MSL Technical Guide 7

Calibrating Standard Weights

Introduction

Reliable weighing depends on the use of standard weights of known mass value. Standard weights are used to calibrate the balances used for weighing, to check these balances between calibrations, and as references for weighing by comparison.

This technical guide describes the calibration of standard weights and the evaluation of the associated uncertainties when the calibration uncertainty is at best 1 part in 10^6. Hence this guide applies to the calibration of weights in OIML (International Organisation of Legal Metrology) International Recommendation R 111-1 accuracy classes F_1, F_2 and lower accuracy classes [1].

The scope of this guide is limited to the calibration of weights by comparison with reference weights having the same nominal mass value, and using direct-reading electronic balances or mass comparators. For other types of weight calibration, refer to the bibliography at the end of the guide.

Standard Weights

Standard weights have a mass value close to a nominal value, such as 200 g or 1 kg, and are made to comply with accepted specifications to ensure that they are suitable for their intended purpose. The most common specification for standard weights is OIML R 111-1 [1], which defines a hierarchy of accuracy classes E_1, E_2, F_1, F_2, M_1, $M_{1\text{-}2}$, M_2, $M_{2\text{-}3}$ and M_3. Class E_1 is the highest accuracy class and each successive class is about three times less accurate. For example, the mass of a class F_1 kilogram must be within 5 mg of 1 kg while an F_2 kilogram must be within 16 mg of 1 kg.

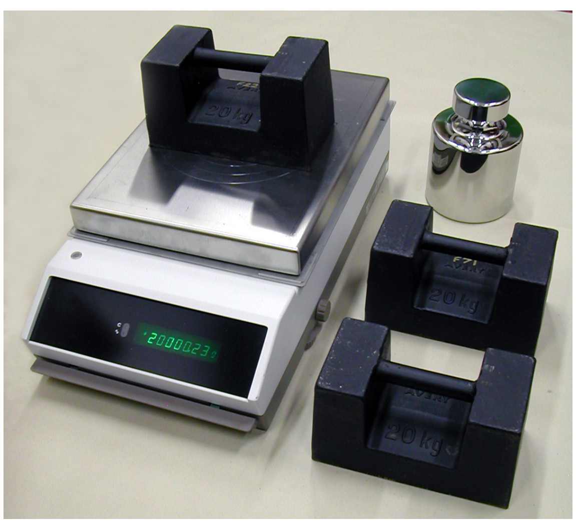

Figure 1. Calibrating class M_1 20 kg weights.

The standard weights considered here are usually calibrated in terms of a so-called conventional mass value – the mass of a weight of density 8000 kg/m^3 that, in air of density 1.2 kg/m^3 at 20 °C, would balance the weight being calibrated. For each accuracy class, R 111-1 specifies how close the density of the standard weight must be to 8000 kg/m^3 to avoid significant errors due to air buoyancy.

R 111-1 also defines other parameters that influence the accuracy and stability of the mass value such as shape, construction, magnetism, material, and surface finish. For example, an F_1 1 kg weight is likely to be cylindrical with a smoothly shaped knob on top and made out of stainless steel, while an M_1 20 kg weight is likely to be rectangular in section with a built-in handle and made of cast iron (see Figure 1).

Model

The comparison between two weights is described by the equation

$$M_{\mathrm{T}} = M_{\mathrm{R}} + \Delta M_{\mathrm{T-R}}\,, \qquad (1)$$

where M_{T} is the conventional mass value of the weight being calibrated and M_{R} is the known conventional mass value of a reference weight. The difference $\Delta M_{\mathrm{T\text{-}R}}$ between these two mass values can be determined from the difference between balance readings for the weight under calibration and for the reference weight.

Weighing Methods

Three weighing methods are suggested. Method 1 is suitable for calibrating weights when the balance repeatability is not the dominant uncertainty or when it is small compared with the desired weighing uncertainty. Otherwise Method 2 is more appropriate. Method 3 is an advanced method developed by MSL that may be used as an alternative to Method 2. More detail is given below on the choice of weighing method.

Before using any of these methods, make sure that the weights and the balance are ready. Check that the weights are clean. If you need to do more than brush them or wipe them with a soft cloth, contact the client and check if they need to know the mass of the weights before cleaning. Weights that need cleaning can be washed or wiped with clean alcohol, distilled water or other solvents. Avoid immersing weights with internal cavities. After cleaning, leave F_2 to M_3 class weights for an hour and F_1 class weights for 1-2 days (see B.4 in [1]). Make sure the weights have had sufficient time to reach the temperature of the laboratory (see Thermal stability below).

Check that the balance has had sufficient time to warm up, is protected from air currents, has zerotracking turned off and has recently had its scale factor adjusted (via the 'CAL' function or equivalent).

The recommended conditions for calibration of a weight are given in Annex C of [1]. For example, for class F_1 weights, the temperature should change by less than 1.5 °C per hour and the relative humidity should be in the range 40 % to 60 % with a maximum change of 15 % per four hours.

Method 1

Compare unknown weight T with reference weight R using a single so-called $ABBA$ weighing, which in this case is $RTTR$, as follows:

- Place weight R on the balance pan and record the balance reading (denoted by r_1) when the balance indication has stabilised.
- Remove R and replace it with weight T and record the balance reading t_1 when the balance indication has stabilised.
- Remove T and return it to the balance pan as though it were a different weight, record balance reading t_2.
- Remove T and replace it with R, record balance reading r_2.

Try to keep the same time interval between each of the four loadings. From the balance readings, calculate

$$\Delta M_{\text{T-R}} = 0.5\left(t_1 + t_2 - r_1 - r_2\right) . \qquad (2)$$

Use this value of $\Delta M_{\text{T-R}}$ in (1) to calculate M_{T}. The $ABBA$ method compensates for drift in the balance zero so you do not need to tare the balance before loading each weight.

With Method 1, several weights with the same nominal mass value can be calibrated in the same weighing, using the sequence $RT_1 \ldots T_J T_J \ldots T_1 R$. For each weight, the analysis follows (2). The number J of test weights should not be more than 5.

Method 2

Compare the unknown weight T with the reference weight R using n $RTTR$ weighings. Each $RTTR$ weighing follows Method 1 and n is 5 or more (to give sufficient degrees of freedom). Calculate a value of $\Delta M_{\text{T-R}}$ for each weighing using (2) and the mean value $\overline{\Delta M_{\text{T-R}}}$ using

$$\overline{\Delta M_{\text{T-R}}} = \sum_{i=1}^{n} \Delta M_i \Big/ n , \qquad (3)$$

where ΔM_i is the value of $\Delta M_{\text{T-R}}$ for the i^{th} $RTTR$ weighing. Use $\overline{\Delta M_{\text{T-R}}}$ in (1) to calculate M_{T}.

As with Method 1, several weights with the same nominal mass value can be calibrated in the same weighing using five or more $RT_1 \ldots T_J T_J \ldots T_1 R$ weighings, although this is not recommended if the balance zero drifts significantly.

Method 3: Alternative to Method 2

MSL has developed a circular weighing approach [2] as an alternative to ABBA weighings for Method 2. Like ABBA weighings, the circular weighing approach is insensitive to drift in the balance reading but it has the advantage of giving drift and repeatability values that can be used to assess the quality of each circular weighing. Furthermore, it is easily extended to simultaneous calibration of up to four weights of the same nominal value.

A circular weighing for two weights is $(AB)_5$, that is, weight A is loaded then weight B, repeated until each has been loaded five times. This weighing has about $\nu = 7$ degrees of freedom. For more weights, the circular weighing sequences are $(ABC)_4$, $(ABCD)_3$ and $(ABCDE)_3$ (with about 8, 7, and 9 degrees of freedom respectively). The disadvantage of circular weighings is the complexity of the analysis although this is easily managed in software such as Excel.

Contact MSL if you would like advice on using the circular weighing approach.

Sources of Uncertainty

The sources of uncertainty associated with calibrating a standard weight are discussed below. The main sources of uncertainty are discussed first, followed by the sources of uncertainty that can usually be made insignificant.

Uncertainties are evaluated following [3], which gives guidelines for calculating a standard uncertainty for each uncertainty component and for combining these standard uncertainties.

Repeatability of the Balance

Repeatability is a measure of the random variations in a balance reading and is the only Type A contribution to the measurement uncertainty. Repeatability measurements are usually carried out six-monthly as part of the in-service checks on the balance (see [4]). The standard uncertainty u_{R} due to repeatability is normally evaluated as the sample standard deviation of the balance readings for p successive loadings of the same weight. That is

$$u_{\text{R}} = \sqrt{\sum_{i=1}^{p} \left(r_i - \overline{r}\right)^2 \Big/ \left(n-1\right)} , \qquad (4)$$

where r_i is the i^{th} balance reading, and $\overline{r}$ is the average of the p balance readings. When $p = 10$, u_{R} has $\nu = 9$ degrees of freedom. Since u_{R} may vary with load, it is advisable to measure it at each of the different loads used for calibrating weights. The uncertainty in a weighing due to repeatability depends on the weighing method.

For Method 1, the standard uncertainty in $\Delta M_{\text{T-R}}$ due to repeatability is

$$u_{\text{R}\Delta} = 0.5\sqrt{u_{\text{R}}^2 + u_{\text{R}}^2 + u_{\text{R}}^2 + u_{\text{R}}^2} = u_{\text{R}} . \qquad (5)$$

A historical value of u_{R} may be used (such as from the most recent check on the balance) although it is preferable to re-measure u_{R} nearer to the time of the weighings. When a historical value is used for u_{R}, it is desirable to check that the balance performance has not degraded since u_{R} was measured. Two estimates of $\Delta M_{\text{T-R}}$ from the $ABBA$ weighing, $t_1 - r_1$ and $t_2 - r_2$, can be used for this purpose. An acceptance criterion is

$$\left| \left(t_1 - r_1\right) - \left(t_2 - r_2\right) \right| < 4\, u_{\text{R}} . \qquad (6)$$

That is, the magnitude of the difference between the two estimates is less than $4u_{\text{R}}$. This criterion is based on an

F-test at a 2 % level of significance and assumes that the zero drift of the balance is small ($t_1 - r_1$ and $t_2 - r_2$ are sensitive to drift in the balance zero). A 2 % level of significance is used to give more latitude in the performance of the balance because Method 1 is used when u_R is not the dominant uncertainty. For the normal 5 % level of significance, the limit in (6) would be $3.2u_R$.

For Method 2, the standard uncertainty in $\overline{\Delta M}_{T-R}$ due to repeatability is calculated from the variability in the weighing results and is given by

$$u_{\overline{R\Delta}} = u_{R\Delta} \big/ \sqrt{n} \ , \qquad (7)$$

where

$$u_{R\Delta} = \sqrt{\sum_{i=1}^{n} (\Delta M_i - \overline{\Delta M}_{T-R})^2 \Big/ (n-1)} \ . \qquad (8)$$

The value obtained for $u_{\overline{R\Delta}}$ should be checked against a historical reference value to detect any degradation in weighing performance. A pooled variance of previous weighings is used to give a reference value $u_{\overline{R\Delta}(\mathrm{Ref})}$ (see C.6.1.4 in [1]). An acceptance criterion is

$$u_{\overline{R\Delta}} < 2u_{\overline{R\Delta}(\mathrm{Ref})} \ . \qquad (9)$$

This criterion is an F-test at a 5 % level of significance and assumes that $u_{\overline{R\Delta}}$ is calculated from five $RTTR$ weighings ($\nu = 4$) and that $u_{\overline{R\Delta}(\mathrm{Ref})}$ is calculated from two sets of five $RTTR$ weighings ($\nu = 8$).

For $RT_1 \ldots T_J T_J \ldots T_1 R$ weighings, the results for each of the J weights are pooled to calculate the standard uncertainty $u_{\overline{R\Delta}}$ due to repeatability in each mean $\overline{\Delta M}_{T-R}$ (see C.6.1.4 in [1]).

For the circular weighing method (Method 3), the standard uncertainty due to repeatability in each ΔM_{T-R} is evaluated as part of the analysis.

Resolution of the Balance

The standard uncertainty in a balance reading due to the resolution δr is $u_{RS} = \delta r/\sqrt{12} = 0.29\ \delta r$, with infinite degrees of freedom (see F.2.2.1 in [3]). In general we consider errors due to resolution to be fully correlated. Hence the standard uncertainty due to resolution $u_{RS\Delta}$ associated with an $RTTR$ weighing is

$$u_{RS\Delta} = 0.5\ (u_{RS} + u_{RS} + u_{RS} + u_{RS}) = \delta r/\sqrt{3}, \qquad (10)$$

and this uncertainty is not usually reduced by repeat $RTTR$ weighings.

Reference Weight Mass Value

The standard uncertainty u_{Mcal} in the mass of the reference weight is usually obtained from the calibration certificate. If an expanded uncertainty U_{Mcal} is given, then calculate $u_{\mathrm{Mcal}} = U_{\mathrm{Mcal}}/k$. If the coverage factor k is not given than assume $k = 2.0$. Look up the value for the degrees of freedom ν corresponding to k (see for example Table G.2 in [3]). For $k = 2.0$, $\nu = 60$.

When several reference weights are used together, the standard uncertainty due to the calibration for the combination is simply the sum of the individual standard uncertainties (because their values are correlated).

Mass Instability of the Reference Weight

The mass value of the reference weight may change after the calibration, introducing some uncertainty in its value during subsequent use. Estimating this uncertainty is difficult and usually relies on the calibration history of the weight or of similar weights (see C.6.2 in [1] and 5.2.2 in [5]). Further guidance on estimating the uncertainty due to mass instability is given in [6].

MSL mass calibration reports include an estimate of the expanded uncertainty U_{Minst} and coverage factor k associated with possible changes in the mass of the weight between calibrations. From this, the standard uncertainty u_{Minst} associated with instability of the mass value of the reference weight is

$$u_{\mathrm{Minst}} = U_{\mathrm{Minst}}/k. \qquad (11)$$

When several reference weights are used together, the standard uncertainty due to instability in the mass of the combination is simply the sum of the individual standard uncertainties (because the mass instability may have a common source).

Buoyancy

If we include the effects of air buoyancy in (1), then we obtain

$$M_T = M_R + \Delta M_{T-R} + \delta M \ , \qquad (12)$$

where the air buoyancy correction δM is given to good approximation by

$$\delta M = M_T \left(\rho_a - \rho_0\right) \left[\frac{1}{\rho_T} - \frac{1}{\rho_R}\right]. \qquad (13)$$

The term ρ_a is the air density in kg/m^3 during the calibration, $\rho_0 = 1.2$ kg/m^3 is the conventional value for air density, and ρ_T and ρ_R are the densities of the weight being calibrated and the reference weight respectively.

Commonly, densities ρ_a, ρ_T, and ρ_R are not known and must be estimated to evaluate an uncertainty u_B due to neglecting δM. For most laboratory locations in New Zealand, where the altitude is less than 300 m, ρ_a on average is close to 1.2 kg/m^3 and hence the most likely value for δM is zero.

Assuming a rectangular distribution for the values of δM, we evaluate the standard uncertainty due to air bouyancy u_B as

$$u_B = \delta M_{MAX} \big/ \sqrt{3} \ , \qquad (14)$$

where δM_{MAX} is the maximum likely deviation of δM from zero.

By considering the allowed values of ρ_T and ρ_R for weights that comply with the density requirements of OIML R 111-1, we obtain

$$u_B = 0.08 MPE \ , \qquad (15)$$

where MPE is the maximum permissible error for the weight being calibrated. Equation (15) assumes that the reference weight is one class more accurate than the weight being calibrated and that the maximum deviation of ρ_a from ρ_0 is 0.05 kg/m^3. The degrees of freedom associated with u_B is ~ 50.

There are a several other sources of uncertainty, as described below, that can usually be made insignificant by design. An uncertainty component is considered insignificant if it is less than 1/3 the largest source of uncertainty.

Off-Centre Loading or Pan Position Error

The pan position error is usually measured as part of the balance calibration. From these measurements, determine how close to the centre of the pan the weights must be loaded when calibrating weights to keep the pan position error small compared with other uncertainties. For this calculation, you can assume that the pan position error is proportional to the off-centre distance and to the load. It may be desirable to mark the pan to facilitate centring of the weights. Additional detail on centring weights and scaling pan position error is given in [6].

Balance Non-Linearity

The quality of modern balances is such that balance non-linearity is unlikely to contribute significantly to the uncertainty, particularly for weights with mass values that comply with the maximum permissible errors in R 111-1.

When calibrating weights that do not comply with R 111-1, it may be necessary to measure the balance sensitivity (see 5.2 in [5]) or to use extra reference weights with the main reference weight (and/or with the weight being calibrated) to reduce the value of $\Delta M_{\text{T-R}}$ so that linearity errors are negligible.

Balance Scale Factor

Make sure that the scale factor of the balance is regularly adjusted (by using the "CAL" facility). Check how much this scale factor changes with time by periodically checking the balance reading with a weight near full load. Modify the frequency of adjustment of the balance accordingly. For most balances used to calibrate F_1 to M_3 class weights, adjustment prior to use each day should be sufficient.

Other Weight Properties

As long as the properties of the weight comply with OIML R 111-1, undesirable effects such as magnetic forces are unlikely to cause significant uncertainties in the measured mass values. Contact MSL if you are in any doubt about the quality of a weight.

Thermal Stability

Weights that are different in temperature from the balance may create convection forces that can bias the measured mass value. Table B.2 in OIML R 111-1 [1] gives recommendations on thermal equilibrium times. For example, the thermal stabilization time is up to four hours for Class F_1 weights brought into the weighing laboratory up to 5 °C hotter (or colder) than the laboratory temperature.

Combining Uncertainties

The combined standard uncertainty u_c in the measured mass value M_T of the weight being calibrated is given by

$$u_c = \sqrt{\frac{u_{R\Delta}^2}{n} + u_{RS\Delta}^2 + u_{MCal}^2 + u_{Minst}^2 + u_B^2}\ , \qquad (16)$$

where n is the number of *RTTR* weighings (one for Method 1 and typically five for Method 2). The combined standard uncertainty is multiplied by a coverage factor k to give an overall expanded uncertainty $U = ku_c$. To calculate the coverage factor for u_c, we require the effective degrees of freedom of u_c.

The effective degrees of freedom ν_{eff} of u_c is estimated using the Welch-Satterthwaite formula (see G.4 in [1]). This formula can be approximated by

$$\nu_{eff} = \nu_{R\Delta} n^2 u_c^4 / u_{R\Delta}^4 \qquad (17)$$

because, apart from $u_{R\Delta}$, the number of degrees of freedom for each uncertainty is large ($\nu \geq 50$). If $u_{R\Delta} = 0$, then use $\nu_{eff} = 50$.

For Method 1, ν_{eff} is large and $k = 2.0$. For Method 2, ν_{eff} is calculated using (17) and the corresponding k value is taken from the *t*-distribution for a 95 % level of confidence. A value for k may be obtained from Table G.2 in [3] or calculated using the Excel function TINV(0.05, ν_{eff}).

Reporting

The mass value and uncertainty for each weight that is calibrated is given to the client in a calibration report. To indicate that the report gives conventional mass values, it should include a statement like:

> *The measured mass was determined as the mass of a weight of density 8000 kg/m³ which, in air of density 1.2 kg/m³, would balance the weight calibrated.*

Assuring Calibration Quality

Checks are required to confirm the stability of the reference weights.

- The reference weights require regular recalibration and their calibration history must be monitored to confirm that the mass values are sufficiently stable.
- In the interval between calibrations, the most frequently used weights should be checked occasionally by comparison with other weights. For example, a 100 g weight can be compared with a 100 g weight from another weight set or with the combination 50 g, 20_1 g, 20_2 g and 10 g.

Choice of Weighing Method

Method 1 is suitable for calibrating weights when the balance repeatability is not the dominant uncertainty or when it is small compared with the desired weighing uncertainty.

Balance repeatability is considered dominant when $u_{R\Delta} > f \cdot u_{Rest}$ where $f \sim \sqrt{3}$ and u_{Rest} is all the other uncertainties combined, that is

$$u_{Rest} = \sqrt{u_{RS\Delta}^2 + u_{MCal}^2 + u_{Minst}^2 + u_B^2}\ . \qquad (18)$$

Method 2 (or Method 3) is preferred over Method 1 when $u_{R\Delta} > \sqrt{3} \cdot u_{Rest}$ and when repeat weighings are required to reduce the uncertainty due to repeatability. Method 2 would normally be used for calibrating weights of class F_1. Five *ABBA* weighings are usually sufficient.

If more than about nine weighings seem necessary then contact MSL for advice.

Examples

Here we give two examples; calibration of a 20 kg Class F_2 weight and calibration of a 200 g Class F_1 weight.

Example 1: Calibration of an F_2 20 kg Weight

- The balance used has a capacity of 30 kg and a resolution of 0.01 g. From (10), resolution uncertainty $u_{RS\Delta} = 0.006$ g. Measured repeatability $u_{R\Delta} = 0.03$ g.
- The reference weight used is Class F_1 with $M_R = 20000.039$ g, $u_{MCal} = 0.015$ g and $k = 2.0$ from the calibration certificate. From the history of this and similar weights, $u_{Minst} = 0.03$ g.
- Both the reference weight and the weight under calibration are known to comply with the density requirements of R 111-1. Since $MPE = 0.3$ g for an F_2 20 kg weight, using (15) gives $u_B = 0.024$ g.
- Weighing Method 1 is used because $u_{R\Delta} < \sqrt{3} \cdot u_{Rest}$ and hence $n = 1$.
- The maximum measured pan position error is 0.04 g/80 mm at 5 kg (equivalent to 2 mg/mm at 20 kg). The pan position error is kept below 0.01 g by centring the weights to within ± 5 mm (using lines drawn on the pan).

The $RTTR$ weighing result is given below.

$$
\begin{aligned}
r_1 &= 20000.02 \text{ g} \\
t_1 &= 20000.18 \text{ g} \qquad \Delta M_{T\text{-}R} = 0.18 \text{ g} \\
t_2 &= 20000.22 \text{ g} \qquad M_T = 20000.22 \text{ g} \\
r_2 &= 20000.02 \text{ g}
\end{aligned}
$$

The combined standard uncertainty in the value for M_T is given by

$$
\begin{aligned}
u_c &= \sqrt{0.03^2 + 0.006^2 + 0.015^2 + 0.03^2 + 0.024^2} \text{ g} \\
&= 0.051 \text{ g}
\end{aligned}
$$

A value for the effective degrees of freedom associated with u_c is calculated from (17) as $\nu_{eff} = 9(0.051)^4/(0.03)^4 > 50$ which gives $k = 2.0$. Hence, the overall expanded uncertainty associated with the mass of the 20 kg Class F_2 weight is $U_c = 0.10$ g. The mass of this weight is within the OIML R 111-1 maximum permissible error of 0.3 g and the expanded uncertainty associated with this weight is $\leq MPE/3$ as required by [1].

Also, the weighing passes the check (6) on the balance performance since $4u_R = 0.12$ g is greater than $|(t_1 - r_1) - (t_2 - r_2)| = 0.04$ g.

Example 2: Calibration of an F_1 200 g Weight

- The balance used has a capacity of 200 g, a resolution of 0.1 mg ($u_{RS\Delta} = 0.058$ mg for one $RTTR$ weighing), and repeatability $u_{R\Delta} = 0.15$ mg from the six-monthly checks.
- The calibration report for the class E_2 reference weight gives a mass value $M_R = 200.000024$ g. It also gives a calibration uncertainty $U_{MCal} = 0.06$ mg and mass instability uncertainty $U_{Minst} = 0.2$ mg, both with $k = 2.0$.
- Air buoyancy uncertainty $u_B = 0.08$ mg, calculated from $MPE = 1$ mg for an F_1 200 g weight.

Method 2 is used here because Method 1 will give an overall uncertainty $U_c = 0.4$ mg, which is more than one-third of the 1 mg MPE.

With five $RTTR$ weighings in Method 2, the repeatability uncertainty is reduced to $u_{\overline{R\Delta}} = 0.067$ mg, reducing the overall uncertainty to $U_c = 0.3$ mg.

References and Bibliography

[1] International Recommendation OIML R 111-1, "Weights of classes *E_1, E_2, F_1, F_2, M_1, $M_{1\text{-}2}$, M_2, $M_{2\text{-}3}$, and M_3*, Part 1: Metrological and technical requirements", Organisation Internationale de Métrologie Légale, Edition 2004 (E). Available at http://www.oiml.org/publications/.

[2] C M Sutton and M T Clarkson, "A General Approach to Comparisons in the Presence of Drift", *Metrologia*, **30**, 487–493, 1993/94.

[3] "Guide to the Expression of Uncertainty in Measurement", Geneva: International Organisation for Standardization, 1995.

[4] C M Sutton and J E Robinson, 2012, *Assuring the Quality of Weighing Results*, MSL Technical Guide 12, version 4, http://msl.irl.cri.nz/training-and-resources/technical-guides.

[5] E C Morris and K M T Fen, "The Calibration of Weights and Balances", Monograph 4: NML Technology Transfer Series, Third Edition, Sydney: CSIRO, 2003.

[6] C M Sutton, J E Robinson and G F Reid, 2012, *Calibrating Balances*, MSL Technical Guide 25, http://msl.irl.cri.nz/training-and-resources/technical-guides.

Further Information

If you want to know more about balances and weighing, contact MSL and book in for a Balances and Weighing Training Workshop. See the MSL website http://msl.irl.cri.nz/.

Version 3: C M Sutton, J E Robinson and G F Reid, October 2012.