement assurance nor for validating uncertainties at this level. If the laboratory plans to provide internal calibration results or extended service to customers, additional data must be obtained and analyzed. Process statistics determined with limited data must have uncertainties that reflect the actual degrees of freedom. The laboratory may limit their application of this process to the smaller mass levels (e.g., 1 kg to 1 mg, or 100 g to 1 mg). LAP problems 1 and 2 may be combined if a limited range will be used for this level in the laboratory. For example, ten runs from 10 kg to 1 mg would be another acceptable approach. If the laboratory does not plan to use the Mass Code for avoirdupois standards, LAP problem 3 is not required.

Alternative ranges may be selected in discussions with the Office of Weights and Measures, for example build-up calibrations to 500 lb is a common supplemental or alternative set of data in a laboratory with the other ranges already in place.

In addition to the measurements described here, a full summary assessment, and technical audit should be completed (such as updating the Advanced Mass Seminar pre-work).

Establishing Measurement Controls

Process Evaluation

For each combination of a weighing design, specific load, and specific balance (e.g., a 4-1 design, at a 1 kg load, on serial number ##### balance), a measurement control process should be defined, and data collected to characterize both the standard uncertainty of the process, s_w, and the standard uncertainty for the standard over time, s_t. These values are essential for correctly reducing measurement data and calculating the uncertainty assigned to each mass value in a report. This process is more than simple statistical process control because the process is evaluated with an integrated F-test in each series and assigned values for the standards and check standards are evaluated with each run using a t-test.

In a surveillance test of reference/working standards from 1 kg to 1 mg, there are seven series as characterized below (a 5-2-2-1 set of standards may be substituted for the 5-3-2-1 set):

Table 2. Example weighing designs for a 1 kg to 1 mg set of standards.

Series	Weighing Design (Tech Note 952)	Decade Balance Loads
1	A.1.2 (1, 1, 1, 1) (4-1)	1 kg
1A (option)	A.1.4 (1, 1, 1, 1, 1) (5-1)	1 kg
2	C.2 (5, 3, 2, 1, 1, 1)	600, 500, 400, 200 g
(optional)	A.1.2 or A.1.4	100 g
3	C.2 (5, 3, 2, 1, 1, 1)	60, 50, 40, 20 g
4	C.2 (5, 3, 2, 1, 1, 1)	6, 5, 4, 2 g
5	C.2 (5, 3, 2, 1, 1, 1)	600, 500, 400, 200 mg
6	C.2 (5, 3, 2, 1, 1, 1)	60, 50, 40, 20 mg
7	C.1 (5, 3, 2, 1, 1) or C.2	6, 5, 3, 2, 1 mg

NOTE: When data is reduced using the Mass Code, switching balances within a series may result in artificially low process standard deviations and will result in errors in between-time calculations.

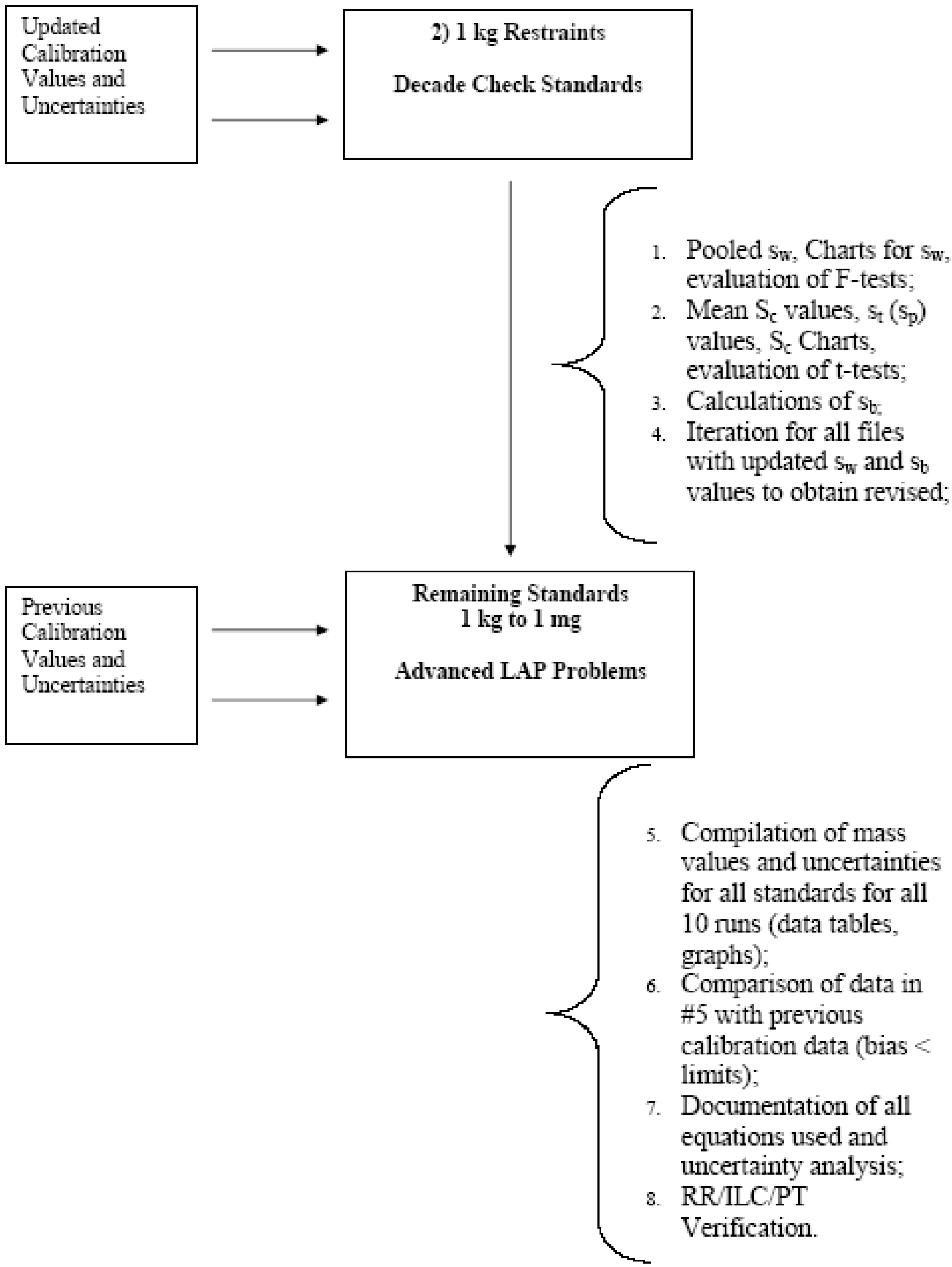

Figure 1. Graphic of Data and Analysis Needed for Advanced LAP Problems.

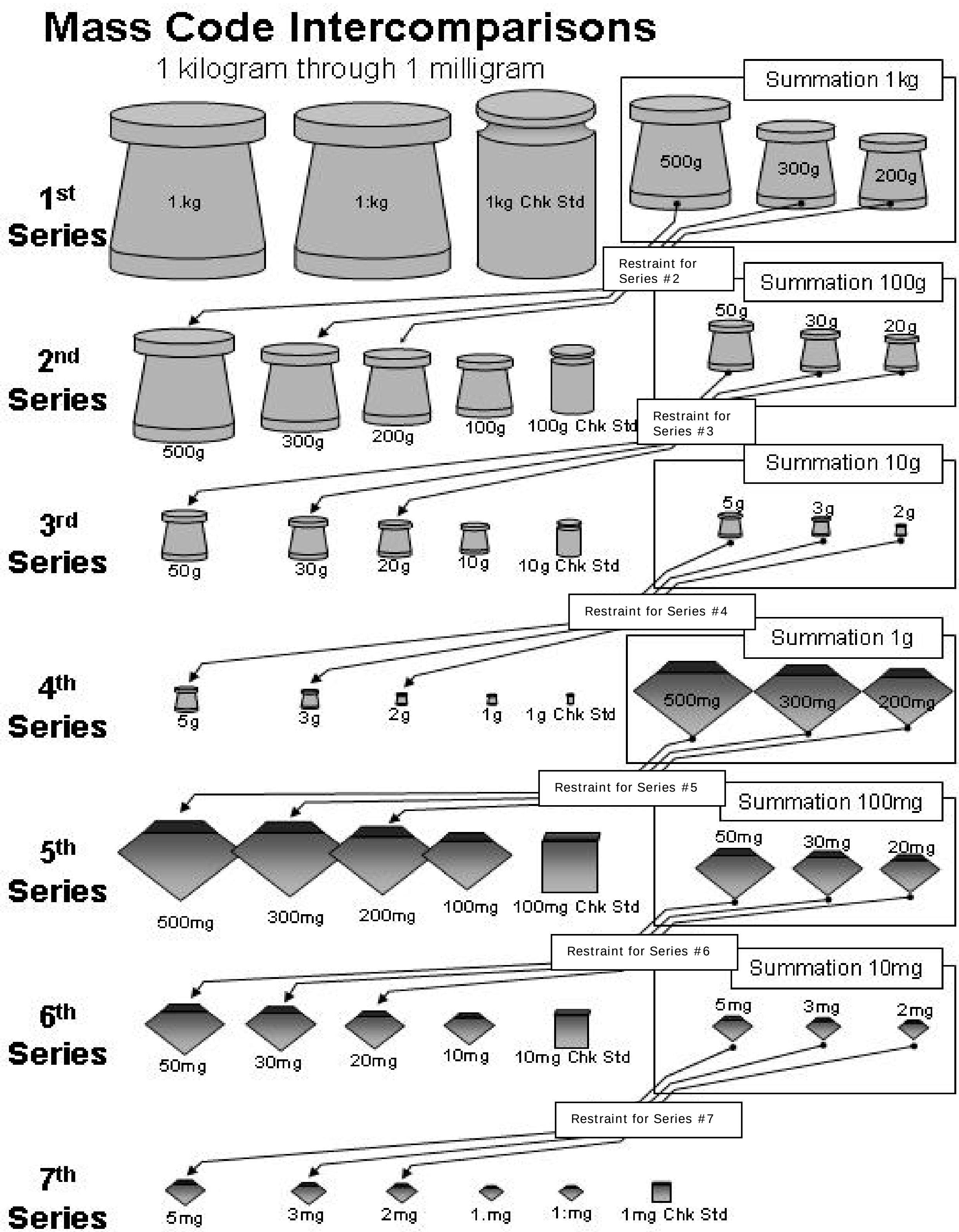

Figure 2. Example of Dissemination from 1 kg Standards.

Measurement Matrix

1 kg through 1 mg

Series 1	P1.kg	P1:kg	C1kg	Σ1kg	
1st Double Sub	+	−			1 kg load
2nd Double Sub	+		−		1 kg load
3rd Double Sub	+			−	1 kg load
4th Double Sub		+	−		1 kg load
5th Double Sub		+		−	1 kg load
6th Double Sub			+	−	1 kg load

Series 2	P500g	P300g	P200g	P100g	C100g	Σ100g	
1st Double Sub	+	−	−	+	−		600 g load
2nd Double Sub	+	−	−		+	−	600 g load
3rd Double Sub	+	−	−	−		+	600 g load
4th Double Sub	+	−	−				500 g load
5th Double Sub	+		−	−	−	−	500 g load
6th Double Sub		+	−	+	−	−	400 g load
7th Double Sub		+	−	−	+	−	400 g load
8th Double Sub		+	−	−	−	+	400 g load
9th Double Sub			+	−	−		200 g load
10th Double Sub			+	−		−	200 g load
11th Double Sub			+		−	−	200 g load

Series 3	P50g	P30g	P20g	P10g	C10g	Σ10g	
1st Double Sub	+	−	−	+	−		60 g load
2nd Double Sub	+	−	−		+	−	60 g load
3rd Double Sub	+	−	−	−		+	60 g load
4th Double Sub	+	−	−				50 g load
5th Double Sub	+		−	−	−	−	50 g load
6th Double Sub		+	−	+	−	−	40 g load
7th Double Sub		+	−	−	+	−	40 g load
8th Double Sub		+	−	−	−	+	40 g load
9th Double Sub			+	−	−		20 g load
10th Double Sub			+	−		−	20 g load
11th Double Sub			+		−	−	20 g load

Series 4	P5g	P3g	P2g	P1g	C1g	Σ1g	
1st Double Sub	+	−	−	+	−		6 g load
2nd Double Sub	+	−	−		+	−	6 g load
3rd Double Sub	+	−	−	−		+	6 g load
4th Double Sub	+	−	−				5 g load
5th Double Sub	+		−	−	−	−	5 g load
6th Double Sub		+	−	+	−	−	4 g load
7th Double Sub		+	−	−	+	−	4 g load
8th Double Sub		+	−	−	−	+	4 g load
9th Double Sub			+	−	−		2 g load
10th Double Sub			+	−		−	2 g load
11th Double Sub			+		−	−	2 g load

Series 5	P500mg	P300mg	P200mg	P100mg	C100mg	Σ100mg	
1st Double Sub	+	−	−	+	−		600mg load
2nd Double Sub	+	−	−		+	−	600mg load
3rd Double Sub	+	−	−	−		+	600mg load
4th Double Sub	+	−	−				500mg load
5th Double Sub	+		−	−	−	−	500mg load
6th Double Sub		+	−	+	−	−	400mg load
7th Double Sub		+	−	−	+	−	400mg load
8th Double Sub		+	−	−	−	+	400mg load
9th Double Sub			+	−	−		200mg load
10th Double Sub			+	−		−	200mg load
11th Double Sub			+		−	−	200mg load

Series 6	P50mg	P30mg	P20mg	P10mg	C10mg	Σ10mg	
1st Double Sub	+	−	−	+	−		60 mg load
2nd Double Sub	+	−	−		+	−	60 mg load
3rd Double Sub	+	−	−	−		+	60 mg load
4th Double Sub	+	−	−				50 mg load
5th Double Sub	+		−	−	−	−	50 mg load
6th Double Sub		+	−	+	−	−	40 mg load
7th Double Sub		+	−	−	+	−	40 mg load
8th Double Sub		+	−	−	−	+	40 mg load
9th Double Sub			+	−	−		20 mg load
10th Double Sub			+	−		−	20 mg load
11th Double Sub			+		−	−	20 mg load

Series 7	P5mg	P3mg	P2mg	P:1mg	P.1mg	C1mg	
1st Double Sub	+	−	−	+	−		6 mg load
2nd Double Sub	+	−	−		+	−	6 mg load
3rd Double Sub	+	−	−	−		+	6 mg load
4th Double Sub	+	−	−				5 mg load
5th Double Sub	+		−	−	−	−	5 mg load
6th Double Sub		+	−	+	−	−	4 mg load
7th Double Sub		+	−	−	+	−	4 mg load
8th Double Sub		+	−	−	−	+	4 mg load
9th Double Sub			+	−	−		2 mg load
10th Double Sub			+	−		−	2 mg load
11th Double Sub			+		−	−	2 mg load

Figure 3. Measurement Matrix 1 kg to 1 mg.

As noted earlier, the Advanced LAP problems are the minimum recommendations for collecting data that will begin to characterize the measurement process. Ten complete runs on the reference/working standards, 1 kg through 1 mg (all seven series reduced using the Mass Code), must be made. An additional series at a 100 g load is essential to establish a secondary starting point for the calibration of 100 g kits. This means that initial data will be collected for at least these eight series.

The section on Calculations in SOP No. 28 or the Excel job aid show how to calculate each of the standard uncertainty values that should be entered in the Mass Code data file once data from the initial ten runs is available. Meanwhile, because statistical data for the new process and check standards may be unavailable, simulated values, based on knowledge of each measurement process, may be used for the Mass Code to reduce data. For the first ten runs, process standard deviation values, based on previous SOP 5 (3-1 weighing designs) are satisfactory. Using this process, only the size of the uncertainties and the statistical tests will be affected when data is reduced; the mass values are not affected. However, once actual data is available, the simulated data in the data file must be replaced, and final reports generated. All reports should be considered as “draft” during the process of gathering the initial data; the statistics and uncertainty values must be updated prior to performing the final data analysis.

Proper characterization of the measurement process is more critical when using advanced weighing designs in decade series with the Mass Code than when using routine mass comparisons with one-to-one standards. This is primarily because the statistics used by the Mass Code distribute the uncertainty from the starting restraint (reference standard) proportionally among all the weights in each series. Also, the standard uncertainty of the process (previously called random error) is distributed among all the weights in each series. This type of data reduction allows the Mass Code to assign smaller uncertainties; however, the validity of these uncertainties is very dependent upon a well-characterized measurement process. A major difference between advanced weighing designs using the Mass Code, and routine calibrations such as the 3-1 weighing design and the double substitution, is that the Mass Code uses the uncertainty of the starting restraint only. The 3-1 design and the double substitution use the uncertainty of a standard(s) at each denomination. In a 1 kg to 1 mg Mass Code calibration, the starting restraint portion of the uncertainty is distributed among all other denominations and soon becomes negligible around five grams. Therefore, weights below five grams are primarily dependent upon the standard uncertainty of the process when assigning an uncertainty to a test weight.

Uncertainty Data Input

The NIST Mass Code has four input requirements. (Line numbers may change in future versions of the Mass Code data files, so be sure to verify line numbers with data file instructions).

1) Line eight: "ran err" or "random error" for the standard is entered first; this isn’t used and zero must be entered;
2) Line eight: "sys err" or "systematic error" for the standard, s_r, is entered second; this value is taken from the [NIST] calibration report and should be divided by the coverage factor, k, that was used in reporting the uncertainty (see calculations in SOP No. 28 for formula to be used when using more than one restraint);
3) Line thirteen: the within-process standard deviation is the same as the standard uncertainty of the process, s_w. This value is used for the F-test and is entered at the beginning of line thirteen on the first series (at the beginning of the line that contains the sensitivity weight data for subsequent series). The value is based on pooled data for observed standard deviations for the process (see calculations in SOP No. 28 for formula to be used); and
4) Line thirteen: the between-time standard deviation, s_b, is entered at the end of line thirteen on the first series (at the end of the line that contains sensitivity weight data for subsequent series) and is the calculated value that measures the variation of the value for the check standard over time, s_t, less the contribution from the standard uncertainty of the process (see the calculations in SOP No. 28 for formula to be used.)

Note: the standard deviation from the measurement process, taken from the check standard, s_t, is used with the within-process standard deviation to determine the between-time standard deviation. However, the standard deviation of the check standard, s_t, is not entered into the data file. The NIST Mass Code recombines the within process and between time components to regenerate the standard deviation for the check standard based on the design, restraint, and check standard positions.

The NIST Mass Code software has been modified to conform to the NIST policy on uncertainty (see NIST Technical Note 1297) as far as possible. References to "systematic error" have been changed to "type B uncertainty"; references to "random error" have been changed to "type A uncertainty" (although recall, that no random error component is entered for the starting restraints); and references to "uncertainty" have been changed to "expanded uncertainty." The type B uncertainties are calculated as one standard deviation estimates for systematic error, and the type A uncertainties are calculated as one standard deviation estimates for random error. The expanded uncertainties are calculated as the root sum squares of the type A and type B uncertainties multiplied by a coverage factor of two.

However, to preserve the integrity of the statistical control procedure for mass calibrations, the operation of the Mass Code deviates from the NIST policy in the evaluation of type A and type B uncertainties. The policy defines type A (random) uncertainties in a global manner; i.e., as a function of both local phenomena (balance precision and long-term measurement precision) and hierarchical phenomena (uncertainties associated with previously assigned values of reference standards). The test for statistical control for each series requires a standard deviation based only on local phenomena. The Mass Code does not distinguish between these two requirements, and it will produce an improper t-test if the hierarchical uncertainties are treated as random components. The solution to this conflict is to distinguish between local and hierarchical uncertainties and to define hierarchical uncertainties as type B uncertainties.[7]

Handling the Output

To establish measurement controls once the Mass Code has been run, certain data must be extracted from each series and placed in a spreadsheet or database for storage, analysis, measurement assurance, and uncertainty calculations:

- Test Number;
- Operator ID;
- Date of Test;
- Balance ID;
- Restraint ID;
- Check Standard ID;
- Check Standard Nominal Value;
- Starting Restraint Number;
- Calibration Design ID;
- Average Corrected Temperature in Degrees C;
- Average Corrected Pressure in Pascal;

[7] Croarkin, M. C., Internal NIST Correspondence, July & August 1993.

- Average Corrected Humidity in Percent;
- Average Computed Air Density in mg/cm^3;
- Observed Standard Deviation of the Process;
- Accepted Standard Deviation of the Process;
- Degrees of Freedom;
- F-Ratio;
- Observed Correction of the Check Standard;
- Accepted Mass Correction of the Check Standard;
- t-Value; and
- Expanded Uncertainty assigned to the Check Standard.

When the Mass Code is run, it produces two files in addition to the report, one called "control" and one called "statis." The "control" file contains a string of data for each series that contains: the month, day, year, check standard identification, observed value for the check standard, balance identification, process standard deviation, degrees of freedom, weighing design identification, average temperature, range of temperature during the measurement, average pressure, range of pressure during the measurement, average humidity, range of humidity during the measurement, air density, range of air density during the measurement, operator identification, and a flag for process control results (0 = ok; a number 1, 2, or 3 (depending on the version of the Mass Code used) flags that the check standard failed, the observed standard deviation failed, or both failed). This file can be imported to a spreadsheet using a parse function so that each item is entered into a separate cell and saved. The "statis" file contains F-test and t-test data for each series run in the Mass Code. Unless an assignable blunder is detected in the data, all out of control series must be saved in the control chart file or the statistical limits will gradually become artificially small resulting in an increased number of failed tests.

Each of these files should be saved by another name and/or in another directory immediately after each run of the Mass Code.

Reviewing Mass Code Reports

Several sections within the Mass Code report must be reviewed for adequacy. Key areas are observation values, the F-test, and the t-test. If unusually high t-test or F-test values are observed, one should check for data entry errors first.

Evaluate the F-test to make sure that the observed standard deviation agrees statistically with the accepted standard deviation of the process. An F-value is quoted and immediately below this value in the report is a statement that shows whether the F-test passed or failed.

Evaluate the t-test to verify that the observed mass correction of the check standard agrees statistically with the accepted mass correction of the check standard. A t-value is quoted and immediately below the t-value in the report is a statement that says whether the check standard is in-control or out-of-control.

Should either test fail, and no data-entry errors are found, the series should be rerun. If the process is gradually changing, the t-values or F-values will usually fall in a range from two to nine. If statistical data is graphed properly, trends can be identified before they become critical. If many tests fail, it could suggest a change in the measurement process in which case the data is combined with the other data to define the new measurement parameters. Failed data with no attributable causes must still be tracked and graphed in the control files, as the F-test and t-test are both based on a 95 % confidence interval rather than a typical control chart limit of 99.73 %. Therefore, if failed data with no attributable errors is routinely and incorrectly discarded, statistical limits will be artificially reduced. Do not discard failing check standard values unless obvious data entry errors were found and corrected or there are other corrected causes such as out of limits environmental conditions that stopped work. Uncertainty values based on minimized variability due to a practice of discarding failed F-tests and t-tests will be invalid.

Graphs and Control Charts

Critical Graphs

Standard Control Charts – When advanced mass calibrations are used for surveillance testing, each weight (500 g to 1 mg) involved may be graphed separately. These data and charts can be used to verify calibration values and to determine appropriate calibration intervals. When appropriate for reference/working standards, a new accepted value may be calculated as either the mean of all values, or the predicted value from the linear fit of the data at some time six months in the future.

When the measurement process has been sufficiently characterized and advanced mass calibrations are used for routine calibrations, a graph must be prepared for the check standard at each decade to properly characterize each measurement process. Analysis of the data provides the statistics for calculating the between-time standard deviations and can be used to verify standard and check standard values. The standard deviation over time is calculated from either the standard deviation about the mean, or the residual standard deviation from the linear fit. The standard deviation is later compared with the process deviation to detect if there is a between-time component of error in the measurement process. As noted for surveillance testing, the accepted values can be calculated in one of two ways.

Measurements and control charts for two external or "monitoring" 1 kg check standards are recommended to verify the calibration values for the two 1 kg reference standards. Measurements are made between these two external standards and the reference standards using a 4-1 or 5-1 weighing design. The external kilogram standards may be part of a circulating mass package with recent NIST calibration (as used in proficiency testing) or may be maintained in the laboratory but must receive less frequent (less than 25 % as often) use than the reference standards. Analysis of calibration results over time as recorded on control charts can provide a realistic estimate of calibration uncertainty and allow the investigation of drift for standard values due to time and use. Historical analysis can also help set calibration intervals.

Process charts – Control charts for the process standard deviation (for each series/balance) are used to establish process variability and an accepted within-process standard deviation, s_w; each point can note the degrees of freedom differentiating between series. Standard deviations for each balance are plotted versus time. Plots are critiqued for outliers and degradation. A new accepted value is calculated by pooling the standard deviation for each balance.

Optional Graphs

Summation Graphs – With a measurement control program in place for 3-1 weighing designs that shows the values for summations of standards, data from the Mass Code reports may be plotted with those values (when the 3-1 design uses the summation as the check standard). A number of items should be plotted versus time: the mean for the 3-1 summation values, the upper and lower control/warning limits, the Mass Code summation values and uncertainties, and a line showing the NIST calibration value. This graph may look cluttered, but it provides enormous insight to the relationships between the NIST value, the Mass Code values, and the 3-1 values for the same group of weights. It also provides a comparison between the separate measurement processes. The 3-1 values may show greater precision than the Mass Code values. This is because in some cases the 3-1 values may be assigned using a balance with a lower standard deviation of the process than the balance used for the Mass Code.

F-values – Graphing the F-values for a series can show trends in the process and can evaluate the appropriateness of the assigned process parameters. The F-values are plotted chronologically with the mean. The graph should be analyzed for trends and for uniform distribution above and below the F-value of 1.00.

t-values – Graphing the t-values of each series allows visualization of the extent of agreement between the observed and accepted value of the check standard. It also permits a comparison with the current measurement process. The t-values are plotted chronologically with the mean. The graph should be analyzed for trends and for uniform distribution above and below the t-value of 1.00.

Proficiency Tests (PT)

Interlaboratory comparisons (proficiency tests) need to be conducted at least once every four years at the advanced level and may consist of two kits with an entire set of standards: 1 kg through 1 mg to enable Youden-type analyses. Proficiency testing reports and results are prepared using OWM template reports and approved by NIST prior to final distribution and reporting. Proficiency test results provide useful information about potential errors in the laboratory, about the uncertainty reported, and about drift of artifacts during the intercomparisons. This information provides opportunity for evaluating a laboratory's capability to meet stated uncertainties.

Evaluation Criteria for Proficiency Tests
Verification of the Reference Value

The Reference value for each PT will be verified by using the OWM procedures for selecting the Reference value. Whenever possible, Reference values for Proficiency Tests will be based on calibrations from a National Metrology Institute.

The normalized error (E_n) concept (which is used internationally) is used to evaluate the submitted value for each laboratory. The normalized error is a ratio of the difference between the observed and accepted values and the combined uncertainty (at k = 2) in the process combined by the root sum square method.

Verification of Laboratory Values

The Normalized Error, E_n, may be used to periodically 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 uncertainty of the reference value is taken from the calibration certificate and the uncertainty of the mean value is determined using the following equation, using the standard deviation of the process from the control chart, s_p, and 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 using Equation 1.

$$u_c = \sqrt{\frac{s_p^2}{n} + u_s^2 + u_o^2 + u_o^2} \qquad \text{Eqn. (1)}$$

$$U_{\bar{S}_c} = u_c * k$$

E_n is calculated using Equation 2:

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

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

Verification of Laboratory Precision

This criterion evaluates and validates the reported uncertainty of the laboratory for its suitability to the level against which the laboratory is being evaluated. At the highest level of mass calibration, the uncertainty assigned by the participating laboratory should be less than the tolerance. There are several perspectives regarding the use of tolerance and uncertainty ratios; this is only a general guideline and is not intended to be a requirement. However, if the laboratory will determine compliance to OIML or ASTM standards, the expanded uncertainty at $k = 2$, must be less than 1/3 of the applicable tolerance as shown in Equation 3.

$$U_{lab} < Tolerance \quad \text{OR} \quad U_{lab} < Tolerance/3 \qquad \text{Eqn. (3)}$$

File Management

Good observation and data records must be maintained in the laboratory consistent with technical record maintenance and document retention requirements.

Once the ten initial runs are made on the standards (LAP Problem 1) and the process parameters are defined, the original data files must be updated with the new accepted process parameters (to replace the initial data entry), and final reports can be generated. These new reports will contain the same mass values, but the quoted uncertainties will then be used to calculate the final uncertainties and reflect the true process.

Software Management

Distribution

The NIST Mass Code is only distributed by the NIST Office of Weights and Measures (OWM) upon acceptable completion of the Advanced Mass Metrology seminar. All practical steps have been taken at NIST to ensure correct results when the software is used with proper data files. However, each laboratory must verify this independently and must document the verification. Each participant should refrain from copying, transferring, sharing, or distributing software to others who have not participated in Advanced Mass training. OWM will only provide technical support to metrologists who have participated in the Advanced Mass training from the NIST Office of Weights and Measures.

Software Updates

The Mass Code will be periodically updated, and new versions will be released to trained metrologists when available. The software will not be provided by OWM to laboratories where trained staff are no longer employed.

Approved Weighing Designs

The OWM recognizes a variety of weighing designs such as those found in NBS Technical Notes 844 and 952, and the NIST/SEMATECH e-Handbook. However, weighing designs are used throughout the world with variations from those presented in NIST publications. Metrologists should use good judgment in developing unusual weighing designs and may submit them to NIST for review or validation whenever appropriate. Any designs submitted to NIST for review should be accompanied by enough experimental data to provide adequate evaluation. Metrologists should consider, and be able to justify, variations in weighing designs, the selection and use of check standards, length of designs and time restraints particularly with respect to drift, selection of balance, use of sensitivity weights, etc., to suit particular calibration applications. With the use of electronic mass comparators, length of time during a design and fatigue of the operator (which affect design selection) are of less concern than with the older mechanical balances. When developing new designs, another consideration should be that weights of equal nominal values should have the same uncertainty. NIST/OWM strongly recommends designs that incorporate check standards for process and standard verification.

Any unusual weighing designs not submitted to NIST for review will be subject to critical review during on-site assessments and may not be used in support of traceability hierarchies without suitable outside review, validation, and approval. When new designs are submitted to OWM for review, they will also be submitted for review and validation of associated K_n factors by the Statistical Engineering Division before OWM approval is granted.

Documentation of Standard Operating Procedures

SOP No. 28 “Recommended Standard Operating Procedures for Using Advanced Weighing Designs” was developed to help with laboratory compliance to NIST Handbook 143, NVLAP Handbook 150 and ISO/IEC 17025:2017. SOP 28 is an appendix to this document.

Metrological Traceability and Calibration Intervals

Ensuring metrological traceability to the SI and providing documentation of how metrological traceability is maintained is a critical concern for customers and for accreditation bodies when evaluating a laboratory. Metrological traceability for mass measurements can be maintained through two reference kilograms that have been calibrated at NIST provided appropriate mass calibration and measurement assurance procedures are used (and documented). The process described in this document provides guidance for laboratories on how to maintain adequate metrological traceability with sufficiently small uncertainties to meet the laboratory and customer requirements. See NISTIR 6969, GMP 11 and GMP 13 for additional examples of traceability hierarchies and calibration interval requirements. Each laboratory must have an implementation policy that covers the requirements in GMP 11 and GMP 13 to ensure metrological traceability with appropriate uncertainties and calibration intervals. Both GMPs are posted on the NIST website.

The assigned LAP problems can be used to initiate an appropriate measurement assurance program and prepare graphs. The data from the analysis and preparation of the graphs must be evaluated against original NIST values for the reference/working mass standards and check standards as appropriate. Additional data collected periodically is added to the original graphs. Data should be updated periodically and evaluated. How often data are updated depends on the laboratory workload.

Each laboratory must document which standards are used at each level in their traceability hierarchy process, what specific measurement assurance is in place at each step, and how often intercomparisons are conducted. The measurement assurance program, as described, is fully integrated into the actual calibration process. Therefore, this is not an exercise used to provide data for an accreditation body but provides checks on the system that the laboratory can use to ensure that each measurement performed for a customer is accurate and traceable with validated measurement results and uncertainties.

It is critical for laboratories to participate in interlaboratory comparisons that provide periodic checks on the measurement process. The data must be correlated with the measurement assurance program to be meaningful. Any discrepancies indicate the need for further investigation and possible need for calibration of the reference standards. Laboratories must participate in this level of intercomparison on a frequency no greater than four years (per OWM policy published in NISTIR 7082, Proficiency Test Policy and Plan (for State Weights & Measures Laboratories.)

Formulae and Calculations

The following items are calculated using formulae located in SOP No. 28 "Recommended Standard Operating Procedure for Using Advanced Weighing Designs" and in the NISTIR 5672 Supplementary Equations spreadsheet. SOP 5 and SOP 28 are included as appendices to this publication and the supplementary spreadsheets are available on the NIST OWM website.

- s_r – The standard uncertainty of the starting restraint in the first series.
- s_w – The within-process standard deviation.
- s_b – The between-time standard deviation for each series.
- Effective densities for summation standards.
- Effective cubical coefficients of expansion for summation standards.

SOP 5

Recommended Standard Operating Procedure for Using a 3-1 Weighing Design

1 Introduction

1.1 Purpose

The 3-1 weighing design is a combination of three double substitution comparisons of three weights of equal nominal value; a standard, an unknown weight, and a second standard called a check standard. (The check standard may be made up of a summation of weights.) The weights are compared using an equal-arm, single-pan mechanical, full electronic, or a combination balance utilizing built-in weights and a digital indication. The specific SOP for the double substitution procedure for each balance is to be followed. The 3-1 weighing design provides two methods of checking the validity of the measurement using an integrated F-test to monitor the within process standard deviation and a t-test to evaluate the stability of the standard and check standard. Hence, the procedure is especially useful for high accuracy calibrations for weights in OIML[1] Classes E_2 to F_2 and ASTM[2] Classes 0 and 1, in which it is critical to assure that the measurements are valid and well documented. This procedure is recommended as a minimum for precision calibration of laboratory working standards that are subsequently used for lower level calibrations and for routine calibration of precision mass standards used for balance calibration. For surveillance of reference standards, such as OIML E_1, ASTM 00 or better, working standards, and for calibration of precision mass standards used to calibrate other mass standards, see SOP 28 for the use of higher level weighing designs.

1.2 Prerequisites

1.2.1 Verify that valid calibration certificates with appropriate values and uncertainties are available for all reference standards used in the calibration. All standards must have demonstrated metrological traceability to the international system of units (SI), which may be through a National Metrology Institute such as NIST.

1.2.2 Standards must be evaluated to ensure that standard uncertainties for the intended level of calibration are sufficiently small. Reference standards should only be used to calibrate the next lower level of working standards in the laboratory and should not be used to routinely calibrate customer standards.

[1] OIML is the International Organization for Legal Metrology. Weight classes are published in OIML R111, which is freely available at http://www.oiml.org.

[2] ASTM International (formerly the American Society for Testing and Materials) publishes the E617 standard for mass specifications and tolerances.

1.2.3 The balance that is used must be in good operating condition with sufficiently small process standard deviation as verified by F-test values, pooled short term standard deviations, and by a valid control chart for check standards or preliminary experiments to ascertain its performance quality when new balances are put into service.

1.2.4 The operator must be experienced in precision weighing techniques. The operator must have specific training in SOP 2, SOP 4, SOP 5, SOP 29, and be familiar with the concepts in GMP 10. (These procedures are published in NISTIR 6969, Selected Laboratory and Measurement Practices, and Procedures to Support Basic Mass Calibrations.)

1.2.5 Verify that the laboratory facilities comply with the following minimum conditions to meet the expected uncertainty possible with this procedure and to comply with the balance manufacturer's operating conditions specified for the balance. Equilibration of balances and weights requires environmental stability of the laboratory within the stated limits for a minimum of 24 hours before a calibration.

Table 1. Environmental conditions.

Echelon[3]	Temperature Requirements During a Calibration	Relative Humidity (%)
I	OIML E_1, ASTM 000, 00, 0 Lower and upper limits: 18 °C to 23 °C Maximum changes: ± 0.5 °C / 12 h and ± 0.3 °C / h	40 to 60 ± 5 / 4 h
	OIML E_2, ASTM 1 Lower and upper limits: 18 °C to 23 °C Maximum changes: ± 1 °C / 12 h and ± 0.7 °C / h	
II	Lower and upper limits: 18 °C to 23 °C Maximum changes: ± 2 °C / 12 h and ± 1.5 °C / h	40 to 60 ± 10 / 4 h

It is important that the difference in temperature between the weights and the air inside the mass comparator is as small as possible. Keeping the reference weight and the test weight inside, or next to, the mass comparator before and during the calibration can reduce this temperature difference. Standards and test artifacts must be allowed to reach equilibration in or near the balance before starting measurements.

[3] Echelon I corresponds to weights of Classes OIML E_1 and E_2, Echelon II corresponds to weights of Classes OIML F_1 and F_2.

2 Methodology

2.1 Scope, Precision, Accuracy

This method can be performed on any type of balance or mass comparator using the appropriate double substitution method for the weighing instrument. Because considerable effort is involved, this weighing design is most useful for calibrations of the highest accuracy. The weighing design uses three double substitutions to calibrate a single unknown weight. This introduces redundancy into the measurement process and permits two checks on the validity of the measurement; one on accuracy and stability of the standard and the other on process repeatability. A least-squares best fit analysis is done on the measurement outputs to assign a value to the unknown weight. The standard deviation of the process depends upon the resolution of the balance and the care exercised to make the required weighings. The accuracy will depend upon the accuracy and uncertainty of the calibration of the standard weights and the precision of the comparison.

2.2 Summary

A standard weight, S, an unknown weight, X, and a check standard, S_c, are intercompared in a specific order using several double substitution processes. The balance and the weights must be prepared according to the appropriate double substitution method for the balance being used. Once the balance and weights have been prepared, all readings must be taken from the reading scale of the balance without adjusting the balance or weights in any way during the process. Within a double substitution all weighings are made at regularly spaced time intervals to minimize effects due to instrument drift. The 3-1 weighing design includes air buoyancy corrections.

2.3 Apparatus/Equipment Required

2.3.1 Precision analytical balance or mass comparator with sufficient capacity and resolution for the planned calibrations.

2.3.2 Calibrated reference standards or working standards, of nominally equal mass to the unknown mass standards being calibrated. Calibrated tare weights are used as needed to ensure that the standard(s) and test artifacts are of equal nominal mass (See NISTIR 6969, SOP 34 for suitable limits).

2.3.3 Calibrated sensitivity weights and tare weights selected to comply with the requirements of SOP 34.

2.3.4 Uncalibrated weights to be used to adjust the balance to the desired reading range if needed.

2.3.5 Forceps to handle the weights or gloves to be worn if the weights are moved by hand.

2.3.6 Forceps and gloves must be selected to avoid damage or contamination to mass standards.

2.3.7 Stop watch or other timing device to observe the time of each measurement (calibration not required; this is used to ensure consistent timing of the measurement). If an electronic balance is used that has a means for indicating a stable reading, the operator may continue to time readings to ensure consistent timing that can minimize errors due to linear drift.

2.3.8 Calibrated barometer with sufficiently small resolution, stability, and uncertainty (See NISTIR 6969, SOP 2, e.g., accurate to ± 66.5 Pa (0.5 mmHg)) to determine barometric pressure.[4]

2.3.9 Calibrated thermometer with sufficiently small resolution, stability, and uncertainty (see SOP 2, e.g., accurate to ± 0.10 °C) to determine air temperature.[4]

2.3.10 Calibrated hygrometer with sufficiently small resolution, stability, and uncertainty (see SOP 2, e.g., accurate to ± 10 percent) to determine relative humidity.[4]

[4]NISTIR 6969 includes SOP 2 for the calculation of air density. The barometer, thermometer, and hygrometer are used to determine the air density at the time of the measurement. The air density is used to make an air buoyancy correction. The limits specified are recommended for high precision calibration.

2.4 Symbols

Table 2. Symbols used in this procedure.

Symbol	Description
S	standard reference weight
X	weight to be calibrated
S_c	check standard
t	small calibrated tare weight, A subscript s or x is used to indicate the larger weight with which it is associated
sw	small calibrated weight used to evaluate the sensitivity of the balance
M	the mass (true mass) of a specific weight. Subscripts s, x, t, sw are used to identify the weight (equals Nominal plus Correction)
N	the nominal value of a specific weight. Subscripts s, x, are used to identify the weight
C	the correction for a specific weight. Subscripts s, x, are used to identify the weight
CM	the conventional mass of a specific weight. Subscripts s, x, t, sw are used to identify the weight
ρ_a	density of air at time of calibration
ρ_n	density of normal air (1.2 kg/m^3)
ρ	density of masses; subscripts s, x, t_s, t_x, sw are used to identify the weight

2.5 Procedure

2.5.1 Preliminary Procedure

2.5.1.1 Weights are visually inspected for cleanliness and damage. Follow the laboratory policy and customer contract review process to determine if and when standards will be cleaned, or standards with inadequate cleanliness are returned without calibration, and when "as found" and "as left" values will be obtained through duplicate calibrations.

2.5.1.2 If cleaning weights, it is important to clean weights before any measurements are made, unless "as found' data is to be measured, because the cleaning process may change the mass of the weight. Cleaning should not remove any significant amounts of weight material. Weights should be handled and stored in such a way that they stay clean. Before calibration, dust and any foreign particles shall be removed by blowing air across the surface or by brushing with a clean soft bristled brush. Care must be taken not to change the surface properties of the weight (i.e., by scratching the weight). If a weight contains significant amounts of dirt that cannot be removed by the methods cited above, the weight or some part of it can be washed with clean alcohol, distilled water, or other solvents. Weights with internal cavities should normally not be immersed in the solvent to avoid the possibility that the fluid will penetrate the opening. If there is a need to monitor the stability of a weight in use, the mass of the weight should, if possible, be determined before cleaning.

2.5.1.3 If weights are cleaned with solvents, they must be stabilized for the times given in the following table:

Table 3. Cleaning stabilization time.

Weight class	**E_1**	**E_2**	**F_1**	**F_2 to M_3**
After cleaning with alcohol	7 to 10 days	3 to 6 days	1 to 2 days	1 hour
After cleaning with distilled water	4 to 6 days	1 to 3 days	1 day	1 hour

2.5.1.4 Prior to performing any calibration tests, the weights need to be equilibrated to the ambient conditions of the laboratory. Specifically, weights of classes F_1 (or better) should be close to the temperature in the weighing area. The minimum times (in hours) required for temperature stabilization (depending on weight size, weight class and on the difference between the initial temperature of the weights and the room temperature in the laboratory) are shown in the table below (with appropriate documented evidence). As a practical guideline, a waiting time of 24 hours is recommended. If weights are extremely hot or cold, additional equilibration is required to address problems with varying surface moisture. Weights must be completely dry prior to calibration.

Table 4. Minimum equilibration times.[5]

ΔT[a]	Nominal Mass[6]	OIML Class E_1 (time in h)	OIML Class E_2 (time in h)	OIML Class F_1 (time in h)	OIML Class F_2 to M_3 (time in h)
± 20 °C	1 000, 2 000, 5 000 kg	-	-	79	5
	100, 200, 500 kg	-	70	33	4
	10, 20, 50 kg	45	27	12	3
	1, 2, 5 kg	18	12	6	2
	100, 200, 500 g	8	5	3	1
	10, 20, 50 g	2	2	1	1
	< 10 g	1	1	1	0.5
± 5 °C	1 000, 2 000, 5 000 kg	-	-	24	1
	100, 200, 500 kg	-	40	10	1
	10, 20, 50 kg	36	18	4	1
	1, 2, 5 kg	15	8	3	1
	100, 200, 500 g	6	4	2	0.5
	10, 20, 50 g	2	1	1	0.5
	< 10 g	0.5	0.5	0.5	0.5
± 2 °C	100 kg to 5 000 kg	-	16	1	0.5
	10, 20, 50 kg	27	10	1	0.5
	1, 2, 5 kg	12	5	1	0.5
	100, 200, 500 g	5	3	1	0.5
	< 100 g	2	1	1	0.5
± 0.5 °C	100 kg to 5 000 kg	-	1	0.5	0.5
	10, 20, 50 kg	11	1	0.5	0.5
	1, 2, 5, kg	7	1	0.5	0.5
	100, 200, 500 g	3	1	0.5	0.5
	< 100 g	1	0.5	0.5	0.5

[a]ΔT = Initial difference between weight temperature and laboratory temperature.

2.5.2 Weighing Design Matrix

The following table shows the intercomparisons to be made in the 3-1 design, in a matrix format as shown in NBS Technical Note 952, Designs for the Calibration of Standards of Mass, J. M. Cameron, M. C. Croarkin, and R. C. Raybold, 1977:

[5] Consider equivalent ASTM Classes for equilibration times.
[6] Nominal masses in the 1, 2, 5 combinations include intermediate values such as 3.

Table 5. Weighing Design Matrix.

Weight ID Comparison	***S***	***X***	S_c
a_1	+	-	
a_2	+		-
a_3		+	-
Standard (Restraint)	+		
Check Standard			+

This design is represented as design ID "A.1.1"in Technical Note 952, with the exception that the design order is reversed and Restraint B is used[7]. The restraint is another name for the "standard" used in the comparison that may be found in NBS Technical Note 952. The following table may be useful for data reduction purposes if using the NIST Mass Code. When creating a data file for this design, the restraint, check, and design vectors will appear in the order and as the numbers in the following table:

Table 6. Weighing Design Vector for Mass Code.

Row in Mass Code Design	First Entry	Second Entry	Third Entry
Restraint	1	0	0
Check	0	0	1
Following series sum	0	0	0
Report	0	1	1
1st double sub	1	-1	0
2nd double sub	1	0	-1
3rd double sub	0	1	-1

2.5.3 Measurement Procedure

Perform the measurement as a series of 3 double substitutions. Review SOP 4 (NISTIR 6969) if needed. Record the pertinent information for the standard, S, unknown, X, and check standard, S_c, as indicated on a suitable data sheet such as the one in the Appendix of this SOP. Measure the laboratory ambient temperature, barometric pressure, and relative humidity, before and after the set of three double substitutions (i.e., before and after all twelve observations). If performing buoyancy corrections on *each* double substitution within the design, record environmental data before observations 1, 5, 9 and after measurement 12. Perform the mass measurements in the order shown in the following table.

[7] Additional 3-1 weighing designs are published. This procedure provides the calculations and solutions for the position of mass standards/artifacts in this design only.

Table 7. 12 Measurement Observations.

Double Substitution	Measurement Number	Weights on Pan	Observation
a_1: *S vs X*	1	$S + t_s$	O_1
	2	$X + t_x$	O_2
	3	$X + t_x + sw$	O_3
	4	$S + t_s + sw$	O_4
a_2: *S vs* S_c	5	$S + t_s$	O_1
	6	$S_c + t_{sc}$	O_2
	7	$S_c + t_{sc} + sw$	O_3
	8	$S + t_s + sw$	O_4
a_3: *X vs* S_c	9	$X + t_x$	O_1
	10	$S_c + t_{sc}$	O_2
	11	$S_c + t_{sc} + sw$	O_3
	12	$X + t_x + sw$	O_4

3 Calculations

Note: As the NIST Mass Code software can be used to perform the calculations for this process, the balance of this section is provided as reference to be used for manual or computerized calculations using other software packages.

3.1 Calculate the average air density, ρ_A, as described in the Appendix to NISTIR 6969, SOP No. 2, Option B.

3.2 Calculate the measured differences, a_1, a_2, and a_3, for the weights used in each double substitution using the following formula (note: do not confuse this formula with the calculations used in SOP 4, NISTIR 6969; the signs will be opposite SOP 4, Option A.):

$$a_x = \frac{(O_1 - O_2 + O_4 - O_3)}{2} \frac{M_{sw}\left(1 - \frac{\rho_A}{\rho_{sw}}\right)}{O_3 - O_2} \qquad \text{Eqn. (1)}$$

3.3 Calculate the within process standard deviation, s_w, for the 3-1 weighing design. This standard deviation has one degree of freedom.

$$s_w = 0.577\left(a_1 - a_2 + a_3\right) \qquad \text{Eqn. (2)}$$

3.4 Calculate the F statistic which compares the observed within process standard deviation, s_w, to the pooled (accepted) within process standard deviation. (See NISTIR 6969, Sections 8.4, 8.5, and 8.9.2, for a discussion of the statistics used in weighing designs.)

$$F = \frac{s^2_{w\,\mathrm{Observed}}}{s^2_{w\,\mathrm{Accepted}}} \qquad \text{Eqn. (3)}$$

The calculated F-statistic must be less than the F-value obtained from an F-table at 95 % confidence level (Table 9.12, NISTIR 6969) to be acceptable. The F-value is obtained from the F-table for numerator degrees of freedom equal one, and denominator degrees of freedom equal to the number of degrees of freedom in the pooled within process standard deviation. If the data fails the F-test and the source of the error cannot be determined conclusively, the measurement must be repeated.

3.5 Calculate the least-squares measured difference d_{sc} for S_c.

$$d_{Sc} = \frac{-a_1 - 2a_2 - a_3}{3} \qquad \text{Eqn. (4)}$$

3.6 Calculate the observed mass of S_c, M_{sc}.

$$M_{Sc} = \frac{M_s\left(1 - \frac{\rho_A}{\rho_s}\right) + d_{Sc} + M_{t_s}\left(1 - \frac{\rho_A}{\rho_{t_s}}\right) - M_{t_{sc}}\left(1 - \frac{\rho_A}{\rho_{t_{Sc}}}\right)}{\left(1 - \frac{\rho_A}{\rho_{Sc}}\right)} \qquad \text{Eqn. (5)}$$

3.7 Calculate the Conventional Mass of S_c, CM_{Sc}:

$$CM_{Sc} = \frac{M_{Sc}\left(1 - \frac{0.0012}{\rho_{Sc}}\right)}{0.999850} \qquad \text{Eqn. (6)}$$

3.8 Evaluate the mass $M_{Sc,}$ or conventional mass CM_{Sc}, of S_c.

The observed mass or conventional mass is evaluated with a t-test in the procedure based on comparison with the accepted mass or conventional mass value determined from the mean of the control chart. The mass or conventional mass (depending on what is tracked in the laboratory) is also plotted on the control chart and must lie within the control limits (See NISTIR 6969, SOP 9). If the value is not within limits, and the source of error cannot be found, measurement must be stopped until suitable corrective action is taken. Corrective action is demonstrated through evaluation of additional measurement results that are within limits.

3.9 Calculate the least-squares measured difference, d_x, for X.

$$d_x = \frac{-2a_1 - a_2 + a_3}{3} \qquad \text{Eqn. (7)}$$

3.10 Calculate the mass of X, M_x.

$$M_x = \frac{M_s\left(1-\frac{\rho_A}{\rho_s}\right) + d_x + M_{t_s}\left(1-\frac{\rho_A}{\rho_{t_s}}\right) - M_{t_x}\left(1-\frac{\rho_A}{\rho_{t_x}}\right)}{\left(1-\frac{\rho_A}{\rho_x}\right)} \qquad \text{Eqn. (8)}$$

3.11 Calculate the conventional mass[8] of X, CM_x. The conventional mass should be reported.

$$CM_x = \frac{M_x\left(1-\frac{0.0012}{\rho_x}\right)}{0.999850} \qquad \text{Eqn. (9)}$$

3.12 Calculate apparent mass versus brass only if requested. This value should only be provided when requested by the customer for use when calibrating mechanical balances that have been adjusted to this reference density. (This is rare.) Apparent mass versus brass (8.3909 g/cm^3 at 20 °C)

$$AM_x\, vs. 8.4 = \frac{M_x\left(1-\frac{0.0012}{\rho_x}\right)}{\left(1-\frac{0.0012}{8.3909}\right)} \qquad \text{Eqn. (10)}$$

4 Measurement Assurance

4.1 The within process standard deviation is incorporated into the NIST Mass Code and is used to conduct an F-test of the observed standard deviation versus the pooled/accepted standard deviation of the process at a 95 % confidence level. If calculations are performed manually or using other software, follow the process described in section 3 for F-test evaluation.

[8] Conventional Mass: "The conventional value of the result of weighing a body in air is equal to the mass of a standard, of conventionally chosen density, at a conventionally chosen temperature, which balances this body at this reference temperature in air of conventionally chosen density." The conventions are: artifact reference density 8.0 g/cm^3; reference temperature 20 °C; *normal* air density 0.0012 g/cm^3. Conventional mass was formerly called "Apparent Mass versus 8.0 g/cm^3" in the United States. (*See OIML D28 (2004)).*

4.2 SOP 5 weighing design integrates a suitable check standard (See NISTIR 6969, GLP 1, SOP 9, and SOP 30).

4.3 The check standard value is calculated and immediately evaluated on the control chart to verify that the mass is within established limits. A t-test may be incorporated to check the observed value of the check standard against the accepted value. It is evaluated using the following equation and a 95 % confidence level. All values must be entered in the control chart, even if failing this statistic to ensure the variability obtained for the process is truly representative of the process and not unduly reduced over time. The observed value of the check standard is compared to the accepted mean value of the check standard and divided by the standard deviation for the check standard observations over time. This equation monitors stability over time but should not be used to assess for bias. A calculated t-value less than two is within the warning limits of the process. A calculated t-value between two and three represents a value between the warning limits and control/action limits. A calculated t-value exceeding three represents a value outside of the control/action limits and suitable action must be taken. Calculated values of the t-statistic may also be monitored over time to determine the presence of drift.

$$t = \frac{\left(S_c - \overline{S_c}\right)}{s_p} \qquad \text{Eqn. (11)}$$

4.4 Check standard measurement results obtained over time are used to calculate the standard deviation of the measurement process, s_p.

4.5 The mean value of the check standard over time is also compared to an appropriate reference value of the check standard with respect to applicable expanded uncertainties to evaluate bias and drift over time. Excessive drift or bias must be investigated and followed with suitable corrective action. (See NISTIR 6969, SOP 9, Section 4.2 for assessment methodology.)

5 Assignment of Uncertainty

The limits of expanded uncertainty, U, include estimates of the standard uncertainty of the mass standards used, u_s, estimates of the standard deviation of the measurement process, s_p, and estimates of the effect of other components associated with this procedure, u_o. These estimates should be combined using the root-sum-squared method (RSS), and the expanded uncertainty, U, reported with a coverage factor to be determined based on the degrees of freedom, which if large enough will be 2, (k = 2), to give an approximate 95 percent level of confidence. (See NISTIR 6969, SOP 29, "Standard Operating Procedures for the Assignment of Uncertainty", for the complete standard operating procedure for calculating the uncertainty.

When the 3-1 weighing design is used in conjunction with specialized software for data reduction, see SOP 28, "Recommended Standard Operating Procedure for Using Advanced

Weighing Designs", for detailed instructions on calculating the uncertainty components which are required.

5.1 The expanded uncertainty for the standard, U, is obtained from the calibration certificate. The combined standard uncertainty, u_c, is used and not the expanded uncertainty, U, therefore the reported uncertainty for the standard will usually need to be divided by the coverage factor k. When multiple standards are used, see SOP 29 for evaluation of dependent and independent conditions and combining methods for the standard uncertainty of the standard. Usually only one standard is used as the restraint for the 3-1 weighing design, the uncertainty of the check standard is not included in assigning an uncertainty to the unknown mass. Where the coverage factor or confidence interval is not given, the laboratory should either contact the calibration provider to obtain the correct divisor or use a value of $k = 2$, assuming that the expanded uncertainty was reported with an approximate 95 % confidence interval (95.45 %).

5.2 The value for the standard deviation of the measurement process s_p is obtained from the control chart data for check standards using only 3-1 weighing design measurements (see SOP No. 9.) The within-process standard deviation, s_w, is only used as a part of the process variability evaluation using the F-test unless between time components are also determined. In that case, the standard deviation of the process, s_p, is treated as s_t (standard deviation over time) and see SOP 28 for details. Statistical control must be verified by the measurement of the check standard in the 3-1 design.

5.2.1 Where the standard deviation of the measurement process from the control chart is less than the resolution of the balance being used, the laboratory may round up to the nearest balance division to represent the standard deviation, or use the larger of the standard deviation of the process and one of the following estimates for repeatability is used to represent the standard deviation of the process:

5.2.1.1 If the laboratory prefers a conservative approach, or when the current and representative degrees of freedom are less than 30, the larger of the control chart s_p and the result from Eqn. 12 should be used, where d is the smallest balance division. For example, if the balance division is 0.1 mg, the smallest standard deviation may be 0.06 mg. If the laboratory calculated standard deviation is 0.075 mg, then 0.075 mg is used.

$$s_p = \frac{d}{\sqrt{3}} \approx 0.6d \qquad \text{Eqn. (12)}$$

5.2.1.2 When the laboratory has the confidence associated with a well characterized measurement process and has 30 or more degrees of freedom to represent the process, the larger of the observed s_p or Eqn. 13 is used. For example, if the balance division is 0.1 mg, the smallest standard deviation may be 0.03 mg. If the laboratory observed standard deviation is larger than 0.03 mg, then that is the value to be used.

$$s_p = \frac{d}{2\sqrt{3}} \approx 0.3d \qquad \text{Eqn. (13)}$$

5.3 Uncertainty due to air buoyancy corrections and air density. Select one of the following options in priority preference for calculating the uncertainty associated with air buoyancy.

5.3.1 Option 1, preferred. Use the formulae provided in OIML R111, C.6.3-1, C.6.3-2, and C.6.3-3.

5.3.2 Option 2. Calculate the uncertainty by quantifying estimated impacts associated with the uncertainties of the air temperature, barometric pressure, relative humidity, and the air density formula based on laboratory uncertainties and calculations given in NISTIR 6969, SOP 2 and the SOP being used. Note: this may be done using a simplified baseline “what if” approach or a Kragten analysis.[9]

5.4 Uncertainty associated with the density of the standards and the unknown test weights, u_{ρ}. Uncertainties associated with the density of the standards used in the calibration may be incorporated into the estimated calculations in section 5.3.

5.5 Uncertainty associated with bias, u_d. Any noted bias that has been determined through analysis of control charts and round robin data must be less than limits provided in SOP 29 and may be included if corrective action is not taken. (See NISTIR 6969, SOP 29 for additional details.)

5.6 Example components to be considered for an uncertainty budget table are shown in the following table.

[9] A baseline “what if” approach calculates the estimated impact of each variable in the final measurement result by individually changing each variable of interest by the uncertainty quantity. (See the EURACHEM/CITAC Quantitative Guide to Uncertainties in Analytical Methods (QUAM, 2012) for a discussion of Kragten spreadsheets).

Table 8. Example Uncertainty Budget Table.

Uncertainty Component Description	Symbol	Source	Typical Distribution
Uncertainty of the standard mass(es) (5.1)	u_s	Calibration certificate	Expanded U divided by coverage factor
Accepted standard deviation of the process (5.2, 5.2.1)	s_p	Control chart, standard deviation chart OR estimates when s_p is smaller than balance division	Normal OR Rectangular, one-half rectangular
Uncertainty of the air buoyancy correction (5.3)	u_b	OIML R111	Rectangular
Air temperature (for air density)	u_t	SOP 2 or OIML R111	Rectangular
Air pressure (for air density)	u_p	SOP 2 or OIML R111	Rectangular
Air relative humidity (for air density)	u_{RH}	SOP 2 or OIML R111	Rectangular
Air density (formula)	$u_{\rho a}$	SOP 2 or OIML R111	Rectangular
Mass densities (5.4)	$u_{\rho m}$	Measured and reported value OIML R111 Table B.7 Typically, 0.03 g/cm^3 to 0.05 g/cm^3	Rectangular
Uncertainty associated with bias (5.5)	u_d	Control chart, proficiency tests	NISTIR 6969 SOP 29

5.7 Draft a suitable uncertainty statement for the certificate (e.g.,)

The uncertainty reported is the root sum square of the standard uncertainty of the standard, the standard deviation of the process, and the uncertainty associated with the buoyancy corrections, multiplied by a coverage factor of 2 (k = 2) for an approximate 95 percent confidence interval. Factors not considered in the evaluation: magnetism (weights are considered to meet magnetism specifications unless measurement aberrations are noted), balance eccentricity and linearity (these factors are considered as a part of the measurement process when obtaining the standard deviation of the process when using a check standard with adequate degrees of freedom.

NOTE: Where inadequate degrees of freedom are available, k, is determined using the appropriate degrees of freedom and the 95.45 % column in the table from Appendix A of NISTIR 6969, SOP 29.

6 Certificate

6.1 Report results as described in NISTIR 6969, SOP No. 1, Preparation of Calibration Certificates. Report the mass, conventional mass, environmental conditions during the calibrations, mass density used (reported, measured, or assumed (specify which applies)), and calculated expanded uncertainties with coverage factor(s).

6.2 Conformity assessments.

Evaluate compliance to applicable tolerances as needed or required by the customer or by legal metrology requirements. Decision criteria for uncertainty and tolerance evaluations include two components: 1) the expanded uncertainty, U, must be < 1/3 of the applicable tolerances published in ASTM E 617 and OIML R111 documentary standards and 2) the absolute value of the conventional mass correction value plus the expanded uncertainty must be less than the applicable tolerance to confidently state that mass standards are in or out of tolerance. Compliance assessments must note the applicable documentary standard and which portions of the standard were or were not evaluated.

Appendix A

Observation Sheet for

3-1 Weighing Design When Tare Weights Are Used

Laboratory data and conditions:

			Before	After
Operator				
Date		Temperature °C		
Balance		Barometric Pressure mmHg		
Load		Relative Humidity %		
Pooled within process s.d., s_w=		Degrees of freedom for process s.d.		
Check standard s.d., s_p =		Degrees of freedom from control chart		

Mass standard(s) data:

ID (Insert Set or SN)	**Nominal**	**Mass Correction**	**Expanded Unc: from cal. certificate**	**Unc: *k* factor**	**Density g/cm^3**	**Unc Density *k*=1**
S						
t_s						
X						
t_x						
S_c						
t_{Sc}						
sw						

Laboratory observations:

Balance Observations, Units_______					
Time:					
$S - X = a_1$		$S - S_c = a_2$		$X - S_c = a_3$	
$S + t_s$		$S + t_s$		$X + t_x$	
$X + t_x$		$S_c + t_{sc}$		$S_c + t_{sc}$	
$X + t_x + sw$		$S_c + t_{sc} + sw$		$S_c + t_{sc} + sw$	
$S + t_s + sw$		$S + t_s + sw$		$X + t_x + sw$	
Time:					
a_1 =		a_2 =		a_3 =	

Appendix B – Flow Chart of the 3-1 Weighing Design Process

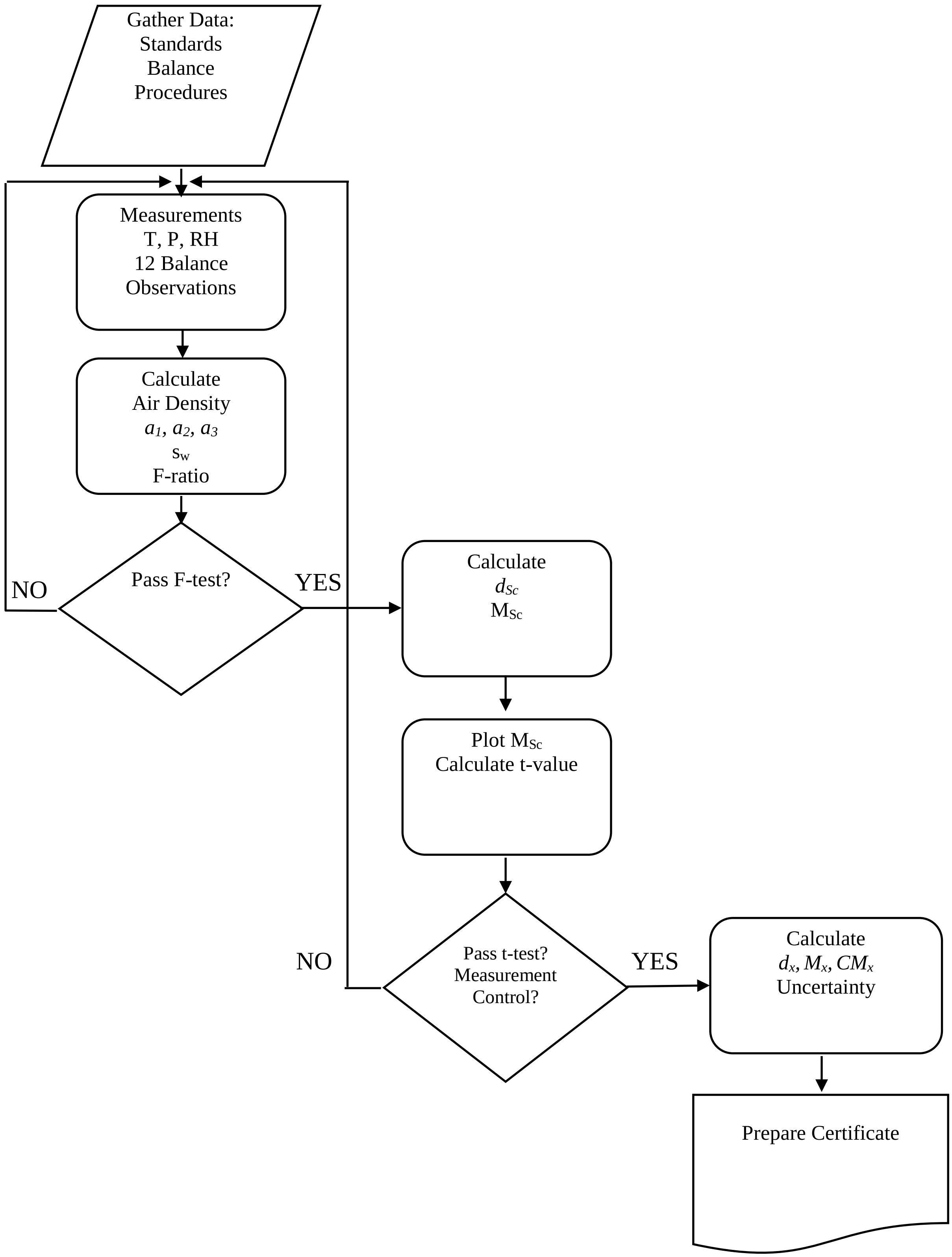

Figure 1. Appendix B – Flow Chart of the 3-1 Weighing Design Process.

SOP No. 28

Recommended Standard Operating Procedure for Using Advanced Weighing Designs

1 Introduction

1.1 Purpose

Advanced weighing designs use a combination of double substitution comparisons of weights of equal nominal value or a series of weights in ascending or descending order; standard(s), unknown weights, and an additional standard called a check standard. The weights are intercompared using an equal-arm, single-pan mechanical, full electronic, or a combination balance utilizing built-in weights and a digital indication. The specific SOP for the double substitution procedure for each balance is to be followed. Weighing designs provide two methods of checking the validity of the measurement using an F-test to check the measurement process and a t-test to evaluate the stability of the standard and check standard. Hence, the procedure is especially useful for high accuracy calibrations in which it is critical to assure that the measurements are valid and well documented. This procedure is recommended for precision calibration of laboratory working standards that are subsequently used for lower level calibrations, for routine calibration of precision mass standards used for calibration of other mass standards, and for surveillance of mass reference and working standards.

1.2 Prerequisites

1.2.1 Calibrated mass standards, with recent calibration values and which have demonstrated metrological traceability to the international system of units (SI), which may be to the SI through a National Metrology Institute such as NIST.

1.2.2 Standards must be evaluated to ensure that standard uncertainties for the intended level of calibration are sufficiently small. Reference standards should only be used to calibrate the next lower level of working standards in the laboratory and should not be used to routinely calibrate customer standards.

1.2.3 Verify that the balance used is in good operating condition with sufficiently small process standard deviation as verified by F-test values, pooled short-term standard deviations, and by a valid control chart for check standards, or preliminary experiments to ascertain its performance quality when new balances are put into service. See NISTIR 5672[1] for a discussion on the performance levels expected for use of these procedures as part of a

[1]Fraley, K. L., Harris, Georgia G. L., NISTIR 5672, Advanced Mass Calibration and Measurement Assurance Program for State Calibration Laboratories.

laboratory measurement assurance program to ensure traceability of laboratory standards.

1.2.4 Verify that the operator is experienced in precision weighing techniques, and has had specific training in SOP 2, SOP 4, SOP 5, SOP 29, and is familiar with the concepts in GMP 10. Further, the operator must have been trained in the creation of data files and the operation of the NIST Mass Code when it is used for data reduction as recommended. Example data sets and sample observation sheets are available in the Advanced Mass Seminar offered by the NIST Office of Weights and Measures.

1.2.5 Laboratory facilities must comply with the following minimum conditions to meet the expected uncertainty possible with this procedure and to comply with the balance manufacturer's operating conditions specified for the balance. Equilibration of balances and weights requires environmental stability of the laboratory within the stated limits for a minimum of 24 hours before a calibration.

Table 1. Environmental conditions.

Echelon[2]	**Temperature Requirements During a Calibration**	**Relative Humidity (%)**
I	OIML E_1, ASTM 000, 00, 0 Lower and upper limits: 18 °C to 23 °C Maximum changes: ± 0.5 °C / 12 h and ± 0.3 °C / h	40 to 60 ± 5 / 4 h
	OIML E_2, ASTM 1 Lower and upper limits: 18 °C to 23 °C Maximum changes: ± 1 °C / 12 h and ± 0.7 °C / h	

1.2.5.1 It is important that the difference in temperature between the weights and the air inside the mass comparator is as small as possible. Keeping the reference weight and the test weight inside, or next to, the mass comparator before and during the calibration can reduce this temperature difference. Standards and test artifacts must be allowed to reach equilibration in or near the balance before starting measurements.

[2] Echelon I corresponds to weights of Classes OIML E_1 and E_2, Echelon II corresponds to weights of Classes OIML F_1 and F_2.

2 Methodology

2.1 Scope, Precision, Accuracy

This method can be performed on any type of balance or mass comparator using the appropriate double substitution method for the weighing instrument. Because considerable effort is involved, this weighing design is most useful for calibrations of the highest accuracy. The weighing design utilizes a combination of double substitutions to calibrate a single unknown weight, or a group of related weights in a decade. This method introduces redundancy into the measurement process and permits two checks on the validity of the measurement; one on accuracy and stability of the standard using an integrated t-test and the other on process repeatability using an F-test. A least-squares best fit analysis is done on the measurements to assign a value to the unknown weights. The standard deviation of the process depends upon the resolution of the balance and the care exercised to make the required weighings. The accuracy will depend upon the accuracy and uncertainty of the calibration of the restraint weights and the precision of the comparison.

2.2 Summary

A restraint weight, S, in some cases two restraint weights, S_1 and S_2, an unknown weight, X, or group of unknown weights, and a check standard, S_C are compared in a specific order typically using the double substitution procedure although other procedures may be appropriate. The balance and the weights must be prepared according to the appropriate method for the balance being used. Once the balance and weights have been prepared, all readings must be taken from the reading scale of the balance without adjusting the balance or weights in any way. Within a double substitution all weighings are made at regularly spaced time intervals to minimize effects due to instrument drift. Because of the amount of effort required to perform weighing designs, the procedure includes an air buoyancy correction using the average air density as determined immediately before and after the weighings, drift-free equation for calculating the observed differences, correction for the cubical coefficient of expansion when measurements are not made at 20 °C, an average sensitivity for the balance over the range of measurements made, and the international formula for air density.[3]

2.3 Apparatus/Equipment Required

2.3.1 Precision analytical balance or mass comparator with sufficient capacity and resolution for the calibrations planned.

[3] See NISTIR 6969, Selected Procedures for Mass Calibrations, SOP 2, Air Density Calculations. Use Option B. The difference between Option A and Option B in SOP 2 is less than the uncertainty associated with assumptions made in the air density equations.

2.3.1. Calibrated reference standards (usually starting at 1 kg or 100 g), calibrated check standards for each decade (e.g., 1 kg, 100 g, 10 g, 1 g, 100 mg, 10 mg, 1 mg for the seven series between 1 kg and 1 mg) and working standards.

2.3.2. Calibrated sensitivity weights and tare weights selected to comply with the requirements of SOP 34. Note: The calculations performed by the Mass Code do not take into consideration the value of any tare weights used in the weighing design. Additional calculations will be required when tare weights are used.

2.3.3. Uncalibrated weights to be used to adjust the balance to the desired reading range if needed.

2.3.4. Forceps to handle the weights, or gloves to be worn if the weights are moved by hand. Forceps and gloves must be selected to avoid damage or contamination of mass standards.

2.3.5. Stop watch or other timing device to observe the time of each measurement (calibration not required; this is used to ensure consistent timing of the measurement). If an electronic balance is used that has a means for indicating a stable reading, the operator may continue to time readings to ensure consistent timing that can minimize errors due to linear drift.

2.3.6. Calibrated barometer with sufficiently small resolution, stability, and uncertainty (See SOP 2, e.g., accurate to ± 66.5 Pa (0.5 mmHg)) to determine barometric pressure.[4]

2.3.7. Calibrated thermometer with sufficiently small resolution, stability, and uncertainty (see SOP 2, e.g., accurate to ± 0.10 °C) to determine air temperature.[4]

2.3.8. Calibrated hygrometer with sufficiently small resolution, stability, and uncertainty (see SOP 2, e.g., accurate to ± 10 percent) to determine relative humidity.[4]

2.4 Procedure

2.4.1 Preliminary Procedure

2.4.1.1 Weights are visually inspected for cleanliness and damage.

2.4.1.2 If cleaning weights, it is important to clean weights before any measurements are made because the cleaning process may change the mass of the weight. Cleaning should not remove any significant

[4] The barometer, thermometer, and hygrometer are used to determine the air density at the time of the measurement. The air density is used to make an air buoyancy correction. The limits specified are recommended for high precision calibration.

amounts of weight material. Weights should be handled and stored in such a way that they stay clean. Before calibration, dust and any foreign particles shall be removed. Care must be taken not to change the surface properties of the weight (i.e. by scratching the weight). If a weight contains significant amounts of dirt that cannot be removed by the methods cited above, the weight or some part of it can be washed with clean alcohol, distilled water, or other solvents. Weights with internal cavities should normally not be immersed in the solvent to avoid the possibility that the fluid will penetrate the opening. If there is a need to monitor the stability of a weight in use, the mass of the weight should, if possible, be determined before cleaning.

2.4.1.3 If weights are cleaned with solvents, they must be stabilized for at least 7 to 10 days.

2.4.1.4 Before making measurements, place the test weight and standards in the balance chamber or near the balance overnight to permit the weights and the balance to attain thermal equilibrium, or use a thermal soaking plate next to the balance with weights covered. Thermal equilibration time is particularly important with weights larger than 1 gram. An alternative heat source such as a heat lamp may further improve temperature stability in front of the balance. Conduct preliminary measurements to determine the size of the sensitivity weight and any tare weights that are required following SOP 34; adjust the balance to the appropriate reading range of the balance indications and exercise the balance. Refer to the appropriate double substitution method for details.

2.4.2 Weighing Designs

Table 2 shows the most common comparisons to be made as referenced in NBS Technical Note 952, Designs for the Calibration of Standards of Mass, J. M. Cameron, M. C. Croarkin, and R. C. Raybold, 1977. Each series is characterized by the number of observations, *n*, the degrees of freedom, *d.f.*, associated with the standard deviation, the number of weights in each design, k (not shown in this table), the number of restraints (standards), and check standards, along with appropriate positions within the design*.

*Positions for check standards must be carefully considered as subsequent equations may be dependent on the position of use.

Table 2. Common weighing designs.

Design ID	Description	*n* Observations	*d.f.* Degrees of freedom	Restraint Position	Check Std Position[a]
A.1.1	3-1 Weighing Design[5]	3	1	1	3
A.1.2	4-1 Weighing Design	6	3	1, 2	3 or 4
A.1.4	5-1 Weighing Design	10	6	1, 2	3, 4, or 5
A.2.1	6-1 Weighing Design	8	3	1, 2	3, 4, 5, or 6
C.1[a]	5, 3, 2, 1, 1 Design (descending)	8	4	1, 2, 3	5 or 4
C.1	5, 3, 2, 1, 1 Design (ascending)	8	4	5	4
C.2	5, 3, 2, 1, 1, 1 Design (descending)	11	6	1, 2, 3	4, 5, or 6
C.2	5, 3, 2, 1, 1, 1 Design (ascending)	11	6	6	4 or 5
C.9[a]	5, 2, 2, 1, 1 Design (descending)	8	4	1, 2, 3, 4	5
C.9	5, 2, 2, 1, 1 Design (ascending)	8	4	5	4
C.10	5, 2, 2, 1, 1, 1 Design (descending)	8	3	1, 2, 3, 4	5 or 6
C.10	5, 2, 2, 1, 1, 1 Design (ascending)	8	3	6	4 or 5

[a]If these designs are NOT the last in a series, there is no position for a check standard.

The “restraint” is another name for the standard used in the comparison. Matrices are shown in Technical Note 952. Determine the best design prior to beginning the series. The series shown allow calibration of any commonly found set of mass standards in the 5, 2, 2, 1 combination or the 5, 3, 2, 1 combination. See the supplemental job aid spreadsheet associated with NISTIR 5672 for many additional weighing designs that identify associated restraint and check standard positions, with the associated K_1 and K_2 values needed for calculating between-time standard deviations.

2.4.3 Measurement Procedure

Record the pertinent information for all weights being intercompared on a suitable data sheet unless an automated data collection system is being used to collect the data and create a data file. Record or collect the laboratory ambient temperature, barometric pressure, and relative humidity immediately before and immediately after each series of intercomparisons.

3 Calculations

Calculations are completed by the NIST Mass Code as described in NBS Technical Note 1127, National Bureau of Standards Mass Calibration Computer Software, R. N. Varner,

[5]Design a.1.1. with inverted order (y3, y2, and y1), with restraint in position 1 (B) is detailed in SOP 5.

and R. C. Raybold, July 1980, with updates to conform to the international formula for calculating air density and the ISO/IEC Guide to the Expression of Uncertainties as described in the NIST Technical Note 1297 and minor error corrections to the original code. The Mass Code performs two statistical tests (t-test and F-test) to verify both the value of the restraints and check standards, and to verify that the measurement process was in control during the comparisons.

3.1 Calculating Effective Densities and Coefficients of Expansion for Summations[6]:

Some designs use a summation mass and sometimes the individual masses of this summation will be constructed from different materials that have different densities and coefficients of expansion. The following equations will be used to calculate the effective density and effective coefficient of expansions for the summation that will be needed as input for the data file. The subscripts 5, 3, and 2 refer to the individual masses that comprise the summation. This approach may also be needed with a 5, 2, 2, 1 combination. E.g., a summation of 1 kg, might be 500 g, 300 g, and 200 g (or 500 g, 200 g, 200g, and 100 g). A summation for 100 g would be 50 g, 30 g, and 20 g (or the equivalent of a 5221 series). A metric calibration of a 4 oz weight would need to add up to 113.389 g (probably 100 g, 10 g, 3 g, 300 mg, 100 mg) and then must also address unequal nominal values as shown in the equations in SOP 4, 5, and 7 as applicable.

$$\text{Effective Density} = \frac{M_5 + M_3 + M_2}{\left(\frac{M_5}{\rho_5}\right) + \left(\frac{M_3}{\rho_3}\right) + \left(\frac{M_2}{\rho_2}\right)} \qquad \text{Eqn. (1)}$$

$$\text{Effective Cubical Coefficient of Expansion} = \frac{\left(\frac{M_5}{\rho_5}\alpha_5\right) + \left(\frac{M_3}{\rho_3}\alpha_3\right) + \left(\frac{M_2}{\rho_2}\alpha_2\right)}{\left(\frac{M_5}{\rho_5}\right) + \left(\frac{M_3}{\rho_3}\right) + \left(\frac{M_2}{\rho_2}\right)} \qquad \text{Eqn. (2)}$$

Table 3. Variables for Above Equations.

Variable	Description
M	Mass (g)
ρ	Density (g/cm^3)
α	Cubical Coefficient of Expansion (/°C)

4 Measurement Assurance

4.1. The within process standard deviation is incorporated into the NIST Mass Code and is used to conduct an F-test of the observed standard deviation versus the pooled/accepted standard deviation of the process at a 95 % confidence level. See NISTIR 6969, Section 8.9 for details of the F-test.

[6] Jaeger, K B., and R. S. Davis, NIST Special Publication 700-1, A Primer for Mass Metrology, November 1984.

4.2. Weighing designs integrate a suitable check standard (See GLP 1, SOP 9 and SOP 30).

4.3. The check standard value is calculated and immediately evaluated to verify that the mass is within established limits; a t-test is incorporated to compare the observed value against accepted value at a 95 % confidence level. This statistical test is incorporated in the NIST Mass Code. See NISTIR 6969, Section 8.16 for details and statistics that are used.

4.4. The mean value of the check standard over time is also compared to an appropriate reference value of the check standard with respect to their applicable expanded uncertainties to evaluate bias and drift over time. Excessive drift or bias must be investigated and followed with suitable corrective action. See NISTIR 6969, SOP 9, Section 3.6 for assessment methodology.

4.5. Check standard measurement results obtained over time are used to calculate the standard deviation of the measurement process, s_t.

5 Assignment of Uncertainty

The NIST Mass Code generates uncertainties as a part of the data reduction. Proper input in the data file is critical for obtaining valid results and is dependent upon a well characterized measurement process. See NISTIR 5672 for a discussion on the input for standard uncertainties in the data file. Comprehensive calculations of the uncertainties for weighing designs must be completed outside of the NIST Mass Code at this time, and a supplemental job aid spreadsheet is posted on the NIST OWM website with this SOP for that purpose.

5.1 Calculating the standard uncertainty, u_S, of the starting restraint in the first series:

Usually the starting restraint will be one or several 1 kg (or 100 g) mass standards that have calibrations and density determinations from a National Metrology Institute or an accredited laboratory. The uncertainty of the standard as stated on a calibration report is divided by the appropriate coverage factor, dependent on the confidence interval stated in the calibration certificate.

5.1.1 One starting restraint scheme (a single starting standard), where U_S is the uncertainty from NIST which must be divided by the proper coverage factor, *k*.

$$u_s = \frac{U_s}{k} \qquad \text{Eqn. (3)}$$

5.1.2 Multiple starting restraint schemes with standards calibrated at the same time against the same starting standards, i.e., dependent calibration (more than one starting standard):

$$u_s = \frac{U_{s1}}{k_1} + \frac{U_{s2}}{k_2},$$

or Eqn. (4)

$$u_s = \frac{U_{s1}}{k_1} + \frac{U_{s2}}{k_2} + \frac{U_{s3}}{k_3},\ etc.$$

5.1.3 Multiple starting restraint schemes with standards *NOT* calibrated at the same time as the starting standards, i.e., independent calibration (more than one starting standard):

$$u_s = \sqrt{\left(\frac{U_{s1}}{k_1}\right)^2 + \left(\frac{U_{s2}}{k_2}\right)^2},$$

or Eqn. (5)

$$u_s = \sqrt{\left(\frac{U_{s1}}{k_1}\right)^2 + \left(\frac{U_{s2}}{k_2}\right)^2 + \left(\frac{U_{s3}}{k_3}\right)^2},\ etc.$$

5.2 Calculating the within-process standard deviation, s_w, for each series:

For each weighing design, the observed within process standard deviation, s_w, along with its degrees of freedom, *d.f.*, is pooled using the technique described in NISTIR 6969 Section 8.4.

$$s_w = \sqrt{\frac{(df_1)s_1^2 + (df_2)s_2^2 + \ldots + (df_k)s_k^2}{df_1 + df_2 + \ldots + df_k}} \qquad \text{Eqn. (6)}$$

5.3 Calculating the between-time standard deviation, s_b, for each series:

Establish a standard deviation over time, s_t, for each check standard over time. If a plot of the check standard shows no apparent drift, the between-time standard deviation may be calculated. The following formulae are used to calculate the between-time standard deviation for the series. If s_b^2 is less than zero, then s_b is treated as zero. See the supplemental job aid spreadsheet associated with NISTIR

5672 for many additional weighing designs, and variations on the designs noted in this section. The file lists restraint and check standard positions, with the associated K_1 and K_2 factors needed for calculating between-time standard deviations. Because of the way the K_n values are used to solve for s_b, it is possible to get a negative value in which case, zero is used as the value for s_b. NOTE: K factors as used in this section are not the same coverage factors that use the lower case *k* in uncertainty calculations.

5.3.1 For the 3-1 design with a single restraint, and a check standard that is either another single weight or a summation, the between time standard deviation is calculated using the following formula. The check standard may be in any other position.

$$s_b = \frac{1}{K_2}\sqrt{s_t^2 - K_1^2 s_w^2}$$
$$K_1 = 0.8165$$
$$K_2 = 1.4142 \qquad \text{Eqn. (7)}$$
$$s_b = \frac{1}{1.4142}\sqrt{s_t^2 - 0.8165^2 s_w^2}$$

5.3.2 Using a 4-1 design with two restraints, and the check standard is the difference between the two restraints, the next equation may be used to calculate the between-time standard deviation. If another weight in the series is used as the check standard, another equation is needed.

$$s_b = \frac{1}{K_2}\sqrt{s_t^2 - K_1^2 s_w^2}$$
$$K_1 = 0.7071$$
$$K_2 = 1.4142 \qquad \text{Eqn. (8)}$$
$$s_b = \frac{1}{1.4142}\sqrt{s_t^2 - 0.7071^2 s_w^2}$$

5.3.3 Using a 4-1 design with two restraints, and with a single check standard occupying any of the remaining positions, the next equation may be used to calculate the between-time standard deviation.

$$s_b = \frac{1}{K_2}\sqrt{s_t^2 + K_1^2 s_w^2}$$
$$K_1 = 0.6124$$
$$K_2 = 1.2247 \qquad \text{Eqn. (9)}$$
$$s_b = \frac{1}{1.2247}\sqrt{s_t^2 - 0.6124^2 s_w^2}$$

5.3.4 Using a 5-1 design with two restraints, and the check standard is the difference between the two restraints, the next equation may be used to calculate the between-time standard deviation. If another weight in the series is used as the check standard, another equation is needed.

$$s_b = \frac{1}{K_2}\sqrt{s_t^2 - K_1^2 s_w^2}$$
$$K_1 = 0.6325$$
$$K_2 = 1.4142$$
$$s_b = \frac{1}{1.4142}\sqrt{s_t^2 - 0.6325^2 s_w^2} \qquad \text{Eqn. (10)}$$

5.3.5 Using a 5-1 design with two restraints, and with a single check standard occupying any of the remaining positions, the next equation may be used to calculate the between-time standard deviation.

$$s_b = \frac{1}{K_2}\sqrt{s_t^2 - K_1^2 s_w^2}$$
$$K_1 = 0.5477$$
$$K_2 = 1.2247$$
$$s_b = \frac{1}{1.2247}\sqrt{s_t^2 - 0.5477^2 s_w^2} \qquad \text{Eqn. (11)}$$

5.3.6 In the second descending series (C.2), six weights are involved (500 g, 300 g, 200 g, 100 g, Check 100 g, and a summation 100 g). Calculate the standard deviations of the mass values for the Check 100 g (s_t) and plot the results to evaluate the presence or lack of drift. If no drift is present, the following formula is used to calculate the between-time standard deviation for this series and all subsequent C.2 series. Subsequent series include the following check standards: 100 g, 10 g, 1 g, 100 mg, 10 mg, and 1 mg. If s_b^2 is less than zero, then s_b is treated as zero.

$$s_b = \frac{1}{K_2}\sqrt{s_t^2 - K_1^2 s_w^2}$$
$$K_1 = 0.3551$$
$$K_2 = 1.0149$$
$$s_b = \frac{1}{1.0149}\sqrt{s_t^2 - 0.3551^2 s_w^2} \qquad \text{Eqn. (12)}$$

5.3.7 If a C.1 descending series is used, the following equation is used to calculate the between-time standard deviation when the check standard is in either of the last two positions:

$$
\begin{aligned}
s_b &= \frac{1}{K_2}\sqrt{s_t^2 - K_1^2 s_w^2} \\
K_1 &= 0.4175 \\
K_2 &= 1.0149 \\
s_b &= \frac{1}{1.0149}\sqrt{s_t^2 - 0.4175^2 s_w^2}
\end{aligned}
\qquad \text{Eqn. (13)}
$$

5.3.8 The between-time formulae shown here are those that are most common and are for descending series only. If another restraint or check standard is used, or if an ascending series is used, another formula will be needed. These formulae are statistically derived, based on the least squares analysis of the weighing design, and assume a normal, non-drifting distribution of measurement results. Equations for some other weighing designs may be calculated using the NIST Electronic Engineering Statistics Handbook. Section 2.3.3.2 "Solutions to Calibration Designs" gives an overview for deriving the solutions to weighing designs. It also provides the unifying equation for s_b (it is called s_{days} in the electronic handbook). To clarify the difference in terminology and notation the unifying equation for s_b is presented as:

$$
s_{days} = \frac{1}{K_2}\sqrt{s_2^2 - K_1^2 s_1^2} \qquad
\begin{aligned} s_{days} &\equiv s_b \\ s_1 &\equiv s_w \\ s_2 &\equiv s_t \end{aligned} \qquad
s_b = \frac{1}{K_2}\sqrt{s_t^2 - K_1^2 s_w^2} \qquad \text{Eqn. (14)}
$$

Section 2.3.4.1 "Mass Weights" provides the solutions for 17 weighing designs used for decreasing weight sets, 6 weighing designs for increasing weight sets and 1 design for pound weights. K_1 is found in the portion of the solution titled "Factors for Repeatability Standard Deviations", and K_2 is found in the portion titled "Factors for Between-Day Standard Deviations".

5.3.9 Note that there is a supplemental job aid spreadsheet that is available and includes many weighing designs with associated K_1 and K_2 values and the equations for calculating between time standard deviations. The latest version of this spreadsheet is posted on the NIST website with this Standard Operating Procedure. An additional spreadsheet job aid is available for final calculations of uncertainty using the output of the Mass Code report.

5.4 Uncertainty associated with bias, u_d. Any noted bias that has been determined through analysis of control charts and round robin data must be less than limits provided in SOP 29 and may be included if corrective action is not taken. (See NISTIR 6969, SOP 29 for additional details.)

5.5 Example components to be considered for an uncertainty budget table are shown in the following table.

Table 4. Example Uncertainty Budget Table.

Uncertainty Component Description	**Symbol**	**Source**	**Typical Distribution**
Uncertainty of the standard mass(es) (5.1)	u_s	Calibration certificate	Expanded uncertainty divided by coverage factor; combined for summations per SOP 29, restraint uncertainties assigned based on design calculations in the associated spreadsheet
Accepted standard deviation of the process as standard deviation over time, within-process standard deviation, and between time standard deviations (5.2, 5.3)	s_p s_t s_w s_b	Control chart, standard deviation chart	Normal Calculated based on K_1 and K_2 values for designs used
Uncertainty of the buoyancy correction	u_b	OIML R111	Rectangular
Air temperature (for air density)	u_t	SOP 2 or OIML R111	Rectangular
Air pressure (for air density)	u_p	SOP 2 or OIML R111	Rectangular
Air relative humidity (for air density)	u_{RH}	SOP 2 or OIML R111	Rectangular
Air density (formula)	$u_{\rho a}$	SOP 2 or OIML R111	Rectangular
Mass densities	$u_{\rho m}$	Measured and reported value OIML R111 Table B.7 Typically, 0.03 g/cm^3 to 0.05 g/cm^3	Rectangular
Uncertainty associated with bias (5.4)	u_d	Control chart, proficiency tests	See NISTIR 6969, SOP 29

5.6 Draft a suitable uncertainty statement for the certificate (e.g.,)

The uncertainty reported is the root sum square of the standard uncertainty of the standard, the standard deviation of the process, and the uncertainty associated with the buoyancy corrections, multiplied by a coverage factor of 2 (k = 2) for an approximate 95 percent confidence interval (95.45 %). Factors not considered in the evaluation: magnetism (weights are considered to meet magnetism specifications unless measurement aberrations are noted), balance eccentricity and linearity (these factors are considered as a part of the measurement process when obtaining the standard deviation of the process when using a check standard with adequate degrees of freedom.

NOTE: Where inadequate degrees of freedom are available, *k*, is determined using the appropriate degrees of freedom and the 95.45 % column in the table from Appendix A of NISTIR 6969, SOP 29.

6 Certificate

6.1 See NISTIR 6969, SOP 1, for Preparation of Calibration Certificates. Report measurement results as printed in Tables I and II as generated by the Mass Code. Actual text of the Mass Code certificate must be modified for each laboratory to be ISO/IEC 17025 compliant and to reflect accurate uncertainty values. Report the mass, conventional mass, environmental conditions during the calibrations, mass density used (reported, measured, or assumed (specify which type is used)), and calculated expanded uncertainties with coverage factors.

6.2 Compliance assessments.

Evaluate compliance to applicable tolerances as needed or required by the customer or by legal metrology requirements. Decision criteria for uncertainty and tolerance evaluations include two components: 1) the expanded uncertainty, U, must be $< 1/3$ of the applicable tolerances published in ASTM E617 and OIML R111 standards and 2) the absolute value of the conventional mass correction value plus the expanded uncertainty must be less than the applicable tolerance to confidently state that mass standards are in or out of tolerance. Compliance assessments must note the applicable documentary standard and which portions of the standard were or were not evaluated.