Uncertainty Propagation for Force Calibration Systems

Henry Zumbrun and Alireza Zeinali
Morehouse Instruments

Figure 1. Measurement Uncertainty Pyramid

Introduction

There are several labs operating throughout the world, whom do not follow a designated guideline for calculating measurement uncertainty for force calibrations. Realizing the need for a guidance document, Morehouse decided to draft this document explaining how to calculate measurement uncertainty and how uncertainty propagation for force calibration systems works. Calibration and utilization of measurement instruments will imply some level of uncertainty. As an instrument calibration is traced back to SI units, a higher number of intermediate calibration stages results in higher levels of measurement uncertainty (Figure 1) [1]. In other words, uncertainty of the unit under test is typically higher than the standard with which it was calibrated. It is not possible for the expanded measurement uncertainty of the unit being calibrated to be less than the machine or force measuring device that is used to calibrate the unit itself. This paper describes the propagation of uncertainties using Calibration and Measurement Capability (CMC) for force measurement instruments through the traceability chain to SI units. For the instrument users who require some minimum level of expanded uncertainty, this paper provides information on what level of calibration is needed for their reference standards.

Test Plan and Equipment

A 445 kN (100k lbf) Morehouse Ultra-Precision Load Cell was chosen for the testing plan. The calibration test setup is shown in Figure 2. The Morehouse load cell provides relatively high stability, resolution, and repeatability. Consequently, the testing plan represents an almost best-case scenario: the lowest level of Calibration and Measurement Capability (CMC) that a load cell user can achieve at each level of the traceability chain. A 89 kN (20k lbf) test point was chosen for analysis based on historical data. This load point was chosen for studying the CMC propagation to follow the ILAC P14 requirements [2]. Morehouse Ultra-Precision 445 kN (100k lbf) systems can often use this load cell in the Tier 2 group from 20 % to 100 % of capacity for force calibration purposes without switching standards. The reference standard of Tier 2 in this paper represents a load cell that is calibrated in accordance with ASTM E74 standard test method with using other load cells with ASTM Class AA designation [3]. Additionally, the 20 % point represents a pivot point for achieving CMC of approximately 0.02 % of applied force. At higher forces, the CMC is typically lower. However, at lower than 20 % of capacity forces, CMC starts to increase; it continues to increase to the 10 % and lower force points, where the CMC becomes higher than 0.05 % of applied force. Therefore, it is often recommended that the end user in Tier 2 only uses the load cell from 20 % through capacity in order to maintain CMC's better than 0.02 % of applied force.

Tier 0: CMC for Primary Standards

In this tier, Calibration and Measurement Capability (CMC) for Morehouse's deadweight calibration systems were determined.

Table 1 contains the uncertainty contributors for this calculation, along with their appropriate divisors. It should be noted that the testing for this study was conducted based on United States customary units, and then converted to SI units in Table 1 to make it more tangible for international users. Degrees of freedom and coverage factors were calculated separately using the Welch-Satterthwaite equation [1]. In this tier, Morehouse had the reference deadweights calibrated directly by N.I.S.T. These weights, pictured in Figure 3, were adjusted for the local gravity, material density, and air buoyancy, and their traceability is derived from the international prototype kilogram (SI unit symbol kg) [3].

When the calibration was performed in a Morehouse deadweight machine, CMC was calculated using these weights. A repeatability study was conducted with three high quality Morehouse load cells (445 kN; 111 kN; and 44 kN capacities) throughout the entire range of the machine. Morehouse's CMC resolution for 89 kN (20k lbf) load was used for UUT resolution in Tier 0 only. This value was determined based on a 111 kN (25k lbf) load cell with 4 mV/V output at capacity and 0.00001 mV/V readability.

The environment was controlled by better than +/- 1.0 °C [3], while the stability of the weights was calculated using historical values for the material and years of wear history from our other deadweight machines. The resolution of the weights was zero since they are physical standards, and

Figure 2. 445 kN (100k lbf) Load Cell in Deadweight Machine Being Calibrated

Figure 3. View of Deadweight Machine

TIER >>>			TIER 0 Primary Standards	TIER 1 Primary Lab	TIER 2 Secondary Lab	TIER 3 Working Standard
UUT Info >>>			No UUT (Deadweight CMC Calculation)	Load Cell Calibrated by Primary Standard (Class AA Assigned)	Load Cell Calibrated by Secondary Standard (Class A Assigned)	Load Cell Calibrated in Force Press
Uncertainty Source		**Divisor**	**Primary Cal (Deadweight)**	**Primary Cal (Deadweight)**	**Working Cal (UCM)**	**Field Cal Lab (scale calibrator)**
Reference	U_{REF}	2	0.396893 N †	1.42 N	17.57 N	27.45 N
Resolution (Reference)	$U_{RES,REF}$	3.464	N/A (deadweight)	1.07 N	1.07 N	1.07 N
Resolution (UUT)	$U_{RES,UUT}$	3.464	0.2780 N ††	1.07 N	1.07 N	1.07 N
UUT Repeatability	U_{REP}	1	0.2567 N	1.7646 N	1.7646 N	1.7646 N
B/W Techs Reproducibility and Repeatability	$U_{R\&R}$	1	0.49 N	3.910 N	3.910 N	3.910 N
Stability	U_{STA}	1.732	0.0178 N	4.45 N	4.45 N	4.45 N
Environmental	U_{ENV}	1.732	Included in U_{REF}	0.667 N †††	0.667 N	0.667 N
Side Load Sensitivity	U_{MISC}	1.732	N/A (deadweight frame)	2.67 N	2.67 N	2.67 N
ASTM Lower Limit Factor (LLF)	U_{ASTM}	2.4		18.296 N (Class AA Assigned)	23.718 N (Class A Assigned)	33.36 N *
Expanded Uncertainty	U	-	0.0016 % (1.42 N) ‡	0.01974 % (17.57 N) ‡‡	0.031 % (27.45 N) ‡‡‡	0.106 % (97.42 N)

Table 1. Uncertainty Propagation Analysis for Load Cell Calibrations [1-5]

Notes:
Expanded Uncertainty may be the CMC of the lab, depending on the Tier.
Table is made based on data from a 445 kN (100k lbf) Morehouse Ultra-Precision load cell. Testing was conducted at 88.964 N (20k lbf) force (20% of capacity).
UCM: Universal Calibrating Machine
† Includes: uncertainty of deadweight calibrations by NIST, uncertainty of material density, uncertainty of gravity, uncertainty of air density
†† CMC Resolution = Capacity (25 kip) / (load cell output) × Readability (This was added on top of the regular UUT resolutions. There is an ongoing debate on whether or not this extra item is necessary.) This is the resolution of the electronic indicator(s).
††† 0.0015% per °C (± 1 °C)
‡ Coverage factor = 2
‡‡ Coverage factor = 1.97
‡‡‡ Coverage factor = 1.98
* LLF was calculated according to the E74 standard method. However, this would not be valid since the standard requires a minimum Class A limit for calibration standard.

the resolution of a good measurement system (Morehouse Ultra-Precision Load Cell coupled with HBM DMP 40 indicator) was used as an uncertainty contributor for UUT resolution. Various technicians' tests were compared to determine the repeatability and reproducibility per point of the Morehouse deadweight calibration machine. All of these efforts, combined with continued process monitoring, yielded a CMC of better than 0.0016 % of applied force.

Tier 1: Using Primary Standard Deadweights to Calibrate a Load Cell

For Tier 1 calibration, the deadweight calibration machine was utilized to calibrate a load cell in accordance with the ASTM E74 standard [3]. More on this calibration procedure is explained online at: http://blog.mhforce.com/2016/02/astm-e74-calibration-procedure.html. A Morehouse 445 kN (100k lbf) load cell was calibrated in this tier by deadweight primary standards known to have a CMC better than 0.016 % of applied load.

To calculate the CMC of the calibration, a repeatability and reproducibility (R&R) study was done for Tier 1 using a 111 kN (25k lbf) Ultra-Precision load cell. Moreover, an environmental condition of ±1 degree Celsius, along with a stability value of 0.005 % (50 parts per million), was used for calculating uncertainty values. The actual resolution of the UUT load cell 1.07 N (0.24 lbf) was employed for uncertainty calculations in Tier 1. It might be noteworthy to mention that the reference uncertainty used in Tier 1 already included the UUT resolution embedded in deadweight CMC calculations. Basically, UUT resolution is considered twice in the calculation of uncertainties for Tier 1–3 [4]. This method is on the conservative side of the uncertainty calculations, and there is ongoing debate about whether or not the resolution from CMC must be included in higher calibration tiers.

Load cell output stability is another of the uncertainty contributors when the cell is calibrated per ASTM E74. Stability is calculated by comparing the load cell output to the previous calibration data [3]. Most Morehouse Ultra-Precision load cells provide a one year stability of around 0.005 % through 0.01 %. Typically, the actual numbers would be used for this evaluation; however, this test was controlled, and the experiment could not wait another year to obtain the actual UUT load cell's stability numbers.

Ideally, load must be applied to the primary loading axis of any load cell in order to produce most repeatable and accurate results. This primary loading axis for shear web load cells such as the one used in this study, generally falls on the axisymmetric axis of the cell. However, in reality, some side loading is traduced into the loading system which can influence the load cell output. Side loading on a shear web load cell is demonstrated in Figure 4. Morehouse Universal Calibrating Machine (UCM) can provide side loading of better than 1/16th of an inch. Additionally, the side load sensitivity of a Morehouse Ultra Precision load cell is 0.05 % of load per inch of side loading. Multiplying 1/16th of an inch by 0.05 % yielded an uncertainty contribution of 0.003 % of applied load.

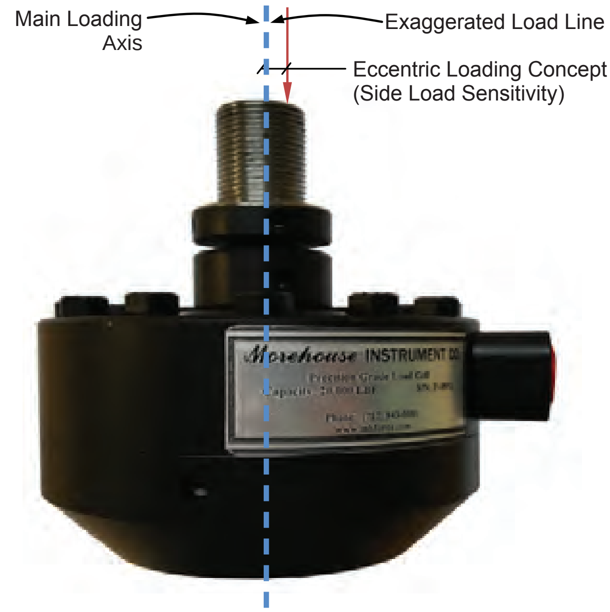

Figure 4. Side Loading on a Load Cell

The ASTM E74 calibration and analysis results in a Lower Limit Factor (LLF), which is the standard deviation of variations in different runs multiplied by a coverage factor of 2.4. The UUT load cell in Tier 1 was assigned a Class AA loading range, which provides a test accuracy ratio (TAR) of better than 5:1 when used to calibrate another load cell in accordance with the ASTM E74 standard. In this range, the calibrated load cell (UUT) can be used to calibrate other load cells that will be used to calibrate force measuring or testing machines [3]. As presented in Table 1, the expanded uncertainty for Tier 1 calibration was 0.01974 % of applied force, or 17.57 N (3.95 lbf) at 89 kN (20k lbf) force. This value was applied as the reference uncertainty in Tier 2 calibration.

Tier 2: Using a Load Cell Calibrated by Primary Standards to Calibrate Other Load Cells

In this tier, the Working Standard load cell was calibrated in accordance with the procedures outlined in the ASTM E74 standard. ASTM E74 fits the data points to a higher order curve using the least squares fit method [3]. This is different than just linearizing a load cell. To run the test, a second Morehouse 100k lbf Ultra-Precision load cell was calibrated using the Morehouse Universal Calibrating Machine (UCM). As previously stated, this paper represents a chain of calibration for high quality instruments and calibrations currently available in the industry. For this reason, the Morehouse Ultra-Precision load cell was used for all calibration levels. Using other instruments with lower performance quality would potentially increase the uncertainty results reported.

In Tier 2 Calibration, identical resolutions were used for both the reference cell and the Unit Under Test (UUT). The first Morehouse Ultra-Precision cell that was calibrated to primary standards in Tier 1 was employed in Tier 2 to calibrate the UUT (the second 445 kN Morehouse Ultra-Precision load cell). The CMC that resulted from Tier 1 calibration (17.57 N) was employed as the reference uncertainty at this level. The same uncertainty contributors were used and a new ASTM LLF was calculated.

Based on the calibration data, the LLF was calculated and an ASTM Class A loading range that provides a test accuracy ratio (TAR) of better than 4:1 was assigned[1]. This calibration produced a working standard with an assigned class A loading range [3]. As shown in Table 1, the resulting expanded uncertainty for Tier 2 calibration is 0.031 % of applied force, or 27.45 N (6.17 lbf) at 89 kN (20k lbf).

Tier 3: Using a Working Standard Load Cell to Calibrate Field Equipment

Tier 3 was meant to simulate the conditions of a field calibration test. In the ASTM E74 pyramid, the working standard that was calibrated in Tier 2 (accredited calibration supplier or secondary standard) could only be used to calibrate testing machines. However, the testing plan presented was conducted in a controlled laboratory environment to simulate the best-case scenario for uncertainty propagation. Thus, the same testing regime, with load cell and UCM, was followed for Tier 3. Nonetheless, an aircraft scale calibrator (such as Morehouse 804000) could have been used. For this calibration, the ASTM LLF was reduced to a pooled standard deviation to perform what would normally be the calibration of a testing machine. Since an identical setup as in Tier 2 was utilized for this test, the uncertainty contributors remained the same; however, the ASTM LLF increased again. The ASTM LLF increase was due to the higher expanded uncertainty bands of the reference.

1 Normal Metrology Practices discourage TAR. ASTM E74 was developed in 1974 and still relies on a method using TAR where the maximum error of primary standards are to be no more than 0.005 % of applied force, Secondary Class AA Standards are no more than 0.05 % and Field Standards are no more than 0.25 % [3]. This equates to TAR's of 10:1, 5:1, and 4:1. Contemporary conventions of metrological science no longer focus on a TAR in establishing decision risk criteria. Most modern practices focus on TUR (Test Uncertainty Ratio) for a measure of adequate decision risk criteria [6].

Repeatability and Reproducibility (R & R) tests were conducted at each tier. In Tier 0, we used the same R & R values as reported in our CMC. In Tiers 1 through 3, we used a R & R study we conducted in house and repeated the number throughout tiers 1 through 3 [1]. The full explanation for B/W Techs Reproducibility and Repeatability can be found in section 7. We would expect the R & R between technicians to grow larger throughout the remaining tiers as well as the resolution of the Unit Under Test because the UUTs at each tier will typically be less accurate than what was used for these tests.

The uncertainty calculations in Table 1 resulted in CMC for Tier 3 equal to 0.106 % of applied force at 89 kN (20k lbf). It might be worth mentioning that actual Tier 3 testing would produce much higher CMC than shown in Table 1 since the stability per point would most likely increase, as would the resolution of the UUT. It is important to note that the end calculation will inevitably be higher than what we have shown.

Explaining CMC Calculations Contributors

All Calibration and Measurement Capabilities were calculated using a combination of A2LA document R205, ILAC P-14, GUM, and the appendix in ASTM E74, which call for the following [1-4]:

1. **Repeatability** – Repeatability was defined as the standard deviation of 10 measurements with the same load cell at a 89 kN (20k lbf) force point. The Tier 0 number was derived from Morehouse's Calibration and Measurement Capability, which was submitted to the company's accreditation body. For Tiers 1–3, repeatability was measured between two technicians, using a 111 kN (25k lbf) load cell, loaded to 89 kN (20k lbf), 10 times each in a 445 kN (100k lbf) Universal Calibrating Machine (UCM).
2. **Resolution** – Resolution was recorded as the resolution of both the Unit Under Test and the Reference Standard. In Tier 1, there was only one contribution from the UUT since the deadweight calibration machine is equipped with intrinsic standards. Per JCGM 200:2012 Resolution is the smallest change in quantity being measurement that causes a perceptible change in the corresponding indication.
3. **Reproducibility** – Reproducibility was determined using an R&R study. Each of the two technicians performed 10 runs of data, and their overall results were compared against one another. A standard deviation of the average was calculated between technicians and used for the final reproducibility number for all tiers.
4. **Reference Standard Stability** – For Tier 0, historical data and Statistical Process Control Data were used

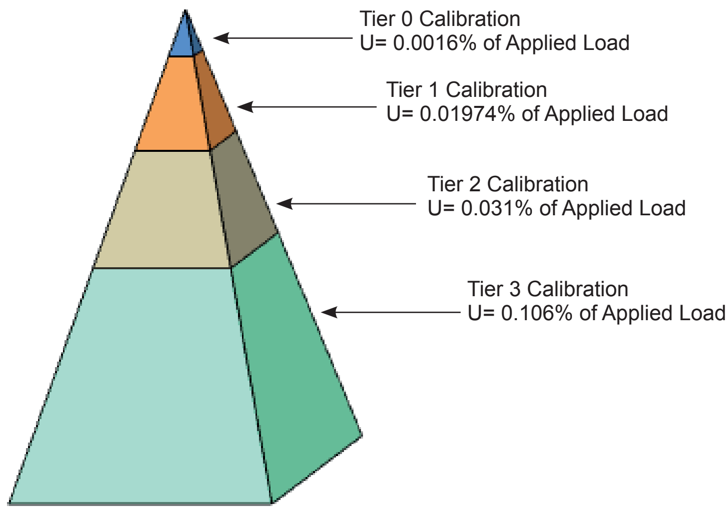

Figure 5. Uncertainty Propagation for 89 kN (20k lbf) through Various Tiers

to calculate stability. For Tiers 1–3, stability of 0.005 % was assumed. This is based on historical data on Morehouse Ultra Precision Load Cells. The number represents an approximation of historical data from several of these load cells.

5. **Environmental Factors** – A change of ±1 degree Celsius was used, and the corresponding effect on load cell output was determined. Generally, load cells of this type have a temperature specification of 0.0015 percent reading per °C.
6. **Miscellaneous Errors** – This consisted of side load sensitivity for the Morehouse calibration machine assuming a maximum of 1/16th of an inch of misalignment.
7. **ASTM LLF** – This is calculated as per the ASTM E74 standard and was reduced to a pooled standard deviation. The ASTM E74 standard can be found online at: https://www.astm.org/Standards/E74.htm. The ASTM E74 standard uses a method of least squares to fit the data points. The standard deviation of the all of the deviations from the predicted values versus the observed values is found by taking the square root of the sum of all of the deviations divided by the number of samples minus the degree of polynomial fit used minus one.
8. **Reference Standard Calibration Uncertainty** – This was calculated using the Welch-Satterthwaite equation, and is a combination of the sum of the squares of all above contributors. The reference standard uncertainty was then transferred from tier to tier, absorbing additional uncertainty contributors per tier.

Conclusions

Based upon the testing information presented from and supported by years of testing, this summary should help guide users in determining what uncertainty they can obtain while using various force standards. If a CMC of better than 0.03 % of applied force is desired, calibration by primary standards (deadweight) is necessary. Figure 5 illustrates the predicted minimum uncertainties that can be achieved by various laboratory tiers. The figure indicates that an additional reference standard would be needed at every 20 % interval to maintain better than 0.02 %. In other words, a 500-kN Universal Calibrating Machine would need reference standard load cells or proving rings with capacities of 445, 89, and 22 kN (100k, 20k, and 5k lbf respectively) to achieve 0.02 % of applied load or better with a force range of 4.450 kN (1k lbf) through 445 kN (100k lbf).

The testing proved the importance of the reference standard in relation to overall expanded uncertainty. Deadweight primary standards are predictably the best possible reference standard. A laboratory using secondary standards—those standards calibrated by deadweight—can achieve CMC's as low as 0.02 % of applied load if they are using several standards. Nonetheless, the downside of using several standards is that this method involves standards to be changed at least once during the calibration. Laboratories that claim CMC's of 0.01 % of applied or better may have to make three to four standard changes, or, they would need to have very expensive reference load cells and meters calibrated direct by a NMI such as N.I.S.T. These changes will add to the overall uncertainty of the force measuring instrumentation being calibrated. Standard changes take time, which often results in higher deviations

How Good Does Your Calibration Provider Have to Be? (T.U.R. Table)								
Calibration Standard Required			Tolerance Required					
			0.010%	0.020%	0.050%	0.100%	0.200%	0.500%
Deadweight	Calibration Lab Capability (CMC)	0.002%	4.329	8.657	21.644	43.287	86.575	216.437
Deadweight		0.005%	1.949	3.897	9.743	19.486	38.972	97.429
Deadweight/Lever		0.010%	0.993	1.987	4.967	9.934	19.868	49.669
High End Load Cell		0.020%	0.499	0.998	2.496	4.992	9.983	24.958
High End Load Cell		0.050%	0.200	0.400	1.000	1.999	3.999	9.997
Good Load Cell		0.100%	0.100	0.200	0.500	1.000	2.000	5.000
This table is based on a Calibration Grade Load Cell with 0.01 lbf Resolution; 0.1 lbf Repeatability. Anything in Red would have too much measurement risk.								

Figure 6. T.U.R. Table

between the test points calibrated with one standard when compared to the test points using the additional standard. This additional error is directly related to timing issues and often raises the ASTM LLF, which affects the Class A loading range [3]. Therefore, if the end user wants the lowest possible loading range, it is recommended that calibration be performed using deadweight primary standards.

Furthermore, the CMC of the calibration laboratory is critical in regards to making statements of compliance. This would be whether or not an instrument is within the required tolerance. ISO/IEC 17025:2005 states "When statements of compliance are made, the uncertainty of measurement shall be taken into account" [5]. Figure 6 shows a table calculating Test Uncertainty Ratios for various CMC's and instrument tolerances. The calculation of T.U.R. involves taking the measuring device's tolerance and dividing by the expanded uncertainty [6]. The CMC discussed in this paper along with the resolution of the unit under test make up the expanded uncertainty. The repeatability of the UUT may be substituted with the repeatability calculated in the CMC for calculation of expanded uncertainty.

$$T.U.R. = \frac{Tolerance}{Expanded\ Uncertainty}$$

Many laboratories often publish their best possible CMCs on their scope of accreditation, or they might publish a reference uncertainty value such as 0.05 % of applied force as it correlates to using a secondary standard with a Class AA loading range. ASTM E74 Class AA operates on a Test Accuracy Ratio (TAR) of 5:1 to ensure that the Class AA standard is at least 5 times better than the force measuring instrument being calibrated [3]. If deadweight calibration is not possible, it is important to ask your calibration provider for the actual measurement process uncertainty, and to find out how many standard changes they will make to assure the attainment of the lowest possible CMC, which will ultimately be transferred to your equipment.

References

[1] JCGM 100:2008 E*valuation of measurement data — Guide to the expression of uncertainty in measurement*.

[2] ILAC P14:01/2013 *ILAC Policy For Uncertainty in Calibration*, http://ilac.org/publications-and-resources/ilac-policy-series/.

[3] ASTM E74-13a *Standard Practice of Calibration of Force-Measuring Instruments for Verifying the Force Indication of Testing Machine*, https://www.astm.org/Standards/E74.htm.

[4] A2LA R205 *R205 - Specific Requirements: Calibration Laboratory Accreditation Program*, American Association for Laboratory Accreditation, 2015, https://www.a2la.org/requirements/17025_calibration_req.pdf.

[5] ISO/IEC 17025:2005 *General requirements for the competence of testing and calibration laboratories*, International Organization for Standardization/International Electrotechnical Commission, 2005.

[6] ANSI/NCSLI Z540.3-2006 *Requirements for the Calibration of Measureing and Test Equipment Handbook*, 2013.

Disclaimer

Any views and opinions expressed in this paper represent those of the authors only, and not necessarily the organizations mentioned in the paper. Morehouse calibration equipment was used to conduct the testing in this study since they were easily accessible to the authors and technicians. However, any laboratory using primary standards better than 0.0016 % of applied force, calibrated by an accredited laboratory, should be able to achieve similar results.

Henry Zumbrun (hzumbrun@morehouse.com) & Alireza Zeinali, Morehouse Instrument Company, 1742 Sixth Avenue, PA 17403, www.mhforce.com.