NPL

National Physical Laboratory

NPL REPORT MS 1

Software for GUM Supplement 1: User Manual

M G Cox, P M Harris and I M Smith

OCTOBER 2011

National Measurement System

Software for GUM Supplement 1: User Manual

M G Cox, P M Harris and I M Smith
Mathematics and Scientific Computing Group

ABSTRACT

This report constitutes a *user manual* for software developed at the National Physical Laboratory to support the use of the 'Guide to the expression of uncertainty in measurement' (GUM) and the first supporting document to the GUM, GUM Supplement 1, concerned with the use of a Monte Carlo method as an implementation of the propagation of distributions. The software enables users to apply the approaches to uncertainty evaluation described in these documents to the four example problems considered in GUM Supplement 1. The software is intended to allow users to reproduce the results presented in tables and figures contained within GUM Supplement 1. It is also intended to help users learn about the methods for uncertainty evaluation described in the GUM and GUM Supplement 1 by enabling them to experiment with (a) different information about the input quantities in the models defining the example problems, and (b) different values for the parameters controlling the application of those methods.

NPL Report MS 1

ISSN 1754–2960

First issued: January 2008
Revised: October 2011

National Physical Laboratory,
Hampton Road, Teddington, Middlesex, United Kingdom TW11 0LW

We gratefully acknowledge the financial support of the UK Department for Business, Innovation and Skills (National Measurement Office)

Approved on behalf of NPLML by Dr Bajram Zeqiri,
Knowledge Leader for the Acoustics and Ionizing Radiation team

Contents

1 Introduction

1.1 Background

The 'Guide to the expression of uncertainty in measurement' (GUM) [2] is the primary document regarding the evaluation and reporting of uncertainty in measurement. Working Group 1, 'Expression of uncertainty in measurement', of the Joint Committee for Guides in Metrology (JCGM) is undertaking work to promote and extend the application of the GUM through the preparation of supporting documents [1]. The first such document, GUM Supplement 1 [3], is concerned with the use of a Monte Carlo method as an implementation of the propagation of distributions for uncertainty evaluation.

GUM Supplement 1 includes four example problems as follows:

1. Additive model [3, subclause 9.2];
2. Mass calibration [3, subclause 9.3];
3. Comparison loss in microwave power meter calibration [3, subclause 9.4];
4. Gauge block calibration [3, subclause 9.5].

This report is a *user manual* for software developed at the National Physical Laboratory to support the use of the GUM and GUM Supplement 1 by enabling users to apply the approaches to uncertainty evaluation described in these documents to the above example problems. The software is intended to allow users to reproduce the results presented in tables and figures contained within GUM Supplement 1. The software is also intended to help users learn about the methods for uncertainty evaluation described in the GUM and GUM Supplement 1. It enables them to experiment with (a) different information about the input quantities in the models defining the example problems, e.g., estimates and associated standard uncertainties for the input quantities in those models, and (b) different values for the parameters controlling the application of those methods, e.g., the number of trials in an application of a Monte Carlo method.

The software does not allow users to undertake uncertainty calculations for measurement models other than those for the above example problems. However, detailed information about the algorithms implemented in the software is available [4].

For each example problem, the software is used to apply the GUM uncertainty framework [3, subclause 5.6] and a Monte Carlo method [3, subclause 5.9] as approaches to the propagation and summarizing stages of uncertainty evaluation [3, subclause 5.1]. Each approach delivers an estimate y of the output quantity Y in the measurement model, the standard uncertainty $u(y)$ associated with the estimate, a coverage interval $[y_{\text{low}}, y_{\text{high}}]$ for Y corresponding to a specified coverage probability p, and an approximation to the probability density function (PDF) for Y.

The Monte Carlo procedure may be applied either non-adaptively or adaptively. In a non-adaptive application of the procedure [3, subclause 7.2], a fixed number of Monte Carlo trials, specified by the user, is undertaken and a test of whether the results obtained have stabilized in a statistical sense performed at the end of the calculation. The test is based on a numerical tolerance $\delta_{\rm stab}$ calculated in terms of a number $n_{\rm stab}$ of significant decimal digits regarded as meaningful in the value of $u(y)$ [3, subclause 7.9.2].[1] In an adaptive application of the procedure [3, subclause 7.9], the calculation is terminated when either (a) the results obtained have stabilized, or (b) a maximum number of trials, specified by the user, has been undertaken.

In the case that the results obtained from the Monte Carlo procedure have stabilized, those results are used to validate the results obtained from the GUM uncertainty framework [3, clause 8]. The basis of the validation is a numerical tolerance $\delta_{\rm val}$ calculated in terms of a number $n_{\rm val}$ of significant decimal digits regarded as meaningful in the value of $u(y)$ [3, subclause 7.9.2].[2] A sufficient number of Monte Carlo trials should be undertaken in obtaining results from a Monte Carlo method for the purpose of validating those from the GUM uncertainty framework, and it is recommended that the numerical tolerances $\delta_{\rm stab}$ and $\delta_{\rm val}$ are chosen to satisfy $\delta_{\rm stab} \le \delta_{\rm val}/5$ [3, subclause 8.2].

The report is organized as follows. Section 2 gives information on installing and uninstalling the software. Section 3 describes how to use the software. General information is provided relating to the software for all four example problems (section 3.1), as well as information specific to each problem and its associated software (sections 3.2–3.5). Section 4 describes the conditions of use of the software that give the results presented in tables and figures contained within GUM Supplement 1.

1.2 Software User Licence Agreement

The software is provided with a Software User Licence Agreement (Ref: MSC/L/08/005) and the use of the software is subject to the terms laid out in that agreement. By installing and running the software, the user accepts the terms of the agreement.

To run the software, the user must install MATLAB's Component Runtime (MCR) libraries (section 2). The user must accept the terms of the MCR Library License as part of the installation of the MCR libraries.

[1] The notation $\delta_{\rm stab}$ and $n_{\rm stab}$ is used in place of the generic notation δ and $n_{\rm dig}$ used in subclause 7.9.2 of GUM Supplement 1 in the context of a test of stabilization of the results obtained from a Monte Carlo method. See section 3.1.5.

[2] The notation $\delta_{\rm val}$ and $n_{\rm val}$ is used in place of the generic notation δ and $n_{\rm dig}$ used in subclause 7.9.2 of GUM Supplement 1 in the context of validating the results obtained from the GUM uncertainty framework. See section 3.1.5.

1.3 Operating the software

The operation of the software, and in particular the numerical and graphical display of results, is intended for the example problems as they are described in GUM Supplement 1, and for instances of those problems that are 'close' to those described therein. Therefore, for each example problem, the values of the expectation and standard deviation of each input quantity in the model for the problem, obtained from the probability distribution that characterizes the quantity, are required to lie within stated intervals. The software undertakes checks on the values set by the user, and will not accept values lying outside those intervals (section 3.1.1). However, the performance of the software and the adequacy of the display of results are not guaranteed. For example, if the user sets a very large value for the number of Monte Carlo trials, the time for the computation to be completed may be unacceptably long. Also, if the user chooses a very small value for the standard uncertainty associated with the estimate of an input quantity, the number of figures to which the value is displayed may be inadequate to distinguish the displayed value from zero. Generally, the number of figures to which y, $u(y)$, y_{low} and y_{high} are displayed is greater than would normally be used to report results, but the intention is to facilitate the comparison of results. Sections 3.2–3.5 give, for the four example problems, the default values and settings that define the example problems, and the intervals in which the expectations and standard deviations for each input quantity are required to lie.

2 Installing and uninstalling the software

The software takes the form of four application programs, corresponding to the example problems listed in section 1.1, called

1. `NPLUnc_additive.exe`,
2. `NPLUnc_mass.exe`,
3. `NPLUnc_comparisonloss.exe`, and
4. `NPLUnc_gaugeblock.exe`.

The application programs have been created by compiling (using the MATLAB compiler) software implemented in the MATLAB programming language [5]. The programs have been created and tested on a personal computer running the Microsoft Windows XP Professional operating system.

To run the application programs it is first necessary to install MATLAB's Component Runtime (MCR) libraries. This is done by running the MCR installation program

```
MCRInstaller.exe
```

once on the target machine, i.e., the machine on which it is intended to run the application programs. It is necessary to have administrative privileges for the target machine because both the system registry and system path are modified as part of the installation process. The MCR installation program installs the MCR libraries, registers the components as needed, and updates the system path to point to the MCR binary directory. The installation process takes some time due to the number of files that are installed. The MCR installation program is about 173 MB in size, and the installed libraries require about 456 MB of disk space.

The software is uninstalled by

- running the MCR installation program `MCRInstaller.exe` and selecting 'Remove' to uninstall the MCR libraries, and
- deleting the application programs.

3 Using the software

3.1 General

An application program is run either by (a) double-clicking on the corresponding executable file (with the extension `.exe`) in Windows Explorer, or (b) opening a DOS window, navigating to the folder containing the program, typing the name of the program (without the extension `.exe`), and pressing `Return`. Running an application program the first time can take longer than when the program is run subsequently.

Each program allows the user to perform the following generic operations:

1. Modify the PDFs used to characterize the input quantities in the measurement model;
2. Modify the coverage probability and type of coverage interval;
3. Run an uncertainty calculation;
4. Exit.

Each of these operations is described below. Information specific to each example problem and its associated program is presented in sections 3.2–3.5.

3.1.1 Modify the PDFs

For example problems 1, 2 and 4 (section 1.1), each input quantity X_i in the measurement model is characterized by a probability distribution, which may be:

1. A Gaussian distribution $\mathrm{N}(\mu, \sigma^2)$ with expectation μ and standard deviation $\sigma \geq 0$ [3, subclause 6.4.7];
2. A rectangular distribution $\mathrm{R}(a, b)$ with lower limit a and upper limit $b \geq a$ [3, subclause 6.4.2];
3. A trapezoidal distribution $\mathrm{Trap}(a, b, \beta)$ with lower limit a, upper limit $b \geq a$ and parameter $\beta,\ 0 \leq \beta < 1$, equal to the ratio of the top width of the trapezoid to the base width [3, subclause 6.4.4];
4. A U-shaped (arc sine) distribution $\mathrm{U}(a, b)$ with lower limit a and upper limit $b \geq a$ [3, subclause 6.4.6];
5. An exponential distribution $\mathrm{Ex}(1/x)$ with expectation $x > 0$ [3, subclause 6.4.10];
6. A gamma distribution $\mathrm{G}(q+1, 1)$ with integer parameter $q \geq 0$ [3, subclause 6.4.11];
7. A curvilinear trapezoidal distribution $\mathrm{CTrap}(a, b, d)$ with lower limit a, upper limit $b \geq a$ and parameter d [3, subclause 6.4.3]. Here, d is defined by $d = r(b-a)/2$ where $r,\ 0 < r < 1$, is the (fractional) 'reliability' of the semi-width $(b-a)/2$ [2, subclauses G.4.2 and H.1.6];
8. A scaled and shifted t–distribution $\mathrm{t}_\nu(\mu, \sigma^2)$ with shift parameter μ, scale parameter $\sigma \geq 0$ and integer degrees of freedom $\nu \geq 3$ [3, subclause 6.4.9].

The input quantities are mutually independent.

For each input quantity, the user is presented with a list of probability distributions with a distribution highlighted. The distribution initially highlighted is the 'current' distribution that characterizes the input quantity and corresponds to either a default distribution (if the PDF has not been modified previously by the user) or a distribution that has been set by the user. This distribution is defined by 'current' values of the parameters for the distribution which either take default values (if the PDF has not been modified previously by the user) or have been set by the user. The current distribution with its current parameter values is selected by pressing `Cancel`.

The PDF used to characterize the input quantity is modified by selecting a distribution from the list and pressing `OK`. Values for the parameters defining the selected distribution are then set as follows:

- When the current distribution is selected, the user may change the current values for the parameters of that distribution. New values are set by changing one or more values and pressing `OK`. The current values are restored in the text boxes by pressing `Cancel`.
- When a distribution other than the current distribution is selected, the user must set values for the parameters of the distribution. Values are set by typing in the values

and pressing `OK`. Values previously entered in the text boxes are cleared by pressing `Cancel`.

Checks are undertaken to ensure the parameter values are valid. For example, $\sigma \geq 0$ for a Gaussian distribution, $b \geq a$ for a rectangular distribution, etc. (see above). Furthermore, to ensure that the problem defined by the user is 'close' to the (default) problem described in GUM Supplement 1, checks are undertaken to ensure that the expectation and standard deviation of each input quantity lie within stated intervals (section 1.3). If a check on the probability distribution set by the user for an input quantity fails, a warning message is displayed and the user must set another distribution for the quantity. A final check is undertaken to ensure that there is at least one input quantity for which the standard uncertainty associated with an estimate of the quantity is strictly positive, i.e., not all the input quantities have values that are known exactly. This condition ensures that the standard uncertainty associated with an estimate of the output quantity is not zero. If this final check fails, a warning message is displayed and the user must set new distributions for all the input quantities.

For example problem 3 (section 1.1), the (vector) input quantity $\boldsymbol{X} = (X_1, X_2)^\top$ is characterized by a bivariate Gaussian distribution $\mathrm{N}(\boldsymbol{x}, \boldsymbol{U_x})$ with $\boldsymbol{x} = (x_1, x_2)^\top$ an estimate of $\boldsymbol{X}$ with associated covariance matrix $\boldsymbol{U_x}$. The covariance matrix $\boldsymbol{U_x}$ is defined by

$$\boldsymbol{U_x} = \begin{bmatrix} u^2(x_1) & \rho u(x_1)u(x_2) \\ \rho u(x_1)u(x_2) & u^2(x_2) \end{bmatrix},$$

where $u(x_i) > 0$ is the standard uncertainty associated with x_i, and ρ, $-1 < \rho < 1$, is the correlation coefficient associated with x_1 and x_2.

The user is presented with a set of default values for $x_1, u(x_1), x_2, u(x_2)$ and ρ, and may change these values. New values are set by changing one or more values and pressing `OK`. The default values are restored in the text boxes by pressing `Cancel`. A number of checks are undertaken to ensure the parameters of the distribution characterizing $\boldsymbol{X}$ are valid (as above).

Sections 3.2–3.5 give the default distributions and default parameter values for the four example problems.

3.1.2 Modify the coverage interval

A coverage interval for the output quantity Y in the measurement model has two attributes, viz., a coverage probability p and an interval type [3, subclause 5.3]. The user may set the coverage probability to be 0.90, 0.95 or 0.99, and the type of coverage interval as the shortest or probabilistically symmetric interval. For each attribute, the user is presented with a list of possible values with a default value highlighted. The attribute is set by selecting a value from the list and pressing `OK`. The attribute is set to its default value, irrespective of whether a different value has been selected, by pressing `Cancel`.

3.1.3 Run an uncertainty calculation

The following steps are undertaken:

1. Information is displayed about the input quantities in the measurement model. The information is either default information set when the program is first run, or that set by the user as part of the operation 'Modify the PDFs' (section 3.1.1).

2. The GUM uncertainty framework for uncertainty evaluation is applied to obtain an estimate of the output quantity, the standard uncertainty associated with the estimate, the effective degrees of freedom attached to the standard uncertainty, and a coverage interval for the output quantity. The attributes of the coverage interval, i.e., the coverage probability and interval type, take either default values set when the program is first run, or values set by the user as part of the operation 'Modify the coverage interval' (section 3.1.2).

3. The results obtained from an application of the GUM uncertainty framework are displayed.

4. The PDF (a Gaussian or scaled and shifted t–distribution) provided by the GUM uncertainty framework to characterize the output quantity is graphed, with the endpoints of the coverage interval shown on the graph.

5. The user may set various controls for an application of a Monte Carlo method for uncertainty evaluation. The controls include:

   (a) Whether the Monte Carlo procedure is applied adaptively (`y` or `Y`) or non-adaptively (`n` or `N`). If the value set by the user is not one of `y`, `Y`, `n` or `N`, a warning message is displayed and the user must set a different value.

   (b) The maximum number of trials undertaken in an application of the Monte Carlo procedure. The number of trials is specified as a multiple of 10^4. In a non-adaptive application of the procedure, the maximum number of trials will be undertaken. In an adaptive application, fewer trials will be undertaken if the results obtained from the procedure have stabilized before the maximum number of trials has been reached. The value set by the user is rounded to the nearest integer M_h, and M_h must satisfy $M_h \geq 2$; otherwise a warning message is displayed and the user must set a different value.

   (c) The number of bins (or classes) for displaying as a scaled frequency distribution the approximation to the PDF for the output quantity obtained from the procedure. The value set by the user is rounded to the nearest integer N_{bin}, and N_{bin} must satisfy $N_{\text{bin}} \geq 10$; otherwise a warning message is displayed and the user must set a different value.

   (d) The initial state of the pseudo-random number generator used by the Monte Carlo procedure. The value set by the user is rounded to the nearest integer S, and S must satisfy $0 \le S \le 2^{32} - 1$; otherwise a warning message is displayed and the user must set a different value.

(e) The number $n_{\rm stab}$ of significant decimal digits in the value of the standard uncertainty associated with an estimate of the output quantity used to determine a numerical tolerance $\delta_{\rm stab}$ for testing the stabilization of the results obtained from the Monte Carlo procedure (section 3.1.5). The value set by the user must satisfy $n_{\rm stab} \ge 1$; otherwise a warning message is displayed and the user must set a different value.

(f) The number $n_{\rm val}$ of significant decimal digits in the value of the standard uncertainty associated with an estimate of the output quantity used to determine a numerical tolerance $\delta_{\rm val}$ for testing whether the results obtained from an application of the GUM uncertainty framework are validated by those obtained from an application of the Monte Carlo procedure (section 3.1.5). The value set by the user must satisfy $n_{\rm val} \ge 1$; otherwise a warning message is displayed and the user must set a different value.

Default values are provided for each of these controls, except the initial state of the pseudo-random number generator, for which a different value is set each time a calculation is run. New values are set by changing one or more values and pressing `OK`. The controls are set to their default values and the calculation is run, irrespective of whether a value has been changed, by pressing `Cancel`.

6. A Monte Carlo method for uncertainty evaluation is applied to obtain an estimate of the output quantity, the standard uncertainty associated with the estimate, and a coverage interval for the output quantity. The attributes of the coverage interval are as for the application of the GUM uncertainty framework (step 2). A progress bar gives an indication of the progress made by the Monte Carlo calculation. In an adaptive application of the Monte Carlo procedure that requires fewer than the maximum number of trials for the results to stabilize, the calculation will complete before the progress bar is full.

7. The results obtained from an application of a Monte Carlo method are displayed. In addition, the following information is displayed:

(a) The attributes of the coverage interval calculated from the applications of the GUM uncertainty framework and a Monte Carlo method.

(b) Information about whether the Monte Carlo calculation has stabilized. The information includes the number $h \times 10^4$ ($h \le M_h$) of trials undertaken and the value of $\delta_{\rm stab}$ used to test for stabilization. It also includes the values $2s_y$, $2s_{u(y)}$, $2s_{y_{\rm low}}$ and $2s_{y_{\rm high}}$ that are compared with $\delta_{\rm stab}$ in the test [3, subclause 7.9.4]. Here, the value s_y is the standard deviation associated with the average of the estimates $y^{(1)}, y^{(2)}, \ldots, y^{(h)}$ of the output quantity obtained from h sets of 10^4 trials, and similarly for $s_{u(y)}$, $s_{y_{\rm low}}$ and $s_{y_{\rm high}}$. The computation is regarded as having stabilized when *all* the values $2s_y$, $2s_{u(y)}$, $2s_{y_{\rm low}}$ and $2s_{y_{\rm high}}$ are no greater than δ_{stab}. In the case that the computation has not stabilized, an examination of the values can help to decide how many more trials might be needed to achieve stabilization of the results for the chosen value of n_{stab}.

(c) In the case that the Monte Carlo calculation has stabilized, information about whether the results obtained from an application of the GUM uncertainty framework are validated by those from an application of the Monte Carlo procedure, and the value of δ_{val} used for the validation.

8. A graph is provided of the length of the coverage interval as a function of the probability at its left-hand endpoint, calculated from the results obtained from an application of a Monte Carlo method (for example, [3, figures 7 and 12]).

9. The PDF provided by a Monte Carlo method to characterize the output quantity is graphed as a scaled frequency distribution, with the endpoints of the coverage interval shown on the graph.

10. The PDFs provided by the applications of the GUM uncertainty framework and a Monte Carlo method are graphed together, with the endpoints of the respective coverage intervals provided by the two approaches shown on the graph.

3.1.4 Exit

Pressing `Exit` will terminate the program and close all windows associated with running the program.

3.1.5 Setting the numerical tolerance

Let n_{dig} denote the number of significant decimal digits regarded as meaningful in $u(y)$. $n_{\text{dig}} = n_{\text{stab}}$ when testing for the stabilization of the results from a Monte carlo calculation, and $n_{\text{dig}} = n_{\text{val}}$ when validating the results obtained from an application of the GUM uncertainty framework against those obtained from an application of the Monte Carlo procedure. The numerical tolerance δ associated with $u(y)$ is defined as follows [3, subclause 7.9.2]:

1. Express $u(y)$ in the form $c \times 10^{\ell}$, where c is an n_{dig} decimal digit integer and ℓ is an integer;

2. Form
$$\delta = \frac{1}{2}10^{\ell}. \tag{1}$$

For example, consider the additive model (example problem 1) with each of its four input quantities characterized by the standard Gaussian distribution $\mathrm{N}(0, 1)$. If two significant decimal digits are regarded as meaningful in $u(y)$, then $n_{\text{dig}} = 2$, $u(y) = 20 \times 10^{-1}$ from which $\ell = -1$ and $\delta = 0.05$. Similarly, if three significant decimal digits are regarded as meaningful, then $n_{\rm dig} = 3$, $u(y) = 200 \times 10^{-2}$ from which $\ell = -2$ and $\delta = 0.005$.

The software uses the following formula to evaluate ℓ:

$$\ell = \text{floor}(\log_{10} u(y)) - (n_{\rm dig} - 1), \tag{2}$$

in which $\text{floor}(z)$ rounds the value z to the nearest integer towards $-\infty$. Thus, in the above example, with $u(y) = 2$ and $n_{\rm dig} = 2$, $\ell = \text{floor}(\log_{10} 2) - (2 - 1) = -1$, and when $n_{\rm dig} = 3$, $\ell = \text{floor}(\log_{10} 2) - (3 - 1) = -2$, as required. However, the formula implemented in the software is not restricted to values of $n_{\rm dig}$ and ℓ that are integers. For example, setting $n_{\rm dig} = 2.7$ gives $\ell = -1.7$ and $\delta \approx 0.01$.

It follows from formulae (1) and (2) that

$$n_{\rm dig} = \text{floor}(\log_{10} u(y)) - \log_{10} 2\delta + 1.$$

This formula can be used to set $n_{\rm dig}$, which is required by the software, in terms of δ, which is the information provided in the descriptions of the example problems given in GUM Supplement 1.

3.2 Additive model

The default settings for the probability distributions for the input quantities are that each input quantity X_i, $i = 1, \ldots, 4$, is characterized by the standard Gaussian distribution $\text{N}(0, 1)$, i.e., $X_i \sim \text{N}(0, 1)$. In the case that the probability distributions are modified by the user, the estimates (expectations) x_i of X_i and associated standard uncertainties (standard deviations) $u(x_i)$ are required to satisfy:

- $x_i \in [-100, 100]$, $i = 1, \ldots, 4$;
- $u(x_i) \leq 100$, $i = 1, \ldots, 4$.

The default settings for the coverage interval are:

- Coverage probability $p = 0.95$;
- Coverage interval of type probabilistically symmetric.

The default settings for the application of the Monte Carlo procedure are:

- The application of the procedure is adaptive;
- $M_h = 1\,000$;

- $N_{\text{bin}} = 50$;
- $n_{\text{stab}} = 2.7$;
- $n_{\text{val}} = 2$.

3.3 Mass calibration

The default settings for the probability distributions for the input quantities are [3, table 5]:

- Conventional mass of reference weight $m_{\text{R,c}} \sim \text{N}(\mu, \sigma^2)$ with $\mu = 100\,000.000\,\text{mg}$ and $\sigma = 0.050\,\text{mg}$;
- Conventional mass of balancing weight $\delta m_{\text{R,c}} \sim \text{N}(\mu, \sigma^2)$ with $\mu = 1.234\,\text{mg}$ and $\sigma = 0.020\,\text{mg}$;
- Air density $\rho_{\text{a}} \sim \text{R}(a, b)$ with $a = 1.1\,\text{kg/m}^3$ and $b = 1.3\,\text{kg/m}^3$;
- Density of weight to be calibrated $\rho_{\text{W}} \sim \text{R}(a, b)$ with $a = 7 \times 10^3\,\text{kg/m}^3$ and $b = 9 \times 10^3\,\text{kg/m}^3$;
- Density of reference and balancing weights $\rho_{\text{R}} \sim \text{R}(a, b)$ with $a = 7.95 \times 10^3\,\text{kg/m}^3$ and $b = 8.05 \times 10^3\,\text{kg/m}^3$.

In the case that the probability distributions are modified by the user, the estimates (expectations) $\widehat{m}_{\text{R,c}}$, $\widehat{\delta m}_{\text{R,c}}$, $\widehat{\rho}_{\text{a}}$, $\widehat{\rho}_{\text{W}}$ and $\widehat{\rho}_{\text{R}}$ of $m_{\text{R,c}}$, $\delta m_{\text{R,c}}$, ρ_{a}, ρ_{W} and ρ_{R} and associated standard uncertainties (standard deviations) are required to satisfy:

- $\widehat{m}_{\text{R,c}} - 100\,000\,\text{mg} \in [-100, 100]\,\text{mg}$ and $u(\widehat{m}_{\text{R,c}}) \le 100\,\text{mg}$;
- $\widehat{\delta m}_{\text{R,c}} \in [-100, 100]\,\text{mg}$ and $u(\widehat{\delta m}_{\text{R,c}}) \le 100\,\text{mg}$;
- $\widehat{\rho}_{\text{a}} \in [0.5, 1.5]\,\text{kg/m}^3$ and $u(\widehat{\rho}_{\text{a}}) \le 0.1\,\text{kg/m}^3$;
- $\widehat{\rho}_{\text{W}} \in [5\,000, 10\,000]\,\text{kg/m}^3$ and $u(\widehat{\rho}_{\text{W}}) \le 1\,000\,\text{kg/m}^3$;
- $\widehat{\rho}_{\text{R}} \in [5\,000, 10\,000]\,\text{kg/m}^3$ and $u(\widehat{\rho}_{\text{R}}) \le 1\,000\,\text{kg/m}^3$.

The default settings for the coverage interval are:

- Coverage probability $p = 0.95$;
- Coverage interval of type shortest.

The default settings for the application of the Monte Carlo procedure are:

- The application of the procedure is adaptive;
- $M_h = 1\,000$;
- $N_{\text{bin}} = 50$;
- $n_{\text{stab}} = 1.7$;
- $n_{\text{val}} = 1$.

3.4 Comparison loss in power meter calibration

The default setting for the probability distribution for the input quantities is that vector input quantity $\boldsymbol{X} = (X_1, X_2)^\top$ is characterized by a bivariate Gaussian distribution $\mathrm{N}(\boldsymbol{x}, \boldsymbol{U_x})$ with $x_1 = 0.010$, $x_2 = 0.000$, $u(x_1) = u(x_2) = 0.005$ and $\rho = 0$. In the case that the probability distribution is modified by the user, the estimates (expectations) x_i of X_i and associated standard uncertainties (standard deviations) $u(x_i)$ are required to satisfy:

- $x_i \in [-0.5, 0.5], \; i = 1, 2$;
- $u(x_i) \leq 0.1, \; i = 1, 2$.

The default settings for the coverage interval are:

- Coverage probability $p = 0.95$;
- Coverage interval of type shortest.

The default settings for the application of the Monte Carlo procedure are:

- The application of the procedure is non-adaptive;
- $M_h = 1\,000$;
- $N_{\text{bin}} = 50$;
- $n_{\text{stab}} = 2$;
- $n_{\text{val}} = 1$.

3.5 Gauge block calibration

The default settings for the probability distributions for the input quantities are [3, table 10]:

- Length of the reference standard $L_\mathrm{s} \sim \mathrm{t}_\nu(\mu, \sigma^2)$ with $\mu = 50\,000\,623\,\mathrm{nm}$, $\sigma = 25\,\mathrm{nm}$ and $\nu = 18$;
- Average length difference $D \sim \mathrm{t}_\nu(\mu, \sigma^2)$ with $\mu = 215\,\mathrm{nm}$, $\sigma = 6\,\mathrm{nm}$ and $\nu = 24$;
- Random effect of comparator $d_1 \sim \mathrm{t}_\nu(\mu, \sigma^2)$ with $\mu = 0\,\mathrm{nm}$, $\sigma = 4\,\mathrm{nm}$ and $\nu = 5$;
- Systematic effect of comparator $d_2 \sim \mathrm{t}_\nu(\mu, \sigma^2)$ with $\mu = 0\,\mathrm{nm}$, $\sigma = 7\,\mathrm{nm}$ and $\nu = 8$;
- Thermal expansion coefficient $\alpha_\mathrm{s} \sim \mathrm{R}(a, b)$ with $a = 9.5 \times 10^{-6}\,{}^\circ\mathrm{C}^{-1}$ and $b = 13.5 \times 10^{-6}\,{}^\circ\mathrm{C}^{-1}$;
- Average temperature deviation $\theta_0 \sim \mathrm{N}(\mu, \sigma^2)$ with $\mu = -0.1\,{}^\circ\mathrm{C}$ and $\sigma = 0.2\,{}^\circ\mathrm{C}$;
- Effect of cyclic temperature variation $\Delta \sim \mathrm{U}(a, b)$ with $a = -0.5\,{}^\circ\mathrm{C}$ and $b = 0.5\,{}^\circ\mathrm{C}$;
- Difference in expansion coefficients $\delta\alpha \sim \mathrm{CTrap}(a, b, d)$ with $a = -1.0 \times 10^{-6}\,{}^\circ\mathrm{C}^{-1}$, $b = 1.0 \times 10^{-6}\,{}^\circ\mathrm{C}^{-1}$ and $d = 0.1 \times 10^{-6}\,{}^\circ\mathrm{C}^{-1}$ corresponding to a (fractional) 'reliability' $r = 0.1$ (section 3.1.1);
- Difference in temperatures $\delta\theta \sim \mathrm{CTrap}(a, b, d)$ with $a = -0.050\,{}^\circ\mathrm{C}$, $b = 0.050\,{}^\circ\mathrm{C}$ and $d = 0.025\,{}^\circ\mathrm{C}$ corresponding to a (fractional) 'reliability' $r = 0.5$ (section 3.1.1).

In the case that the probability distributions are modified by the user, the estimates (expectations) $\widehat{L}_\mathrm{s}$, $\widehat{D}$, $\widehat{d}_1$, $\widehat{d}_2$, $\widehat{\alpha}_\mathrm{s}$, $\widehat{\theta}_0$, $\widehat{\Delta}$, $\widehat{\delta\alpha}$ and $\widehat{\delta\theta}$ of L_s, D, d_1, d_2, α_s, θ_0, Δ, $\delta\alpha$ and $\delta\theta$ and associated standard uncertainties (standard deviations) are required to satisfy:

- $\widehat{L}_\mathrm{s} - 50\,000\,000\,\mathrm{nm} \in [-1\,000, 1\,000]\,\mathrm{nm}$ and $u(\widehat{L}_\mathrm{s}) \leq 1\,000\,\mathrm{nm}$;
- $\widehat{D} \in [-1\,000, 1\,000]\,\mathrm{nm}$ and $u(\widehat{D}) \leq 100\,\mathrm{nm}$;
- $\widehat{d}_1 \in [-10, 10]\,\mathrm{nm}$ and $u(\widehat{d}_1) \leq 10\,\mathrm{nm}$;
- $\widehat{d}_2 \in [-10, 10]\,\mathrm{nm}$ and $u(\widehat{d}_2) \leq 10\,\mathrm{nm}$;
- $\widehat{\alpha}_\mathrm{s} \in [10 \times 10^{-6}, 15 \times 10^{-6}]\,{}^\circ\mathrm{C}^{-1}$ and $u(\widehat{\alpha}_\mathrm{s}) \leq 2 \times 10^{-6}\,{}^\circ\mathrm{C}^{-1}$;
- $\widehat{\theta}_0 \in [-1, 1]\,{}^\circ\mathrm{C}$ and $u(\widehat{\theta}_0) \leq 1\,{}^\circ\mathrm{C}$;
- $\widehat{\Delta} \in [-1, 1]\,{}^\circ\mathrm{C}$ and $u(\widehat{\Delta}) \leq 1\,{}^\circ\mathrm{C}$;
- $\widehat{\delta\alpha} \in [-1 \times 10^{-6}, 1 \times 10^{-6}]\,{}^\circ\mathrm{C}^{-1}$ and $u(\widehat{\delta\alpha}) \leq 1 \times 10^{-6}\,{}^\circ\mathrm{C}^{-1}$;
- $\widehat{\delta\theta} \in [-1, 1]\,{}^\circ\mathrm{C}$ and $u(\widehat{\delta\theta}) \leq 1\,{}^\circ\mathrm{C}$.

The default settings for the coverage interval are:

- Coverage probability $p = 0.99$;
- Coverage interval of type shortest.

The default settings for the application of the Monte Carlo procedure are:

- The application of the procedure is adaptive;
- $M_h = 1\,000$;
- $N_{\text{bin}} = 50$;
- $n_{\text{stab}} = 2$;
- $n_{\text{val}} = 1$.

4 Reproducing the results given in GUM Supplement 1

To reproduce the numerical and graphical results for the example problems given in GUM Supplement 1 it is necessary to set appropriately (a) the distributions for the input quantities, (b) the attributes of the required coverage interval, and (c) the controls for the application of a Monte Carlo method including (d) the value S of the initial state of the pseudo-random number generator used by the software. Regarding (a), (b) and (c), the default values and settings listed in sections 3.2–3.5 provide a starting point. Regarding (d), tables 1–4 give the values S that must be set by the user. For example:

1. For the problem of mass calibration [3, subclause 9.3], to obtain the numerical results contained in table 6 of GUM Supplement 1 and the corresponding graph contained in figure 10, use the default values listed in section 3.3 and the value $S = 89\,436\,961$ given in table 2.

2. For the problem of comparison loss in microwave power meter calibration [3, subclause 9.4], to obtain the numerical results in table 9 of GUM Supplement 1 corresponding to $x_1 = 0.010$ and the corresponding graph contained in figure 15, use the default values listed in section 3.4 but with the PDF for the input quantities modified so that $x_1 = 0.010$ and $\rho = 0.9$, and the value $S = 89\,436\,096$ given in table 3 corresponding to $x = 0.010$.

Method	Initial state S
MCM $\left(M = 10^5\right)$	89 437 050
MCM $\left(M = 10^6\right)$	89 436 096
MCM $\left(M = 10^6\right)$	89 436 961
Adaptive MCM ($n_{\mathrm{dig}} = 2.7$)	89 436 348
Adaptive MCM ($n_{\mathrm{dig}} = 2.7$)	89 437 071

Table 1: Initial states of the pseudo-random number generator used by the Monte Carlo procedure for the additive model [3, subclause 9.2].

Method	Initial state S
Adaptive MCM ($n_{\mathrm{dig}} = 1.7$)	89 436 961

Table 2: Initial state of the pseudo-random number generator used by the Monte Carlo procedure for the example problem of mass calibration [3, subclause 9.3].

Method	x_1	Initial state S
MCM $\left(M = 10^6\right)$	0.000	89 436 961
MCM $\left(M = 10^6\right)$	0.010	89 436 096
MCM $\left(M = 10^6\right)$	0.050	89 437 071

Table 3: Initial states of the pseudo-random number generator used by the Monte Carlo procedure for the example problem of comparison loss in microwave power meter calibration [3, subclause 9.4].

Method	p	Initial state S
Adaptive MCM ($n_{\mathrm{dig}} = 2.0$)	0.99	89 436 961
Adaptive MCM ($n_{\mathrm{dig}} = 2.0$)	0.95	89 436 096

Table 4: Initial states of the pseudo-random number generator used by the Monte Carlo procedure for the example problem of gauge block calibration [3, subclause 9.5].

Acknowledgements

The work described here was supported by the National Measurement Office of the UK Department of Business, Innovation and Skills as part of its NMS Mathematics and Modelling for Metrology programme.

References

[1] Bich, W, Cox M G and Harris P M 2006 Evolution of the 'Guide to the expression of uncertainty in measurement' *Metrologia* **43** S161–7

[2] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML *Evaluation of measurement data — Guide to the expression of uncertainty in measurement* Joint Committee for Guides in Metrology, JCGM 100:2008 `http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf`

[3] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML *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, JCGM 101:2008 `http://www.bipm.org/utils/common/documents/jcgm/JCGM_101_2008_E.pdf`

[4] Cox M G, Harris P M and Smith I M 2010 Software specification for uncertainty evaluation *NPL report MS 7*, National Physical Laboratory, Teddington, UK

[5] MATLAB (Copyright The MathWorks, Inc.) `www.mathworks.com`