# The evaluation of uncertainty in mass calibration: possible approaches in a comparison study

**J A Sousa**[1]**, A B Forbes**[2]**, A S Ribeiro**[3]**, P M Harris**[2]**, F Carvalho**[4]**, L Bacelar**[4]

[1]LREC, 9000-264 Funchal, Portugal
[2]NPL, Teddington TW11 0LW, UK
[3]LNEC, 1700-066 Lisboa, Portugal
[4]TAP Air Portugal, 1704-801 Lisboa, Portugal

E-mail: jasousa@lrec.pt

**Abstract.** In this paper we give the results of four methods of calculating uncertainty associated with a mass calibration problem, three based on different implementations – the first and second order law of propagation of uncertainty and the Monte Carlo method – of the general methodology described by the Guide to the Expression of Uncertainty in Measurement, the fourth based on a Bayesian formulation. The nonlinearities present in the model for the calibration problem means that the first order approach can be an unreliable method for evaluating uncertainties, relative to the other three approaches.

## 1. Introduction

The application of the GUM (Guide to the Expression of Uncertainty in Measurement) [1] uncertainty framework to the evaluation of uncertainty is accepted as adequate in most fields of metrology, including mass metrology. The methodology has, as its starting point, an input-output model describing the measurand as a function of influence quantities and the assignment of probability distributions or probability density functions (PDFs) to the influence quantities. Once these two elements are in place, the PDF associated with the measurand, the output quantity is defined. The issue is how to determine information about this PDF in a computationally convenient way.

The GUM uncertainty framework, as usually implemented,  involves two types of approximation, i) linearization, i.e., first order Taylor expansion, of the function relating the output to the inputs, and ii) an appeal to asymptotic results derived from the Central Limit Theorem to associate a Gaussian distribution to the output quantity. (The GUM also provides a methodology for associating a student *t*-distribution to the output quantity, using a scheme to determine effective degrees of freedom of the *t*-distribution. We do not consider this more general scheme in this paper.) For linear models and Gaussian inputs, the GUM uncertainty framework is exact. In all other cases, it will only provide an approximate solution. The quality of the approximation is however difficult to predict, a priori. The GUM also discusses the use of a second order Taylor expansion, which should lead to more accurate approximations.  GUM Supplement 1 [2] describes the use of a Monte Carlo method (MCM) to propagate the distributions associated with the input quantities through to that associated with the

---

[1]    To whom any correspondence should be addressed.

output quantity. The MCM does rely on approximating the function nor on asymptotic results associated with the output distribution and can be used to test the validity of the GUM approach or as the main computational tool for evaluating the uncertainty [3]. Moreover, an important advantage of MCM is that it goes beyond the evaluation of the measurement result and its associated standard uncertainty since it propagates the probability density functions instead of just the uncertainties of the input quantities. This leads to an estimate of the PDF of the output quantity, and thus any required statistic, including the measurement result, the associated standard uncertainty and coverage intervals, can be obtained from this distribution, as represented by the Monte Carlo sample. Another important advantage of MCM is its applicability regardless of the nature of the model, e.g., those incorporating strong nonlinearities. As mentioned above, the GUM uncertainty framework does in fact encompass second order methods to deal, at least partially, with nonlinearity.

The three approaches mentioned so far, the GUM first and second order and MCM, can be viewed as computational approaches that attempt to evaluate summary information about the same PDF, that associated with the output quantity, defined in terms of the input-output model and the PDFs assigned to the input quantities [4]. In a Bayesian formulation of uncertainty evaluation, the starting point is somewhat different, in that a prior PDF has also to be assigned to the measurand. However, the MCM method coincides with the Bayesian approach for a particular choice of prior and the sample produced using the MCM approach is a sample for the Bayesian posterior distribution constructed using that particular prior [5,6]. The MCM sample can be used to generate a sample from the posterior distribution constructed using another (perhaps more appropriate) prior for the measurand through the application of an extremely simple Markov chain Monte Carlo (MCMC) algorithm [7]. In this context, the GUM and MCM methods can be viewed as approximate computational approaches to Bayesian uncertainty evaluation.

## 2. The mass calibration model

See for example [3, section 9.10]. The model concerns the calibration of a weight $W$ of mass density $\rho_W$ against a reference weight $R$ of mass density $\rho_R$. It is assumed that the two masses are nominally the same. The calibration is performed using a balance operating in air of density $\rho_a$ and provides an estimate of the mass $\delta m_R$ of a small weight $\delta R$, also of density $\rho_R$, needed to achieve a balance. Taking into account buoyancy effects, the application of Archimedes' principle leads to the following model equation.

$$m_W(1 - \rho_a / \rho_W) = (m_R + \delta m_R)(1 - \rho_a / \rho_R) \tag{1}$$

involving the masses $m_W$ and $m_R$ of W and R, respectively. It is usual to work in terms of conventional masses relating to a standard density $\rho_0 = 8\ 000\ \text{kg/m}^3$ for the weight and that of $\rho_{a_0} = 1.2\ \text{kg/m}^3$ for air, so that, for example, the conventional mass $m_{W,c}$ corresponding to $m_W$ is given by

$$m_W(1 - \rho_{a_0} / \rho_W) = m_{W,c}(1 - \rho_{a_0} / \rho_0)$$

Working in terms of conventional masses, (1) becomes

$$m_{W,c}\left(\frac{1 - \rho_a / \rho_W}{1 - \rho_{a_0} / \rho_W}\right) = (m_{R,c} + \delta m_{R,c})\left(\frac{1 - \rho_a / \rho_R}{1 - \rho_{a_0} / \rho_R}\right)$$

from which we obtain (using an approximation adequate for most purposes)

$$m_{W,c} = (m_{R,c} + \delta m_{R,c})\left[1 + (\rho_a - \rho_{a_0})\left(\frac{1}{\rho_W} - \frac{1}{\rho_R}\right)\right] \tag{2}$$

The calibration problem is to determine an estimate of $m_{W,c}$, and its associated uncertainty, from knowledge of $m_{R,c}$, $\delta m_{R,c}$, $\rho_a$, $\rho_W$ and $\rho_R$. Eq. (2) gives $m_{W,c}$ as a function $m_{W,c} = f(\boldsymbol{\xi})$ of the five parameters $\boldsymbol{\xi} = (m_{R,c}, \delta m_{R,c}, \rho_a, \rho_W, \rho_R)^\mathsf{T}$. Thus, Eq. (2) provides the input-output model to enable the GUM and MCM approaches to be executed.

   In this paper, we compare the MCM against the first and second order propagation of uncertainties and also compare the MCM method with a Bayesian evaluation of uncertainty.

## 3. The different approaches used in this study for the evaluation of uncertainty

### 3.1. Uncertainty evaluation using the first and higher order GUM approach

If $\eta = f(\boldsymbol{\xi})$ is a function of n input parameters $\xi_j$, the GUM gives a general approach based on the law of propagation of uncertainty (LPU) for propagating uncertainties associated with estimates of the inputs to that associated with an estimate of the output. In its simplest form, the approach states that if each input parameter $\xi_j$ is associated a probability distribution with mean $x_j$ and standard deviation $u(x_j)$, then the output parameter $\eta$ is associated with a distribution with mean $y = f(\mathbf{x})$, $\mathbf{x} = (x_1, \ldots, x_n)^\mathsf{T}$, and standard deviation

$$u^2(y) = \sum_{j=1}^{n} c_j^2 u^2(x_j), \qquad c_j = \frac{\partial f}{\partial \xi_j}(\mathbf{x}), \tag{3}$$

assuming that the $\xi_j$ are independently distributed. The expression for the uncertainty $u^2(y)$ given in Eq. 3 is derived from a first order approximation to the function $f$. A more accurate estimate can be made by using a higher order approximation. The general approach is as follows [3][8]. If we assume that the input quantities are independently distributed, and make the further assumption that the input distributions are symmetric, there will be important simplifications to take advantage of, and the some of expansion terms will be zero. With these assumptions, a higher order estimate of $\eta$ given by

$$y + \Delta y = f(\mathbf{x}) + \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial \xi_i^2} u^2(x_i),$$

and its associated uncertainty is estimated by

$$u^2(y + \Delta y) =$$

$$\sum_{i=1}^{n} \left(\frac{\partial f}{\partial \xi_i}\right)^2 u^2(x_i) + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(\left(\frac{\partial^2 f}{\partial \xi_i \partial \xi_j}\right)^2 + \frac{\partial f}{\partial \xi_i} \frac{\partial^3 f}{\partial \xi_i \partial \xi_j^2} + \frac{\partial f}{\partial \xi_j} \frac{\partial^3 f}{\partial \xi_j \partial \xi_i^2} + \frac{1}{2} \frac{\partial^2 f}{\partial \xi_i^2} \frac{\partial^2 f}{\partial \xi_j^2}\right) u^2(x_i) u^2(x_j)$$

$$+ \sum_{i=1}^{n} \left(\left(\frac{\partial^2 f}{\partial \xi_i^2}\right) + \frac{1}{3} \frac{\partial f}{\partial \xi_i} \frac{\partial^3 f}{\partial \xi_i^3}\right) E(\Delta \xi_i^4) - \left(\frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial \xi_i^2} u^2(x_i)\right)^2. \tag{4}$$

If $\Delta \xi_j$ is associated with a normal or rectangular distribution centred at 0, then $E(\Delta \xi_j^4) = 3(u^2(x_i))^2$ and $(9/5)(u^2(x_i))^2$, respectively. The calculation also requires the evaluation of $\dfrac{\partial f}{\partial \xi_i}$, $\dfrac{\partial^2 f}{\partial \xi_i \partial \xi_i}$, and $\dfrac{\partial^3 f}{\partial \xi_i \partial \xi_j^2}$, usually the main deterrent in implementing this approach. For the mass calibration problem with the data as in Table 1, the only nonzero higher order partial derivatives that will be evaluated as nonzero (taking into account that $x_3 = \rho_{a_0}$ and $x_4 = x_5$) are

$$\frac{\partial^2 f}{\partial \xi_3 \partial \xi_4} = -(\xi_1 + \xi_2)/\xi_4^2, \qquad \text{and} \qquad \frac{\partial^2 f}{\partial \xi_3 \partial \xi_5} = (\xi_1 + \xi_2)/\xi_5^2.$$

The second order estimate of $\eta$ is the same as the first order estimate, since $\Delta y = 0$, but its associated uncertainty is given by

$$u^2(y + \Delta y) = u^2(x_1) + u^2(x_2) + \left(\frac{\partial^2 f}{\partial \xi_3 \partial \xi_4}\right)^2 u^2(x_3)u^2(x_4) + \left(\frac{\partial^2 f}{\partial \xi_3 \partial \xi_5}\right)^2 u^2(x_3)u^2(x_5) \quad (5)$$

### 3.2. Uncertainty evaluation using MCM

The Monte Carlo Method (MCM) [1][2] is a way of evaluated the uncertainty associated with the estimate $m_{W,c}$ derived from Eq. 2 without linearising approximations or, for the calculation of coverage intervals, assumptions of normality. For many problems, the method is straightforward to implement. If the input variable $\xi_j$ is associated with a distribution with probability density function $p(\xi_j)$, then for $q = 1,....M$, draws $\boldsymbol{\xi}_q = (\xi_{1,q},...,\xi_{n,q})^{\mathsf{T}}$ are made from the input distributions. Then $y_q = f(\boldsymbol{\xi}_q)$ are draws from the distribution associated with $\eta = f(\boldsymbol{\xi})$. Means, standard deviations and coverage intervals can be estimated easily from the corresponding sample statistics derived from $\{y_q, q = 1,...,M\}$.

### 3.3. Uncertainty evaluation using a Bayesian approach

The Bayesian methodology defines a posterior distribution for $m_{W,c}$ to be determined on the basis of the measurement of $\delta m_{R,c}$ and the prior information on the other parameters. The noninformative prior $p(m_{W,c}) \propto 1$ was chosen for $m_{W,c}$. The posterior distribution is related to the distribution $p(\eta)$ for $\eta = f(\boldsymbol{\xi})$ by a factor that depends on the partial derivative of $f$ with respect to $\delta m_{R,c}$. The MCMC approach follows that described in [7] and starts with the draws $y_q = f(\boldsymbol{\xi}_q)$ but also requires $d_q = \dfrac{\partial f}{\partial \delta m_{R,c}}(\boldsymbol{\xi}_q)$ to be recorded. The aim of the MCMC algorithm is to modify the MCM sample in a simple sequential accept-reject scheme involving $d_q$, so that the modified sample represents a sample from the Bayesian posterior distribution for $m_{W,c}$. Means, standard deviations, etc., can be estimated from the sample statistics as in the MCM approach.

## 4. Discussion of results

As an example calculation, we assume that the information about $m_{R,c}$ and $\delta m_{R,c}$ is taken from calibration certificates and for which it is appropriate to assign Gaussian distributions. We assume that the density information is given in terms of upper and lower limits for which it is appropriate to assign rectangular distributions. This information is summarised in Table 1. The uncertainty associated with the density estimate $\rho_W$ is perhaps unrealistically high but it is useful to be able to illustrate how nonlinearities in the model affect the uncertainty evaluations.

**Table 1.** Distributions associated with the input quantities with $m_{W,c}$.

| $\xi_j$ | Distribution | Mean | Standard uncertainty |
|---|---|---|---|
| $m_{R,c}$ | Gaussian | 100 000.000 mg | 0.050 mg |
| $\delta m_{R,c}$ | Gaussian | 1.234 mg | 0.020 mg |
| $\rho_a$ | Rectangular | 1.20 kg/m$^3$ | $(1/\sqrt{3})\times 0.10$ kg/m$^3$ |
| $\rho_W$ | Rectangular | $8.0\times 10^3$ kg/m$^3$ | $(1/\sqrt{3})\times 1.0\times 10^3$ kg/m$^3$ |
| $\rho_R$ | Rectangular | $8.00\times 10^3$ kg/m$^3$ | $(1/\sqrt{3})\times 0.05\times 10^3$ kg/m$^3$ |

For the data in Table 1, the estimate of the deviation of $m_{W,c} - m_0$ from a nominal mass of 100 g evaluated from Eq. 2 is $\delta \hat{m} = 1.234\,0$ mg.

The fact that the sensitivity coefficients associated with the density quantities are zero means that, to first order, the uncertainties associated with these quantities do not contribute to the uncertainty associated with the estimate of $m_{W,c}$ so that $u(\delta \hat{m}) = (0.050^2 + 0.020^2)^{1/2}$ mg $= 0.053\,9$ mg. However, for the second order approach, with the same example data, the measurement uncertainty yields $u(y + \Delta y) = 0.075$ mg, suggesting a clear underestimation entailed by the simpler implementation.

Compared to the second order approach (and even the GUM first order approach), the MCM method is straightforward to implement. Fig. 1 shows the normalised histogram of 100,000 MCM samples generated using the input distributions specified in Tab. 1 and the input-output model given in Eq. (2). These samples were used, along with the partial derivatives $d_q = \dfrac{\partial f}{\partial \delta m_{R,c}}(\xi_q)$, as the input to the MCMC algorithm to produce a sample from the Bayesian posterior distribution for $m_{W,c}$ using a noninformative prior $p(m_{W,c}) \propto 1$. The degree to which the MCMC algorithm modifies the MCM sample depends on the extent that $d_q$ varies. For the mass calibration example, these partial derivatives vary by less than 5 parts in $10^6$, which means that the MCMC sample is essentially the same as the MCM sample. In fact, all of the 100,000 of the MCM samples were accepted so that the two samples are identical. For larger MCM samples, some would be rejected but would not introduce any meaningful differences. The good agreement between the two approaches is due to the fact that the dependence of the measurement equation on $\delta m_{R,c}$ is almost linear. For such cases, the MCM approach is an extremely effective approach for sampling from the Bayesian posterior distribution [6,7].

The results of the four methods of evaluating an estimate of $m_{W,c} - m_0$ are given in Table 2. The first three methods all attempt to provide summary information about the same probability distribution, namely the output distribution associated with Eq. 2. The differences are in the level of approximation, in the case of the GUM methods, a first or second order approximation, in the case of MCM, the approximation of a continuous distribution by a discrete sample. As the number of MCM trials increases, the discrete approximation can becomes more exact. Due to nonlinearities in the model, the first order GUM method significantly underestimates the standard deviation of the output distribution. (For a smaller uncertainty associated with the density quantity $\rho_W$, the 1st order GUM gives a more accurate estimate of the uncertainty.) The GUM 2nd order method gives satisfactory estimates and can be implemented for the mass calibration problem without much complicated calculation. For a general model, there can be a considerable (or prohibitive) amount of work involved.

As indicated above, the MCM and MCMC approaches give identical results. Fig. 1 also shows the Gaussian distributions determined by the GUM 1st and 2nd order methods. It is noted that the GUM 2nd order distribution looks a very good representation of the frequency distribution derived from the Monte Carlo samples. This can be explained by the fact that the dominant influence quantities $m_{R,c}$ and $\delta m_{R,c}$ are associated with Gaussian distributions and these quantities appear linearly in the measurement equation, Eq. 2.
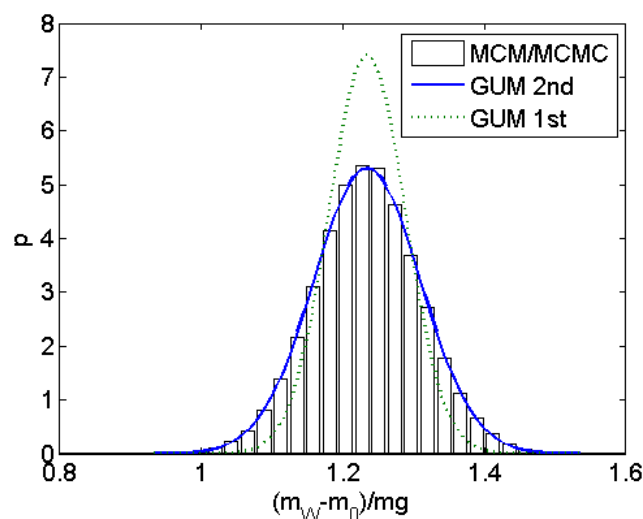


Figure 1. Frequency histogram for MCM compared with the Gaussian distributions determined by the GUM 1st and 2nd order approaches.

Table 2. Parameter estimates and associated uncertainties associated with the mass calibration problem for four methods of evaluation.

| Method | $\delta\hat{m}$ /mg | $u(\delta\hat{m})$ /mg |
|---|---|---|
| GUM(1st order) | 1.234 | 0.054 |
| MCM | 1.234 | 0.075 |
| GUM(2nd order) | 1.234 | 0.075 |
| Bayes | 1.234 | 0.075 |

## 5. Conclusions

Four computational approaches have been applied to evaluating the uncertainty associated with a mass comparison. Three, the GUM 1$^{st}$ order, GUM 2$^{nd}$ order and the Monte Carlo Method (MCM) all involve the same distribution determined by the measurement equation and the assignment of distributions to the influence quantities. The differences between them therefore reflect the degrees of approximation implicit in the computational approaches. The GUM 1$^{st}$ and 2$^{nd}$ order approaches involve truncation approximations and assign Gaussians of the basis of the estimated first and second moments of the distribution. The MCM method is approximate in the sense that only a finite sample is constructed. The MCM method and GUM 2$^{nd}$ order method are in close agreement for the calculated example.

The fourth method involved a Bayesian formulation of the mass comparison measurement and necessarily involves the assignment of a prior distribution for the quantity of interest. If a noninformative prior $p(m_{W,c}) \propto 1$ is chosen, the Bayesian posterior distribution is essentially the same distribution from which the MCM determines a sample. This means that the MCM sample can also be regarded as a sample from the Bayesian posterior corresponding to the assigned prior, as confirmed by the application of the Markov chain Monte Carlo (MCMC) scheme. Thus, three approaches give essentially the same results: GUM 2$^{nd}$ order, MCM and the Bayesian approach. The GUM 1$^{st}$ order approach is not recommended for the mass comparison measurement problem. Of the three approaches, the MCM method is the easiest to implement. The MCMC scheme implementing a Bayesian approach requires a modest extension of the MCM approach.

## 6. References

[1]     JCGM 100:2008 Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM 1995 with minor corrections, Joint Committee for Guides in Metrology).

[2]     JCGM 101:2008 Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method (Joint Committee for Guides in Metrology).

[3]     Cox M G and Harris P M 2006 Software Support for Metrology Best Practice Guide N° 6, Uncertainty Evaluation. NPL Report DEM-ES-011. National Physical Laboratory, UK.

[4]     Forbes A B 2012 Approaches to evaluating measurement uncertainty, *International Journal of Metrology and Quality Engineering*, **3**:71-77.

[5]     Elster C, Wöger W and Cox M G 2007 Draft GUM supplement 1 and Bayesian analysis, *Metrologia*, **44**:L31:L32.

[6]     Forbes A B and Sousa J A 2011 The GUM, Bayesian inference and the observation and measurement equations, *Measurement*, **44**:1422-1435.

[7]     Forbes A B 2012 An MCMC algorithm based on GUM Supplement 1 for uncertainty evaluation, *Measurement*, **45**:1188-1199.

[8]     Nielsen, L 2002 Private communication.