Automating Data Acquisition on a Mechanical Equal-Arm Balance Used in Large Mass Calibrations

Kevin Chesnutwood

Abstract: This paper focuses on efforts by the Mass and Force Group at the National Institute of Standards and Technology (NIST) to enhance the data taking operations of an approximately 40 year-old mechanical two-pan equal-arm balance used for high-precision mass comparisons in the range from approximately 50 kg to 1134 kg (110 lb to 2500 lb). The repeatability and sensitivity of this manually operated balance (called the Russell balance) is exceptional and therefore, the balance remains a core component of the NIST large mass laboratory. In order to enhance the data taking procedure and reduce the potential uncertainties inherent with human involvement in the data collection process, an automated system was designed and installed to obtain the turning points of the balance taken during calibration that ultimately are used to convert scale units to SI mass units. This paper discusses the advantages of the improved system, the challenges that had to be overcome, and the design, operation, and verification of the automated system.

1. Introduction

As part of the mission of the National Institute of Standards and Technology (NIST) Mass and Force Group, the Mass section has the responsibility of disseminating the SI (International System of Units) unit of mass to its customers, with calibration services that span a range from 1 mg to 30 000 kg. The Mass Group uses digital mass comparators for calibrations of artifacts up to approximately 50 kg. However, beyond the 50 kg threshold, mechanical balances are used due to their exceptional repeatability and sensitivity. While this paper discusses equipment used only in the range from 50 kg to 1134 kg, there is the potential for future application of the system to equipment used in calibrations performed at NIST up through 30 000 kg.

The mechanical two-pan equal-arm balance used at NIST for high-precision mass comparisons in the 50 kg to 1134 kg range is called the Russell balance and has been in service for more than 40 years. Section 2 provides an overview of the Russell balance, discusses how measurements are manually performed, and gives insight as to the reasons for automating these measurements. Section 3 outlines the automated approach to taking data on

Kevin Chesnutwood

National Institute of Standards and Technology (NIST)
Mass and Force Group, Mail Stop: 8222
100 Bureau Drive
Gaithersburg, MD 20899 USA
Email: kevin.chesnutwood@nist.gov

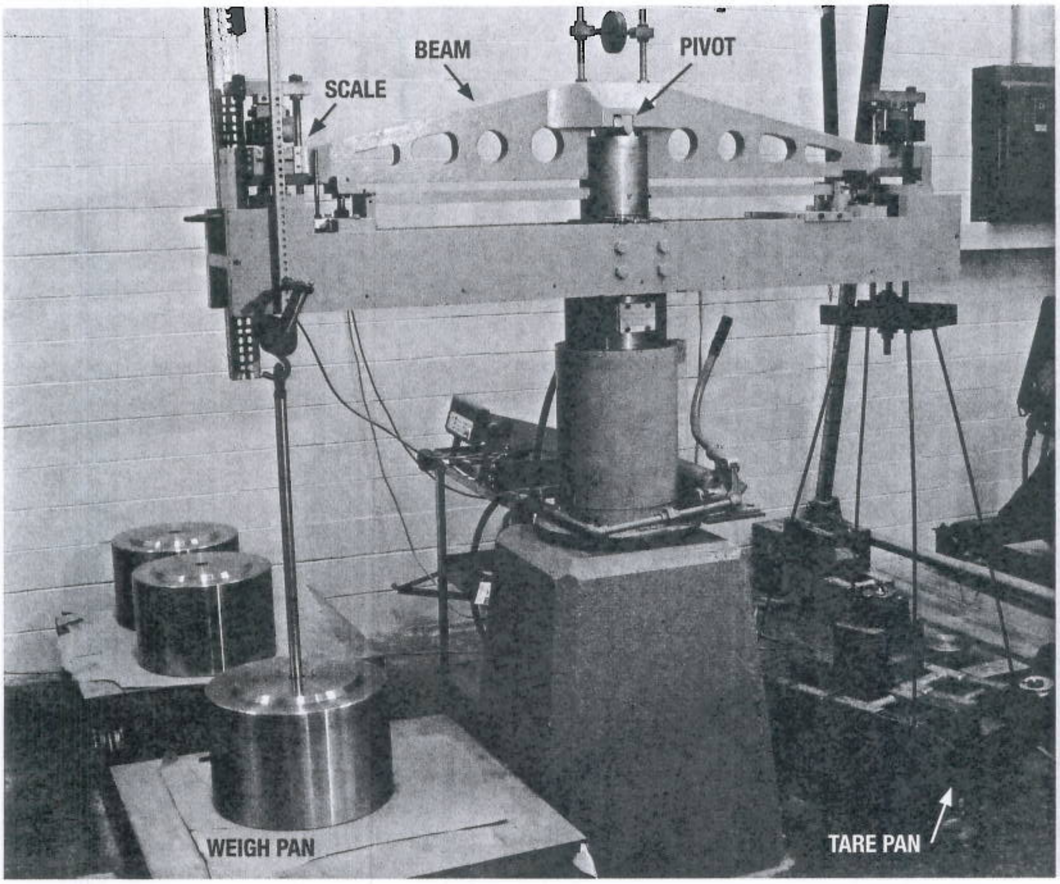

Figure 1. Photograph showing the Russell balance with a 226.8 kg (500 lb) stainless steel weight on the weigh pan.

the Russell balance[1] and the advantages of this approach. Section 4 compares the results computed using data sets acquired simultaneously by the traditional "manual" method and the overlaid automated data acquisition system, prior to putting the new system into service. Section 5 summarizes the conclusions and explores adapting the system to NIST's largest mechanical balance.

The purpose of this paper is to describe a new way to take measurements on an existing balance that reduces the uncertainties inherent in the current level of human interaction in the data-collection process, while preserving the desirable qualities of these types of mechanical balances. Refer to the publications listed in the references for more details and discussion regarding the intricacies of mass calibrations and general mass metrology. [1-3]

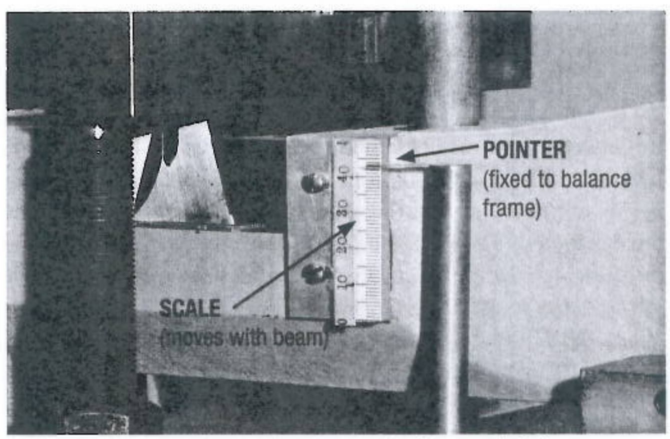

Figure 2. Photograph showing the detail of the scale/pointer assembly that was used when manual measurements were taken.

2. Overview and Operation of the NIST Russell balance

The Russell balance is a mechanical two-pan equal-arm balance used for high precision mass comparisons. To setup and use the balance as a comparator, one pan (sometimes called the tare pan or the counterpoise) of the balance is loaded with enough counterweight to match (or balance) the opposing pan (weigh pan) which is interchanged between reference standards and a weight for which a mass value will be assigned, sometimes called the unknown (see Fig. 1). The beam, when properly balanced, will oscillate (or pivot up and down) about the center knife edge and will eventually come to rest after a potentially long time interval. The damping of the beam oscillation is usually quite slow since there is minimal friction through the knife edges of the balance. Therefore, in order to determine a deflection (or reading) from the balance as it is performing this oscillating motion, an operator can read the magnitude of these oscillations with a pointer indicating on a scale (located on the fixed frame of the balance) as the end of the beam travels

[1] Certain commercial equipment, instruments, software, or materials are identified in this paper. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

up and down in the vertical plane (see Fig. 2). Rather than wait for the beam to come to rest and taking a static reading on the scale (which likely wouldn't be prudent due to friction sticking), it is more practical and accurate to record the highest and lowest readings of the beam travel while the beam is changing direction during it's oscillation. These peak readings are referred to as turning points.

Once the swing is induced, the operator usually waits at least one complete oscillation before beginning to record turning points. As turning points are recorded, the operator must ensure that the points are valid by determining if the balance is in "decay," meaning the amplitude of the oscillation is getting smaller (or converging to the final resting place). Sometimes the balance will accelerate (meaning the last top reading is greater than the previous top reading for example) and requires the discarding of these points and waiting until the balance renders a consistent pattern of decay. Once valid turning points are taken, the operator then takes consecutive Bottom-Top-Bottom (BTB) readings (or Top-Bottom-Top readings (TBT)) and averages the two Bottom readings (or two top readings if TBT) and adds it to the reading taken in between. The result is then divided by two and is called the computed mean. Enough valid turning points are taken to ensure that at least three computed means are calculated. The computed means are averaged to obtain the final average computed mean (where the balance would stop if left to decay long enough).

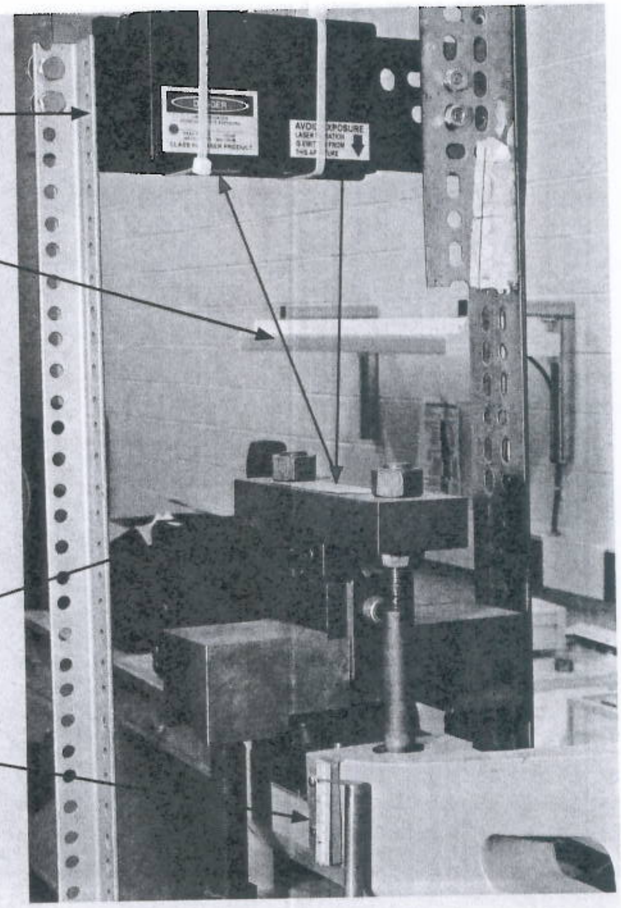

Figure 3. Photograph showing the laser measurement sensor with simulated laser beam path. As the balance beam oscillates up and down, the laser measurement sensor triangulates the height (correlated from a DC voltage) from the fixed sensor to the moving balance beam.

At NIST, mass comparisons in this range are generally performed using the "double substitution" method that entails a comparison between a reference mass and an unknown mass and requires the use of a sensitivity weight. [1] The purpose of the sensitivity weight is to measure the scale deflection per unit mass, or in other words relate scale units to mass units. During setup, with the unknown placed on the weigh pan, the beam oscillation is adjusted to keep the pointer "on scale" by adjusting the counter mass on the tare pan. The sensitivity weight, which is a small reference weight, is then added to the weigh pan to ensure that even with the increased deflection due to the additional mass, the turning points remain "on scale."

After these initial adjustments of the tare pan resulting in an average computed mean somewhere near the scale's midpoint, no additional adjustments are made to the tare pan for the rest of the comparison process. Each double substitution comparison requires four weighings on the balance. Additionally, a weighing design is chosen that incorporates a series of these difference measurements that, by fixing the value of the reference standard, allows the mass of the unknown to be determined by using the method of least squares. [4] These methods were developed at NIST by Cameron et al. in 1979, and a full description can be found in reference [4]. The weighing design used throughout this paper is referred to as a "3-1" indicating that it involves three masses of equal nominal mass. In this design, the third mass used in the comparison is another reference standard called a "check standard," and it is treated as an unknown (meaning in this case that the mass value is not fixed in the calculations).

Once all masses being used in the weighing design are adjusted to produce "on scale" readings, the operator can then execute the proper sequence of the weighing design by interchanging, on the weigh pan, the unknown, the reference standard, the check standard, and all combinations with sensitivity weights.

As can be seen, the process for one comparison involves numerous time consuming measurements. For example, a typical 3-1 weighing design will require twelve changes to the weigh pan, meaning that there are twelve separate average computed means recorded (requiring about 100 turning points all together). Additional time is needed for changing weights, hydraulically raising or lowering the balance, changing lifting hardware, etc. Operators should not be interchanged during

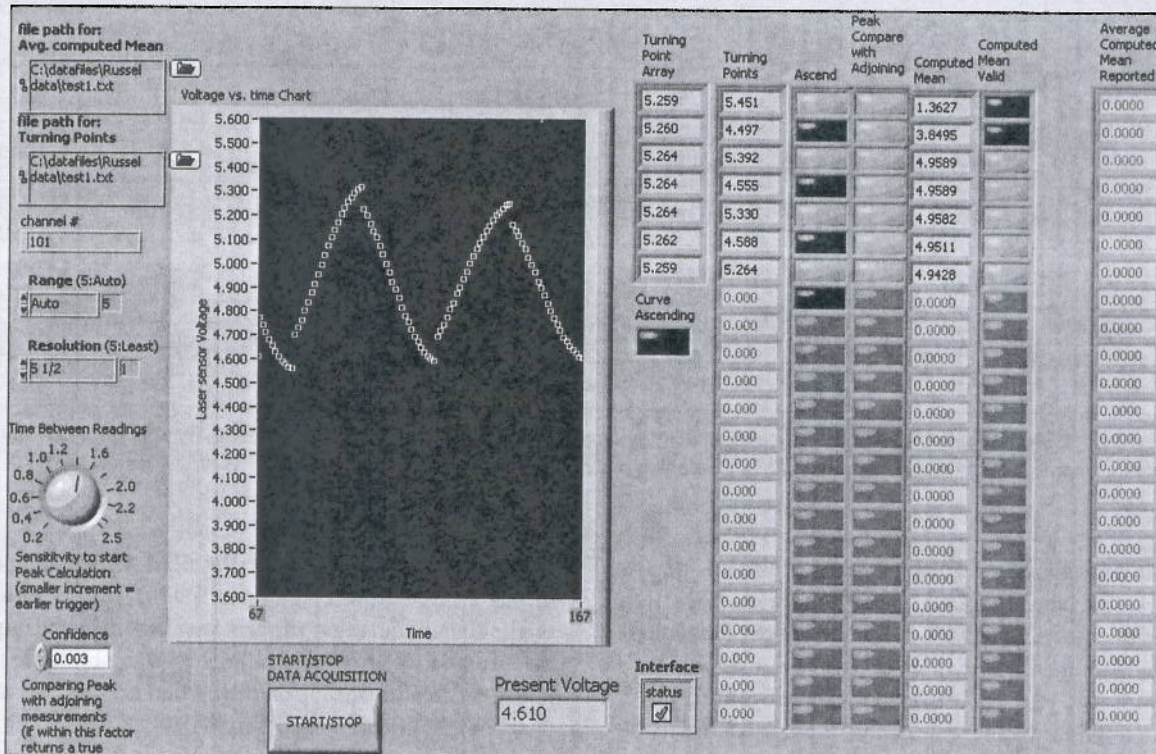

Figure 4. Snapshot of LabView program that controls the data acquisition and calculates the turning points, computed means, and the average computed mean. Note: Chart does not show what one may expect as the typical sine wave due to the program not updating the chart while the high speed acquisitions of the turning point data are being gathered.

this comparison in order to avoid possibly introducing an operator bias due to different operator stance, height, optical characteristics, and operating convention. Therefore, the operator can suffer fatigue and eye strain and actually introduce a bias due to becoming restless and changing body stance which ultimately affects the optical alignment of the eyes with the scale (parallax due to changing lens location). A misread reading and/or changing posture of the operator can lead to poor repeatability and/or increased standard deviations in the averages. Additionally, all measurements are hand recorded during this procedure, increasing the odds of a transcription error and the possibility that the measurements may need to be repeated.

3. Automating the Reading of Turning Points

A solution to the aforementioned problems associated with manual measurements has been accomplished by removing the human aspect of physically reading the scale. An automated system was designed and installed on the balance using a laser measurement sensor to determine the turning points of the balance. A picture of the laser sensor setup is shown in Fig. 3. The sensor applies the principle of triangulation and uses a laser beam reflected off of the top knife edge assembly near the end of the beam to relate the displacement of the beam relative to the fixed sensor, much like the mechanical pointer. However, the same beam displacement is correlated to an analog voltage that can be read at a much higher resolution and more consistently than can be achieved manually by eye. For example, the manual process using the scale and pointer divides the vertical travel of the balance into approximately 300 divisions (the scale shown in Fig. 2 is only usable between 10 and 40 due to mechanical stops). On the other hand, in addition to having a better resolution, the laser sensor is located slightly further out on the beam and gains some additional travel in its measuring range. Therefore, the laser sensor can divide the measuring range into about 1600 divisions, a significant improvement in resolution over the manual system.

The sensitivity weight is still used to correlate the scale units to mass units, in this case to correlate a voltage to mass units. The laser measurement sensor is a DynaVision Model SPR04. A standard 15 Volt DC power supply is used to power the laser while the analog output voltage is read by an Agilent 34970 Data Acquisition and Control Unit containing a precision 5-1/2 Digit Digital Multi-Meter (DMM) capable of adequately reading the resolution of the sensor which is 1 mV. The digital signal is recorded using a laptop computer running a LabView program communicating through a National Instruments Universal Serial Bus (USB)/General Purpose Interface Bus (GPIB) interface.

Date	F-test Manual	F-test Automated	t-test Manual	t-test Automated	Mass Manual (kg)	Mass Automated (kg)	Difference in Mass (Absolute Value, mg)	Uncertainty Reported (mg)
4/07	0.09	0.15	0.72	0.88	45.414 008	45.414 019	11	372
11/06	0.10	0.26	-0.22	-0.30	90.828 536	90.828 537	1	372
1/07	0.01	0.00	0.81	0.86	363.314 961	363.314 972	11	646
1/07	0.00	0.28	-0.07	-0.06	363.315 007	363.314 997	10	646
6/07	1.56	1.99	0.26	0.21	454.143 19	454.143 18	10	646
4/07	0.05	0.07	-0.04	-0.01	453.624 85	453.624 85	0	646
1/07	0.00	0.00	0.39	0.42	500.002 319	500.002 290	29	664

Table 1. Results showing comparisons between the manual method and the automated method. Automated and manual data were taken simultaneously for each individual comparison using a 3-1 weighing design.

In order to determine the turning points using this system, the operator uses a LabView program to take voltage readings from the sensor as the beam is oscillating. The time between readings of the voltage is adjustable by a control on the user interface to the program (see Fig. 4). This allows the operator to find a sufficient reading rate based on how fast the beam is traveling (which has been found to vary under different load conditions). The program then determines if the beam is ascending or descending by comparing the current reading with the previous reading. As the beam loses momentum and nears a direction change (turning point), the voltage measurements change at a slower rate. When this condition occurs, the LabView program triggers the DMM to record voltage measurements at a very high rate (50 readings per second) throughout the period while the balance is changing direction. From this group of readings, the program determines the maximum/minimum (depending on which direction the beam is traveling) and uses that peak value as the turning point.

To ensure that the value of the recorded peak is reasonable and fits the trend of the recorded data, an algorithm in the program compares the peak reading to several of the adjoining readings taken on both sides of the peak. If the algorithm determines an invalid peak recording, the program indicates that the turning point is invalid and it is not used in calculating a computed mean. Additionally, as more turning points are gathered, the program determines whether the balance is in proper decay, and hence whether the turning point can be used in calculating a valid computed mean. When the program determines that it has obtained three consecutive valid turning points, the computed mean is calculated and displayed. The operator then can simply watch the screen and determine when enough valid computed means have been acquired (usually three) and can stop data acquisition. Another screen then appears allowing the operator to choose which valid computed means are to be used in calculating the average computed mean for that measurement. The average computed means are then stored in a separate file for incorporation into the final mass calculations, thus eliminating the need to manually transcribe the data. While the operator still needs to be an active part of the data acquisition process and still must manipulate the balance and weights, the strain of staring at a scale for several hours, calculating means on the fly, and properly recording data were now eliminated and the likelihood of having to repeat measurements is greatly reduced.

4. Validation and Verification of the Automated System

The automated system did not require any changes to the mechanical readout system that is already in place. Therefore, verification of the automated system could easily be determined by direct comparison with manual data taken under identical conditions. In this manner, numerous comparisons were completed using a 3-1 weighing design. The results of these comparisons are shown in Table 1.

As shown in the table, verification was performed using the same statistical tests (metrics) and by direct comparison of the mass and uncertainties calculated from each correlating set of data. Briefly, one metric used is the "F-test" (or "F-ratio") which is used to analyze balance performance during a comparison (more detail in contained in reference [1]). Basically, the standard deviation of the current data set is checked for consistency by comparing it to the long-term standard deviation of the balance at a 95 % confidence interval (k = 2). Therefore, if the automated data compares well with the manual data in each test, then it is safe to assume that it agrees with the long-term performance of the balance. The magnitude of the difference between the automated and manual data for the F-test is small (average = 0.14) and establishes that the data collected using the automated system are consistent with that collected using the manual system and also are in statistical control with the long-term history of the balance.

A second statistical check used to help verify the automated system is the "t-test" (more detail is obtained in reference [1]).

Date	F-test Manual	F-test Automated	t-test Manual	t-test Automated	Mass Correction Manual (mg)	Mass Correction Automated (mg)	Difference in Mass Correction (Absolute Value, mg)	Uncertainty Reported (mg)
10/22/08	0.07	0.01	1.00	1.25	2242	2255	13	646
10/23/08	0.63	0.67	0.72	0.75	1813	1811	2	646
10/27/08	0.07	0.09	1.58	1.46	1893	1901	8	646
10/28/08	0.73	0.61	-0.02	0.03	1953	1967	14	646
11/25/08	Not Taken	0.39	Not Taken	2.12	Not Taken	2140	Not Taken	646

Table 2. Repeatability/reproducibility tests using NIST 226.8 kg stainless steel reference standards in a 3-1 weighing design. Results calculated for a NIST standard acting as the unknown.

Each comparison uses at least two standards, in addition to the unknown. One reference standard (the "restraint") is used to calibrate both the unknown and the remaining reference standard (check standard). The newly derived value for the check standard is then compared to the long-term average of the check standard. Any statistically significant difference in the two values usually indicates a physical change in one (or both) of the standards or, more commonly, that there was an error made during the comparison. The average t-test value difference between the automated and the manual data is 0.05 and further indicates that the automated system is consistent with the manual system.

The last check used to verify the automated system involved directly comparing the calculated mass and uncertainty of the unknown by both methods. As seen in Table 1, the differences in the calculated masses were at least 23 times smaller than the uncertainty in the measurement itself. The average difference was 104 times smaller than the uncertainty (or about 1.8 % of the measurement uncertainty). The magnitude of these differences provides the last step in confirming the precision, consistency, accuracy, and the validation of the automated system in comparison to the manual system.

Additional repeatability and reproducibility comparisons were completed in a 3-1 weighing design that used several NIST stainless steel reference standards. The comparison was repeated over several days using the same weights, weighing design, and setup. Results of the repeatability/reproducibility tests are shown in Table 2.

The same metrics for examining the data were used for this comparison, keeping in mind that one of NIST's stainless steel standards was acting as the unknown. As seen in Table 2, the F-test and t-test comparisons indicate that the automated and manual data were consistent and in statistical control. Table 2 also shows that the differences in the calculated mass corrections (from a nominal mass) were at least 46 times smaller than the uncertainty of the measurement, with an average difference that was 125 times smaller (or 1.4 % of the measurement uncertainty). Additionally, the last line of the table only shows data taken by the automated system because no manual readings were taken during this comparison.

5. Conclusions

The automated data gathering system for the Russell balance is a marked improvement to the data taking procedure for this mass measurement range. It improved operator comfort, as well as eliminated many human factors in the measurement process. Thus, the possibility of introducing random errors inherent with human data-reading/logging were reduced without compromising the advantages of mechanical balances used in these weight ranges. Although the balance characteristics limit a large gain in repeatability from the increased sensitivity of the laser sensor, the system showed enough benefits to warrant possible adaptation to the NIST 30 000 kg balance, a single beam balance that also incorporates reading turning points.

6. Acknowledgments

The author gratefully acknowledges valuable input and direction from Dr. Zeina Jabbour and Mr. Brian Scace of the NIST Mass and Force Group.

7. References

[1] K.B. Jaegar and R.S. Davis, "A Primer for Mass Metrology," *Special Publication 700–1*, National Bureau of Standards, Gaithersburg, MD, November 1984.

[2] Z.J. Jabbour and S.L. Yaniv, "The Kilogram and Measurements of Mass and Force," *Journal of Research of the National Institute of Standards & Technology*, vol. 106, pp. 25-46, 2001.

[3] R.S. Davis, "NIST Measurement Services: Mass Calibrations," *Special Publication 250–31*, National Institute of Standards & Technology, Gaithersburg, MD, January 1989.

[4] J.M. Cameron, M.C. Croarkin, and R.C. Raybold, "Designs for the Calibration of Standards of Mass," *NBS Tech Note 952*, National Institute of Standards and Technology, Gaithersburg, MD, 1977.