IMEKO 20th TC3, 3rd TC16 and 1st TC22 International Conference
Cultivating metrological knowledge
27th to 30th November, 2007. Merida, Mexico.

The relevance of using a monte carlo method to evaluate UNCERTAINTY in mass calibration

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

[1] Laboratório Regional de Engenharia Civil, 9000-264 Funchal, Madeira, Portugal
[2] Laboratório Nacional de Engenharia Civil, 1700-066 Lisboa, Portugal
[3] National Physical Laboratory, Teddington TW11 0LW, UK
[4] TAP Air Portugal - Laboratório de Calibrações, 1704-801 Lisboa, Portugal

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 generally methodology described by the Guide to the Expression of Uncertainty in Measurement, the fourth based on a Bayesian formulation. Nonlinearities 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.

Keywords: modelling, uncertainty, Monte Carlo, mass calibration.

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.

Problems can arise because the GUM uncertainty framework involves two types of approximation, i) linearisation 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. 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 and cases exist where a model *a priori* not particularly suited to the GUM methodology produces very acceptable results, which apparently depends on the relative order of magnitude of the different sources of uncertainty and its probability density function (PDF), which in turn reinforces the relevance and applicability of the MCM as a validation tool in a vast number of problems in metrology [2, 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.

Another important advantage of MCM is its applicability regardless of the nature of the model, e.g., those incorporating strong nonlinearities. The GUM uncertainty framework does in fact encompass higher order methods to deal with nonlinearity. 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.

2. The mass calibration model

See for example [3, section 9.10]. The model concerns the calibration of a weight W of mass density ρ_W against a reference weight R of mass density ρ_R. It is assumed that the two masses are nominally the same. The calibration is performed using a balance operating in air of density ρ_a and provides an estimate of the mass δm_R of a small weight δR, also of density ρ_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 ρ_0 = 8 000 kg/m^3 for the weight and that of ρ_{a_0} = 1.2 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 $m_{R,c}$, $\delta m_{R,c}$, ρ_a, ρ_W and ρ_R. Eq (2) gives $m_{W,c}$ as a function $m_{W,c} = f(\boldsymbol{\xi})$ of the five parameters $\boldsymbol{\xi} = (m_{R,c}, \delta m_{R,c}, \rho_a, \rho_W, \rho_R)^{\mathrm{T}}$.

3. Uncertainty evaluation using the first order GUM approach

If $\eta = f(\boldsymbol{\xi})$ is a function of n input parameters ξ_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 ξ_j is associated a probability distribution with mean x_j and standard deviation $u(x_j)$, then the output parameter η is associated with a distribution with mean $y = f(\mathbf{x})$, $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$, 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 ξ_j are independently distributed. The terms c_j are known as the sensitivity coefficients. For the purposes of calculating coverage intervals, the GUM makes the further assumption that the distribution associated with η is approximated by a Gaussian (normal) distribution, i.e., $\eta \sim N(y, u^2(y))$, so that a 95 % coverage interval is given by $[y - 2u(y), y + 2u(y)]$.

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 ρ_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 associated with $m_{W,c}$.

ξ_j	Distribution	Mean	Standard deviation
$m_{R,c}$	Gaussian	100 000.000 mg	0.050 mg
$\delta m_{R,c}$	Gaussian	1.234 mg	0.020 mg
ρ_a	Rectangular	1.20 kg/m^3	0.10 kg/m^3
ρ_W	Rectangular	8.0×10^3 kg/m^3	$(1/\sqrt{3}) \times 1.0 \times 10^3$ kg/m^3
ρ_R	Rectangular	8.00×10^3 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.

From Eq. (2), the sensitivity of $m_{W,c}$ with respect to the five parameters can be calculated and are given in Table 2. 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.

Table 2.

Partial derivatives and sensitivity coefficients associated with $m_{W,c}$ corresponding to the data in Table 1.

ξ_j	$\partial f / \partial \xi_j$	c_j
$m_{R,c}$	$1 + (\xi_3 - \rho_{a_0})(1/\xi_4 - 1/\xi_5)$	1
$\delta m_{R,c}$	$1 + (\xi_3 - \rho_{a_0})(1/\xi_4 - 1/\xi_5)$	1
ρ_a	$(\xi_1 + \xi_2)(1/\xi_4 - 1/\xi_5)$	0
ρ_W	$-(\xi_1 + \xi_2)(\xi_3 - \rho_{a_0})/\xi_4^2$	0
ρ_R	$(\xi_1 + \xi_2)(\xi_3 - \rho_{a_0})/\xi_5^2$	0

4. Uncertainty evaluation using MCM

The Monte Carlo Method (MCM) [2,3] is a way of evaluating 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 ξ_j is associated with a distribution with probability density function $p(\xi_j)$, then for $q = 1,\ldots M$, draws $\boldsymbol{\xi}_q = (\xi_{1,q},\ldots,\xi_{n,q})^{\mathrm{T}}$ 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 converage intervals can be calculated easily from the corresponding sample statistics derived from $\{y_q, q = 1,\ldots,M\}$. Using this approach, the MCM estimate of $m_{W,c} - m_0$ based on 10^6 trials is 1.234 1 mg and the corresponding uncertainty is 0.075 4 mg. This uncertainty estimate does reflect the uncertainty associated with the density parameters.

5. Uncertainty evaluation using higher order GUM

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 [3,4]. We write $\eta = y + \Delta\eta = f(\mathbf{x} + \Delta\boldsymbol{\xi})$ where $\Delta\boldsymbol{\xi}$ is associated with a distribution with mean $E(\Delta\boldsymbol{\xi}) = 0$ and variance matrix

$$U = \begin{pmatrix} u^2(x_1) & u(x_1,x_2) & \cdots & u(x_1,x_n) \\ u(x_2,x_1) & u^2(x_2) & \cdots & u(x_2,x_n) \\ \vdots & \vdots & \ddots & \vdots \\ u(x_n,x_1) & u(x_n,x_2) & \cdots & u^2(x_n) \end{pmatrix}$$

giving the variances and covariances

$u^2(x_i) = u(x_i, x_i) = E(\Delta\xi_i^2)$, $u(x_i, x_j) = E(\Delta\xi_i \Delta\xi_j)$,

respectively. To a second order approximation $\Delta\eta$ is given by

$$\Delta\eta = \sum_{i=1}^{n} \frac{\partial f}{\partial \xi_i} \Delta\xi_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial \xi_i \partial \xi_j} \Delta\xi_i \Delta\xi_j$$

and, taking expectations, the best estimate of $\Delta\eta$ is given by

$$\Delta y = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial \xi_i \partial \xi_j} u(x_i, x_j),$$

so that the best estimate of η is given by $y + \Delta y$. The variance of $\Delta\eta$ is given by $E(\Delta\eta^2) - (E(\Delta\eta))^2$. Using a fourth order expansion, $E(\Delta\eta^2)$ is estimated by

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial f}{\partial \xi_i} \frac{\partial f}{\partial \xi_j} u(x_i, x_j) + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{\partial f}{\partial \xi_i} \frac{\partial^2 f}{\partial \xi_j \partial \xi_k} E(\Delta\xi_i \Delta\xi_j \Delta\xi_k) +$$
$$\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{l=1}^{n} \left(\frac{1}{4} \frac{\partial^2 f}{\partial \xi_i \partial \xi_j} \frac{\partial^2 f}{\partial \xi_k \partial \xi_l} + \frac{1}{3} \frac{\partial f}{\partial \xi_i} \frac{\partial^3 f}{\partial \xi_j \partial \xi_k \partial \xi_l} \right) E(\Delta\xi_i \Delta\xi_j \Delta\xi_k \Delta\xi_l).$$

If we assume that the input quantities are independently distributed then the only terms in the quadruple sum above that can be nonzero are for the cases a) $i = j = k = l$, b) $i = j$ and $k = l$, c) $i = k$ and $j = l$, and d) $i = l$ and $j = k$. If we make the further assumption that the input distributions are symmetric, then all terms in the triple sum above are zero. With these assumptions, a higher order estimate of η 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

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

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 $\frac{\partial f}{\partial \xi_i}$, $\frac{\partial^2 f}{\partial \xi_i \partial \xi_i}$, and $\frac{\partial^3 f}{\partial \xi_i \partial \xi_j^2}$, 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 η is the same as the first order estimate, since $\Delta y = 0$, but its associated uncertainty is given by

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

so that $u(y + \Delta y) = 0.075$ mg, for the example data. This value compares well with the result given by MCM.

6. Evaluation of uncertainty using Bayesian methods

Renaming the mass correction term $\delta m_{R,c}$ by δm, Eq. (2) can be rearranged to give δm a function $\delta m = \phi(\boldsymbol{\alpha})$ of $\boldsymbol{\alpha} = (m_{R,c}, m_{W,c}, \rho_a, \rho_W, \rho_R)^T$:

$$\delta m = \frac{m_{W,c}}{1 + \left(\rho_a - \rho_{a_0}\right)\left(\frac{1}{\rho_W} - \frac{1}{\rho_R}\right)} - m_{R,c}\,. \tag{4}$$

The measurement experiment involving the balance produces a measurement *y* of $\delta m = \phi(\boldsymbol{\alpha})$ with an associated observation equation of the form $y = \phi(\boldsymbol{\alpha}) + \varepsilon$ where associated to term ε is a Gaussian distribution, i.e., $\varepsilon \in N(0, \sigma^2)$. Modelling the measurement process in this way allows us to calculate the probability $p(y \mid \boldsymbol{\alpha})$ of observing a measurement result *y* for any given value of the parameters. Given a prior probability density $p(\boldsymbol{\alpha})$ for the parameters $\boldsymbol{\alpha}$, Bayes' theorem [5,6] states that the posterior distribution $p(\boldsymbol{\alpha} \mid y)$ for $\boldsymbol{\alpha}$, taking into account the measurement information *y*, is such that $p(\boldsymbol{\alpha} \mid y) \propto p(y \mid \boldsymbol{\alpha}) p(\boldsymbol{\alpha})$. The prior distributions for $m_{R,c}$, ρ_a, ρ_W and ρ_R can be assigned as in Table 1. In addition we need a prior distribution for $m_{W,c}$ which take to be a normal distribution centred on the nominal value of 100 g and a standard deviation $\sigma_W =$ 5 g. We assume that all parameters $\boldsymbol{\alpha}$ are independently distributed. The posterior distribution is such that the nonzero

$$p(\boldsymbol{\alpha} \mid y) \propto \exp\left[-\frac{1}{2}\left\{\left(\frac{y-\phi(\boldsymbol{\alpha})}{\sigma}\right)^2 + \left(\frac{m_{W,c}-100}{\sigma_W}\right)^2 + \left(\frac{m_{R,c}-100}{\sigma_R}\right)^2\right\}\right].$$

The fact that $\phi(\boldsymbol{\alpha})$ is a (mildly) nonlinear function of the parameters $\boldsymbol{\alpha}$ means that this distribution departs (somewhat) from multivariate normal distribution. We can obtain a sample from this distribution using an inverse Monte Carlo (IMC) sampling method [7] as follows:

1. Sample $\mathbf{a}_q$ according to $p(\boldsymbol{\alpha})$, $q = 1,\ldots,M.$. For the mass calibration example this means sampling for univariate Gaussian or rectangular distributions.
2. Sample $y_q \in N(\phi(\mathbf{a}_q),\sigma^2)$, $q = 1,\ldots,M.$
3. Determine the index set $Q_y = \{q : \|y_q - y\| < \tau\}$, of those simulated measured values close to the actual measured value.
4. Determine the subsample $\{\mathbf{a}_q, q \in Q_y\}$. This subsample represents a sample from the posterior joint distribution $p(\boldsymbol{\alpha} \mid y)$ and $\{a_{k,q}, q \in Q_y\}$ is a sample from $p(\alpha_k \mid y)$ and can be used to determine the mean, standard deviation and coverage interval as in the standard Monte Carlo Method.

Fig. 1 shows the histogram of 20 000 IMC samples from $p(m_{W,c} - 100 \mid y = 1.234)$ compared with that derived using MCM. The IMC estimate of $m_{W,c} - m_0$ is 1.233 mg and its associated uncertainty is 0.073 mg, a little less than the MCM estimate.

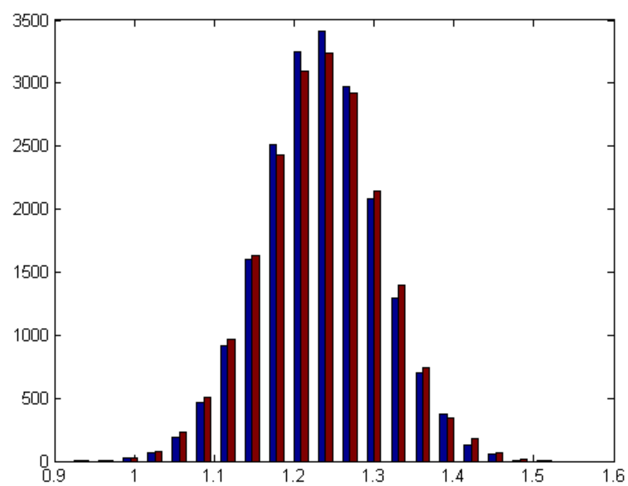

Figure 1. Frequency histogram for MCM (right hand bars) and IMC (left hand bars) samples for the mass calibration problem.

The two frequency histograms are expected to be somewhat different as they represent samples from different distributions. From a Bayesian point of view, Eq. (2) represents a way of estimating $m_{W,c}$ from the measurement data y:

$$\hat{m}_{W,c} = (m_{R,c} + y)\left[1 + (\rho_a - \rho_{a_0})\left(\frac{1}{\rho_W} - \frac{1}{\rho_R}\right)\right] = f(m_{R,c}, y, \rho_a, \rho_W, \rho_R),$$

and the MCM method attempts to propagate the distributions associated with the inputs through to the estimate. However, the distribution associated with y is $N(\phi(\boldsymbol{\alpha}),\sigma^2)$ and depends on $m_{W,c}$, the parameter we are trying to estimate. MCM gets round this difficulty by replacing $\phi(\boldsymbol{\alpha})$ with its measured estimate y. The Bayesian approach introduces no such approximation but uses instead a prior distribution for $m_{W,c}$.

The results of the four methods of evaluating an estimate of $m_{W,c} - m_0$ are given in Table 3, for the theoretical worked example. 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. Due to nonlinearities in the model, the first order GUM method significantly underestimates the standard deviation of the output distribution. The GUM 2nd order method gives satisfactory estimates and can be implemented for the mass calibration problem even if involving some complicated calculation. Of the first three methods applied to this example, MCM seems the most useful.

Table 3.
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	1.233	0.073

The fourth method based on a Bayesian formulation provides summary information for a different distribution, describing the state-of-knowledge about the parameters of the problem rather than the properties of a parameter estimation method. Because of the mild nonlinearities of the problem these two distributions are different. While it can be argued that the Bayesian formulation provides the correct basis for making inferences about the parameters, for the mass calibration problem the two approaches, MCM and IMC, give very similar results.

7. Calibration of an OIML class F2 weight

Experimental data from a calibration laboratory show that in practice the methods can be more consistent with each other. Table 4 shows experimental data related to the calibration of an OIML class F2 weight of 100 g nominal mass.

Table 4.
TAP experimental data related to the calibration of an OIML class F2 weight of 100 g nominal value.

ξ_j	Distribution	Mean	Standard deviation
$m_{R,c}$	Gaussian	99 999.634 mg	0.053 mg
$\delta m_{R,c}$	Gaussian	0.917 mg	0.029 mg
ρ_a	Rectangular	1.170 kg/m^3	0.10 kg/m^3
ρ_W	Rectangular	8.00×10^3 kg/m^3	$(1/\sqrt{3})\times0.10\times10^3$ kg/m^3
ρ_R	Rectangular	7.96×10^3 kg/m^3	$(1/\sqrt{3})\times0.05\times10^3$ kg/m^3

Applying the same methods to this real data produces comparable results between first order GUM and MCM, underlying our remark about the unrealistically high uncertainty in ρ_W. Nevertheless, the potential differences are demonstrated and the use of a validation technique such as MCM is highly recommended.

The experimental data draws our attention to another problem, though. The values assigned to $\delta m_{R,c}$ were obtained using a number n = 3 of repeated measurement, which is not uncommon in mass metrology. This means that the standard deviation and coverage interval are much larger (easily obtained with MCM assigned a *t*-Student distribution to this input variable), due to the fact that there is little evidence to determine the uncertainty associated with the estimate of $\delta m_{R,c}$. In practice, there is likely to be prior information available from similar experiments that, using a Bayesian formulation for example, can supplement the measurement information and potentially reduce the uncertainty associated with the estimate of $\delta m_{R,c}$ [7].

8. Summary and concluding remarks

The first three methods all attempt to provide summary information about the same probability distribution, namely the output distribution associated with Eq. (2). As the number of MCM trials increases, the discrete approximation can become more exact. The GUM would underestimate the uncertainty associated with the output quantity whenever nonlinearities prevail, but for a smaller uncertainty associated the density quantity ρ_W, the 1st order GUM gives a more accurate estimate of the uncertainty. The GUM 2nd order method gives satisfactory estimates but for a general model, there can be a considerable (or prohibitive) amount of work involved. The Monte Carlo method is a more flexible approach and it can be used as an important validation tool.

The Bayesian estimates are derived from inverse Monte Carlo simulation that, like MCM, is simple in concept and generally straightforward to implement. However, alternative simulation techniques such as Markov chain Monte Carlo [6] are much more efficient although conceptually less straightforward.

8. References

[1] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML. Guide to the Expression of Uncertainty in Measurement, 1995. ISBN 92-67-10188-9. Corrected and reprinted.

[2] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML: Propagation of distributions using a Monte Carlo method (GUM Supplement 1). *In press*.

[3] M. G. Cox and P. M. Harris. Software Support for Metrology Best Practice Guide Nº 6, Uncertainty Evaluation. NPL Report DEM-ES-011. National Physical Laboratory, Teddington, UK, September 2006.

[4] Nielsen, L. Private communication, 2002.

[5] Box, G. E. P. and Tiao, G. C.: Bayesian inference in statistical analysis, New York: Wiley, 1973.

[6] Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B.: Bayesian data analysis, Boca Raton: Chapman & Hall, 2004.

[7] Barker, R. M., Cox, M. G., Forbes, A. B. and Harris, P. M.: Software Support for Metrology Best Practice Guide No. 4, Discrete modelling and experimental data analysis, , NPL Report DEM-ES-018, National Physical Laboratory, 2007.