May 2005
(Corrected June 2005)

Using Reference Scales

By Juana Williams

Weights and measures officials may sometimes find it necessary to use alternate procedures to verify the accuracy of commercial weighing and measuring devices when the use of field standard (mass) weights or graduated volumetric provers is not practical. The official follows a test procedure where the indications of performance for the device under test are compared to the indications on a reference scale. In this case, the reference scale becomes the standard. Often a reference scale is used when it becomes difficult to obtain consistent and accurate results from conventional test standards. Testing might require use of a reference scale because of the properties of a measured product (e.g., lube oil or corrosive products) or because of the complicated dynamics of equipment involved in the measurement process (e.g., moving conveyor belt). This test procedure is acceptable if care is exercised in selecting a reference scale with a suitable minimum graduation and sufficient capacity, and there is adequate testing of the reference scale's performance. This article covers what constitutes a good reference scale.

Use of a Reference Scale

What conditions warrant use of a reference scale and where are guidelines that define a reference scale? NIST Handbook 44 "Specifications, Tolerances, and Other Technical Requirements for Weighing and Measuring Devices" and NIST Handbook 133 "Checking the Net Contents of Packaged Goods" provide some guidance on the use of a reference scale.

The test notes of several specific device codes and the Fundamental Considerations for Enforcement of Handbook 44 briefly address the tolerances and acceptable time interval between verification and use of a reference scale. Based on Handbook 44 technical information, a reference scale might be defined as a weighing device or weighing system held to a specified level of accuracy that uses weight measurement to (1) determine the accuracy of a quantity indicated by a device under test or (2) establish the weight of materials for use as a test load.

Reference scales are recognized for use to verify that weighing systems with rapidly moving components such as dynamic monorail weighing systems or belt-conveyor scales are capable of making accurate weighments. For example, a railway track scale or vehicle scale might be used as a reference scale to verify that commodities such as coal are accurately weighed on a belt-conveyor scale. In this instance, the rapid movement of the conveyor belt prohibits the use of sufficient field standard weights to verify the weighing system's accuracy. Consequently, material (coal) is first weighed on the reference scale and then run across the belt-conveyor scale; the indicated weight values from both scales are compared to determine the error in the belt-conveyor scale's performance.

Gravimetric (measurement by weight) tests conducted with reference scale standards are recognized for determining the accuracy of measurements made by meters such as lube oil meters, cryogenic liquid-measuring devices, and mass flow meters. The product delivered through the meter is weighed on the reference scale and then converted to volume. The converted volume is compared to the volume indicated on the meter under test to determine the meter's error. Gravimetric test methods are used when metered products are caustic and likely to react with prover field standard metals. In this instance, the product is metered into a receiving vessel or tank designed to contain corrosive products, then the product, container, etc., are weighed on a reference scale. Test of lube oil meters with volumetric neck-type prover standards is not only time consuming, but the viscous oils make it difficult to repeat the process of draining the standard between each test draft to ensure consistent test results. Consequently, lube oils are tested gravimetrically. Officials may also encounter inspection sites with policies that prohibit commingling of products. To avoid any cross contamination of products, a single product is delivered into a receiving vessel or tank that is dedicated to the storage and/or transport of that particular product, then the vessel is weighed on a reference scale. Finally, the quantity of product weighed on the reference scale is converted to volume and that volume is compared to the volume indicated by the meter under test to determine the meter's error.

Handbook 133 Chapter 2. Basic Test Procedures – Gravimetric Testing provides extensive guidelines for selecting a reference scale. Handbook 133 outlines a step-by-step procedure for selecting a reference scale standard when gravimetric test methods are used to determine the accuracy of the net contents of packaged goods. The Handbook 133 selection process for a reference scale differs slightly from that in Handbook 44 because the reference scale's division size is based on the Maximum Allowable Variation (MAV) for the package under test. This article outlines procedures for selecting and testing a reference scale that is intended for use to test the accuracy of devices.

Selection

Location. The location of the reference scale can affect its performance and test results. A reference scale should be level and not adversely affected by environmental conditions such as wind or vibration. Problems can arise when the reference scale is not on the same premises as the device under test. Some issues that occur because of long distances between the device under test and the reference scale are changes in the tare weight of vehicles transporting the receiving vessel due to fuel consumption or mud and other substances from the roadway adhering to truck tires. In unattended remote locations the conveyor belt for a weighing system may be set up to move materials for many miles. This means that any spillage of materials used as the test load and verified on the reference scale might go unnoticed.

Capacity. Reference scales should have sufficient capacity to test the device over its range of rated minimum and maximum capacity and all possible test draft sizes to include the weight of any receiving tanks, provers, or vessels. The platform (load-receiving element) size should accommodate the dimensions of weighed materials and any container (cart, vehicle, etc.) that is part of the tare weight. A reference scale should not be used if the net load is less than that specified in Handbook 44 2.20 Scales Table 8 Recommended Minimum Load since the use of the reference scale to weigh light loads is likely to result in relatively large errors. More detailed information on the relationship of the minimum load to scale errors follows in the discussion on Scale Division.

Scale Division. Steps should be taken to eliminate the uncertainty associated with reading weight values on the reference scale.

The size of the scale division (*d*) for the reference scale relative to the size of the net load significantly affects the accuracy to which a device can be tested. The size of the *d* for the reference scale also affects the test draft size required to adequately evaluate the performance of the device under test and the rounding error associated with reading the reference scale's indications to the nearest division.

The reference scale normally rounds to the nearest scale division, which introduces a potential error of one-half *d* for each weight determination. Using a scale with a higher resolution, error weights to increase the readability of weight values, or the use of a larger test draft can reduce the rounding error.

These principles can be applied to a real world device such as a mass flow meter. The device under test is a retail motor-fuel dispenser (RMFD) that uses a mass flow meter to meter compressed natural gas (CNG) into a vehicle. The RMFD indicates in mass units of pounds (mass indications are required for official test even though the RMFD indicates in units of gasoline gallon equivalents [GGE]). To test the performance of the RMFD a cylinder tank is used to simulate various stages of delivery into a vehicle. The tank, along with the metered CNG, is weighed on the reference scale to determine the RMFD's error (the weight of the empty tank is also established to determine its tare value).

The size of the division for the reference scale should be one-tenth of the tolerance applied to the device under test (when corrections are not made to errors in the reference scale reading) to reduce rounding errors when reading the results to the nearest scale division. This also ensures that the cumulative errors that can occur when reading the reference scale indications, along with other factors that contribute to uncertainty in the reference scale's performance, do not use up the entire tolerance allowed for the test standard. The error in a test standard used without correction should not be greater than one-third of the smallest tolerance applied to the device under test as specified in NIST Handbook 44 Appendix A, Fundamental Considerations, Section 3.2 Tolerances for Standards.

Consider an example in which the acceptance tolerance of ± 1.5 % for an Accuracy Class 2.0 mass flow meter (MFM) application (CNG) applies and the combined weight of the empty tank and the metered CNG is 22 kg, where the tank weight is 20 kg and CNG product weight is 2 kg. The tolerance is applied to the smallest net load indicated on the device under test (MFM), which in this case is 2 kg. The scale division should be no greater than one-tenth of that amount and is calculated as follows:

Reference scale division (*d*) ≤ Smallest Test Load x Tolerance for the device under test x 1/10
≤ 2 kg x 0.015 x 0.1
≤ 0.003 kg

Thus, the scale division for the reference scale should be no greater than 0.003 kg or 3 g. Since the Scales Code of NIST Handbook 44 requires that the value of a scale division (*d*) be expressed in units of 1, 2, or 5, the reference scale division in this example must be no greater than 2 g.

Consider a second example in which the device under test indicates in inch-pound units and the combined weight of the empty tank and the metered CNG is 50 lb, where the tank weight is 45 lb and the CNG product weight is 5 lb. Using the same formula as above, the maximum scale division for the reference scale is calculated as follows:

Reference scale division (*d*) ≤ Smallest Test Load x Tolerance for the device under test x 1/10
≤ 5 lb x 0.015 x 0.1
≤ 0.0075 lb

Thus, the scale division for the reference scale should be no greater than 0.0075 lb. Since the Scales Code of NIST Handbook 44 requires that the value of a scale division (*d*) be expressed in units of 1, 2, or 5, the reference scale division must be no greater than 0.005 lb.

A combination of approaches such as using a scale with a higher resolution, error weights to increase the readability of weight values, or increasing the test draft size might be used to reduce rounding errors. Each scenario must be evaluated on a case-by-case basis to ensure that you have selected the right approach or combination of approaches. For example, an application in which you are using a vehicle scale as the reference scale will have completely different influences and considerations than an application where a CNG MFM is the device under test. Error weights can be used to reduce the error. In some cases error weights as small as 0.1 *d* might be used; in other cases influences such as environmental factors and scale stability may necessitate the use of a larger denomination error weight (for example, 0.25 *d* of the reference scale division).

Consider an example in which a device under test is to be tested using a reference scale with a division size of 5 g or 0.01 lb. If environmental conditions permit, error weights could be used to increase the readability to the nearest 0.5 g or 0.001 lb, respectively, for the gross and tare weights. Each weight value is thus ± 0.25 g or ± 0.0005 lb, respectively, reading to the nearest 0.5 g or 0.001 lb, but since there are two weighings, gross and tare, the potential for total rounding error is 0.5 g or 0.001 lb.

Next, the test draft size should be determined. To limit the error for each weighing to one-tenth of the tolerance, the minimum test draft size must be no less than:

$$\frac{\text{Readability of scale using error weights (kg or lb) x 10}}{\text{Tolerance for device under test}} = \text{Minimum test draft size (kg or lb)}$$

Substituting the information gathered in the previous example into the equation and assuming a tolerance of ± 1.5 %, the size of the minimum test drafts are calculated as follows:

$$\frac{0.5\ g \text{ x } 10}{0.015} = 333.33 \text{ g or } 0.333 \text{ kg} \qquad \text{OR} \qquad \frac{0.001 \text{ lb x } 10}{0.015} = 0.667 \text{ lb}$$

It is essential to consider the minimum net load weighed on the reference scale. The minimum load requirements in Handbook 44 Table 8 are based on the tolerances that apply to the device under test and the error associated with rounding weight values. The errors associated with rounding weight indications on the reference scale can result in a potential error of ± 0.5 *d* for each weight determination since the reference scale normally rounds weight values to the nearest scale division. One way to determine the maximum potential error that results from rounding error is to use the formula:

$$\frac{\text{Potential Rounding Error } (d) \text{ x } 100}{\text{Applied Load } (d)} = \text{Potential Error as a Percentage (\%) of the Applied Load}$$

To illustrate the importance of examining if the reference scale is suitable to weigh the minimum test draft, the formula is used applying the Table 8 recommended minimum net load of 20 *d* on a Class III reference scale as shown below:

$$\frac{0.5\ d \text{ x } 100}{20\ d} = \text{A potential rounding error that is 2.5 \% of the net load}$$

However, if a test draft that represents 10 *d* were metered through the RMFD, the potential for rounding error increases as follows:

$$\frac{0.5\ d \text{ x } 100}{10\ d} = \text{A potential rounding error that is 5 \% of the net load}$$

Test Standard Error. Handbook 44 Appendix A Fundamental Considerations Section 3.2. Tolerances for Standards specifies that the error in a standard should be known and corrected, or when used without correction, its error should not be greater than one-third the tolerance applied to the device under test. This rule keeps to a minimum the proportion of the tolerance allowed for the device under test from being used up because of errors of the standard (reference scale). For example, when Handbook 44 requires applying an acceptance tolerance of ± 1.5 % to a device, then the reference scale must be accurate to at least 0.5 %.

Safety. Safety is an important consideration in the selection and use (adequate support, etc.) of a reference scale at an inspection site. The reference scale should be intrinsically safe for use in its operating environment.

Testing

Test Procedures. Typically, test procedures for various types of scales that can be selected for use as reference scales are specified in NIST Handbook 112 "Examination Procedures Outlines (EPO) for Commercial Weighing and Measuring Devices." A reference scale should be tested as specified in the EPO for that device type to include the increasing-, shift- or section-, repeatability- and decreasing-load tests. Additional tests should be conducted at the approximate empty (tare) weight of the receiving vessel and the anticipated weight of the receiving vessel when filled with product.

Test Method. There will always be uncertainty in the process of selecting, testing, and using a reference scale. What is important is that those uncertainties be identified and minimized or eliminated to the extent possible. NIST IR 6919 *Recommended Guide for Determining and Reporting Uncertainties for Balances and Scales* (available on the NIST WMD web-site at http://ts.nist.gov/ts/htdocs/230/235/NISTIR6919.pdf) includes a comprehensive description of how to determine and address uncertainties in balances and scales. As noted in NIST IR 6919, there are many possible sources of uncertainty in a balance or scale calibration. While there are other things that can also contribute to uncertainty, among the most common uncertainty contributors are the:

- uncertainty or tolerance of the applied load;
- repeatability of the weighing system;
- readability of the weighing system;
- reproducibility of the weighing system; and
- effects of temperature changes, drafts, or wind, or off-center loading, indicator drift, electrical noise and variation, and vibration.

Page 15 of NIST IR 6919 includes an example of how a cause and effect diagram can be used to help you identify factors that contribute to uncertainties in the use and application of a particular weighing device. This diagram illustrates influences contributing to scale uncertainty through the scale design, installation, staff, and procedures, facility, standards, and method of use. A similar diagram might be used to help you identify uncertainties, associated with a particular reference scale.

Several points should be considered to avoid contributing to the uncertainty in the reference scale's performance. The stability of materials used as the test load to determine the reference scale's accuracy must be examined, especially if the material is affected by factors such as moisture. In instances where sufficient field standard test weights are not available, the substitution and/or strain test load method may be used to test the reference scale to the maximum capacity of the device under test. A small weight kit may be necessary for use as error weights or to test the sensitivity of some reference scales. Any errors in the reference scale's indications at a particular test load should be noted for later use to correct the weighment results for that same test load on the device under test. While the device is under test, the test load should be centered on the reference scale platform and be placed in the same position to ensure additional errors are not introduced into the process.

Requirements. Reference scales must meet the accuracy and performance criteria in Handbook 44 although they are not required to be NTEP devices. Field standards used to test the reference scale should have acceptable accuracy and traceability to national or international standards.

Test Interval. A reference scale should be tested prior to use. A reference scale should be adjusted as close to zero error as practicable. Several Handbook 44 codes provide guidelines on the appropriate time interval between test and use of the reference scale. A retest of the reference scale is always recommended, and may be warranted, if at any point during the verification process for the device under test, the reference scale exhibits any erratic performance or if conditions of use (traffic, other applications, etc.) result in questions about the performance of the reference scale.

If you have any questions about these procedures, please contact Juana Williams at 301-975-3989 or by email at juana.williams@nist.gov.