Using simulation to check uncertainty calculations

B. D. Hall

Measurement Standards Laboratory of New Zealand,
Industrial Research Ltd, PO Box 31-310, Lower Hutt, New Zealand.
(email: b.hall@irl.cri.nz)

Abstract

A simulation approach is described for testing the performance of uncertainty calculations based on the *Guide to the expression of uncertainty in measurement* (GUM). Performance is measured in terms of the long-run success rate of an uncertainty calculation when applied to many simulated independent measurements. An individual calculation is deemed successful if the uncertainty interval obtained covers the measurand used in the simulation. Several examples, including two from the GUM, illustrate the approach. Software implementing the method is described in detail. Simulation is a practical method that can provide useful insights into specific measurement problems when there is any doubt about the validity of GUM calculations.

1 Introduction

The Guide to the Expression of Uncertainty in Measurement (GUM) is widely used by metrology organizations and is referred to in many national and international standards when an evaluation of measurement uncertainty is required [1]. Methods of evaluating uncertainty described in the GUM have been developed for general use. They work well when certain conditions are all satisfied, namely: when the response of a measurement system to small changes in influence quantities is approximately linear; when the distribution of measurement error associated with a result is approximately Gaussian and, if a calculation of effective degrees-of-freedom (the effective sample size) is required, when all influence quantity estimates are independent.

These conditions are very often satisfied in practice. However, doubts about the validity of GUM calculations have been raised from time to time in relation to particular types of measurement. Sometimes, a different method of uncertainty calculation is found to give results that disagree with the GUM. In such cases, it is usually very difficult to determine the correct method of uncertainty calculation analytically. So a general approach to testing GUM calculations would be very useful.

This article describes a straightforward simulation approach that can test whether the GUM can be applied to a particular type of measurement under specific conditions. A computer simulation of the measurement procedure is developed, which includes data-processing and a model of measurement errors. Then, using reasonable choices for the measurand, influence quantities and error distributions, a large number of independent data-sets are generated and an uncertainty interval is calculated for each. The performance of the uncertainty calculation method is evaluated by assessing the proportion of intervals that contain the measurand and comparing this with the nominal coverage (the 'level of confidence'). If the observed coverage found under simulated conditions

is near nominal (e.g., 95%) then a GUM uncertainty calculation can be applied with confidence to an actual measurement of the same type.

A simulation model for testing uncertainty calculations can be developed from information about a measurement procedure and the data provided in an uncertainty budget for the measurement. Such a model is not intended to reproduce the finer details of an actual experimental set-up. However, in as much as the uncertainty budget is a description of the main sources of measurement error and the measurement equation describes how such errors influence results, the model will be an adequate generator of data for testing uncertainty calculations.

This article describes the simulation approach in detail and outlines supporting software. In section 2, the approach is explained using a simple measurement scenario. Then, in section 3, it is applied to two examples from Appendix H of the GUM. The first of these is a detailed analysis of a gauge-block measurement. The GUM notes that several second-order terms make significant contributions to the uncertainty budget, but it is difficult to predict what effect these terms will have on the coverage of the uncertainty calculation and indeed whether the modified procedure described in the GUM is needed to achieve the desired level of confidence. The simulation method is used to assess the importance of these second-order terms.

In the second example, the impedance of an electric circuit component is measured. Estimates of the real and imaginary components of impedance are obtained from a common set of data and are therefore correlated. A small number of observations are made, so the effective sample size is also important in the uncertainty analysis of this problem. Now, the GUM method of calculating the effective degrees-of-freedom for the result (the Welch-Satterthwaite (WS) formula [1, G.2]) cannot be applied when estimates are correlated, so the GUM gives no guidance on how to obtain expanded uncertainty intervals in this case. We use simulation to select the best of several *ad hoc* calculations that might be used to complete the analysis.

2   Simulation method

A simple measurement scenario is used here to illustrate the simulation method. A quantity Y is defined in terms of two observable influence quantities

$$Y = X_1 + X_2 \ .$$

While X_1, X_2 and Y are assumed fixed but unknown, measurements of X_1 and X_2 are subject to Gaussian noise, leading to uncertainty in the estimate of Y.

To measure X_i (X_1 or X_2), a small number n_i of observations are collected. The sample mean x_i and the sample variance s_i^2 are calculated. In the terminology of the GUM, x_i is called the estimate of X_i and $u_i = s_i/\sqrt{n_i}$ is the associated standard uncertainty. GUM uncertainty calculations require three numbers to characterise each input quantity: x_i, u_i and ν_i, where $\nu_i = n_i - 1$.

2.1   Simulating the measurements

A sequence of independent values of x_i, u_i and ν_i must be generated to simulate measurements of X_i that can be used in GUM calculations. While ν_i is constant in this work, we require random number generators for x_i and u_i. Modeling observations by a Gaussian distribution with mean X_i and variance $n_i \sigma_i^2$, we consider the estimates x_i to

be drawn from a Gaussian distribution with mean X_i and variance σ_i^2. Consequently, we consider u_i to be the realisation of

$$\sigma_i \sqrt{\frac{\chi_{\nu_i}^2}{\nu_i}} \;,$$

where $\chi_{\nu_i}^2$ is a chi-square random variable with ν_i degrees of freedom.

A function `sim.gaussian` (see Appendix 5.1), written in the R programming language [2], generates three N-element vectors containing simulated estimates, standard uncertainties and degrees-of-freedom.[1] With this function, a simulation of N measurements of X_1 and X_2 can be written as

```
sim.x1 <- sim.gaussian(N,X1,sigma1,df1)
sim.x2 <- sim.gaussian(N,X2,sigma2,df2)
```

where the the results of N simulated measurements have been assigned to program variables `sim.x1` and `sim.x2`. Each variable refers to a list data structure in memory containing the N-element vectors for x, u and ν. List elements are accessed in R by appending a dollar sign and the elements name (i.e., `$x`, `$u` or `$df`) to the variable. For example, to obtain a vector of N simulated estimates of $Y = X_1 + X_2$ we could write

```
sim.y <- sim.x1$x + sim.x2$x
```

2.2   Calculation of uncertainty and coverage

A software package has been used to simplify the expression of GUM calculations and optimise computations [3]. The package implements the notion of an *uncertain number*, which automates the propagation of uncertainty through an arbitrary number of data processing steps [4]. In this case, the data processing for a measurement of Y, given simulated measurements of X_1 and X_2 is[2]

```
#------------------------------------
x1 <- elementary(
    'x1',
    sim.x1$x,
    sim.x1$u,
    sim.x1$df
)
x2 <- elementary(
    'x2',
    sim.x2$x,
    sim.x2$u,
    sim.x2$df
)
y <- intermediate(
    'y',
    ~ x1 + x2,
    list(x1,x2)
)
```

Here, the program variables `x1` and `x2` refer to N-element vectors of uncertain numbers, representing N the independent estimates x_1 and x_2, which are inputs to the GUM calculation. The variable `y` is an N-element vector of uncertain numbers, representing the N evaluations of $x_1 + x_2$. The calculation of N estimates of Y and the propagation of uncertainty information is done implicity during the call to `intermediate()`.

---

[1]Although the N values of ν_i are the same, the software performing the GUM uncertainty calculation requires a value of ν_i for each of the N simulated experiments.

[2]The term *elementary* describes an uncertain number defined by three numbers: an estimate, an uncertainty and a degrees-of-freedom. The term *intermediate* describes an uncertain number defined as a function of other uncertain numbers [4].

To assess coverage, we test whether the uncertainty interval generated for each experiment covers the measurand. This can be calculated in the following way

```
Y <- X1 + X2
cover <- 100 * sum(
    sim.successes(
        Y,
        value(y),
        uncertainty(y),
        dof(y)
    )
) / N
```

The functions `value(y)`, `uncertainty(y)` and `dof(y)` return N-element vectors containing the estimates of Y, the associated standard uncertainties and degrees-of-freedom, respectively. The function `sim.successes()` tests whether the measurand (`Y <- X1 + X2`) is contained in the 95% coverage interval evaluated for each measurement. It returns an N-element vector of Boolean values (see Appendix 5.1). Applying the R function `sum()` to this vector returns the number of successful intervals. So, finally, the percentage of successful intervals is assigned to `cover`.

2.3   Application

We now look at the performance of GUM calculations in the context of a measurement of Y when estimates of X_1 or X_2 are based on very small samples.

The measurement equation is linear, so the values used for X_1 and X_2 in the simulation can be set to zero without loss of generality, which assigns the measurand to zero. The following simulation parameters are also fixed: $\sigma_1 = 2$, $\sigma_2 = 1$ and $\nu_1 = 3$. Then, $N = 10^5$ simulations are performed for each of $\nu_2 = 1, 2, 3, 5, 10, 20, \infty$. The observed coverage values are shown in Table 1 as a function of ν_2. The results in Table 1 show departures from 95% coverage that are small and unlikely to be of practical concern.

ν_2	1	2	3	5	10	20	∞
coverage (%)	96.4	95.6	95.2	94.7	94.2	93.9	93.6

Table 1: The observed coverage of the uncertainty calculation for an estimate of $Y = X_1 + X_2$ when $\nu_1 = 3$ and $\sigma_1 = 2$, $\sigma_2 = 1$. The standard uncertainty in the observed coverage is approximately 0.07%

Observed coverage values, like those reported in Table 1, vary because a finite number of simulations is performed. Testing each uncertainty interval can be considered a Bernoulli trial, so the observed coverage, which is the result of N trials, has a binomial distribution. If the probability of success is p, and there are N independent trials, the standard deviation of the observed coverage will be $\sqrt{Np(1-p)}$. So, for $p = 0.95$ and $N = 10^5$, the standard deviation in the observed coverage will be 69, or 0.07%.

3   GUM examples

This section looks at two examples from Appendix H of the GUM and uses simulation to check the performance of different methods of uncertainty calculation. The GUM provides an uncertainty budget and a measurement equation in each case, from which a suitable simulation model can be inferred. Our aim is to simulate sets of data that can be regarded as representative of the type and conditions of measurement under consideration.

The information provided in the GUM identifies the physical quantities that contribute to a measurement result. These are imprecisely known, leading to an error in the final measurement result. The uncertainty budget provides estimates of each influence quantity and associates each with a distribution that is considered to represent what is known about the likely errors. We use this information to develop a simulation model.

The first example illustrates how the simulation method can be applied to real measurements by considering a gauge block calibration. There are nine influence quantities, three different types of distribution are associated with measurement errors and the measurement equation is slightly non-linear. The problem is sufficiently complicated to preclude any simple intuition about how coverage may be affected by departures from the assumptions on which the GUM method is based.

The second example is a simplified measurement scenario that is used in the GUM to discuss different approaches to data processing. In one of these approaches, the GUM cannot offer guidance on how to obtain an expanded uncertainty interval from the results. We consider this scenario as an example where several alternative uncertainty calculations could be envisaged as ways of finishing the calculation. Simulation provides a tool to assess which method can be expected to perform best.

3.1 Gauge block measurement

Example H.1 of the GUM describes a gauge block measurement. Analysis in the GUM identifies several second-order terms that are significant when compared with other terms in the uncertainty budget. When these terms are included in the calculations, the standard uncertainty of the result increases from 32 nm to 34 nm. In order to better understand how coverage is affected, we have tested the GUM uncertainty calculation with and without the extra terms using simulated data.

Section 3.1.1 reviews the information provided in the GUM. As much as possible, the GUM's notation is followed, however, it is necessary to make some small changes to clarify aspects of the simulation model that will be developed. Sections 3.1.2 and 3.1.3 describe the simulation model and data processing and section 3.1.4 presents the results obtained.

3.1.1 Description

In H.1 the measurand l is the length at 20 °C of an end gauge being calibrated (see [5] for more about this type of measurement and its uncertainty analysis). The measurement equation is given as

$$l = l_s + d - l_s(\delta_\alpha \theta + \alpha_s \delta_\theta) \, , \tag{1}$$

where the terms are described as follows.

- l_s is the length of the standard.

- d is the length difference between the standard and the gauge being calibrated.

- δ_α is the difference between the coefficients of thermal expansion for the standard and the gauge being calibrated.

- θ is the temperature deviation, from 20 °C, of the gauge being calibrated.

- α_s is the coefficient of thermal expansion of the standard.

- δ_θ is the difference in temperature between the two gauges.

Two of these terms are decomposed further

$$d = \bar{d} + d_1 + d_2 \,, \tag{2}$$

where

- $\bar{d}$ is the arithmetic mean of a number of length difference observations
- d_1 is associated with random comparator errors
- d_2 is associated with systematic comparator errors

and

$$\theta = \bar{\theta} + \Delta \tag{3}$$

where

- $\bar{\theta}$ is the mean difference of the test-bed temperature from 20 °C
- Δ is associated with a cyclical error in the test-bed temperature.

A summary of the information provided in the GUM is shown in Table 2.

term	distribution type	value	uncertainty parameter	degrees of freedom
l_s	Gaussian	$50\,000\,623\,\mathrm{nm}$	$25\,\mathrm{nm}$	18
$\bar{d}$	Gaussian	$215\,\mathrm{nm}$	$5.8\,\mathrm{nm}$	24
d_1	Gaussian	$0.0\,\mathrm{nm}$	$3.9\,\mathrm{nm}$	5
d_2	Gaussian	$0.0\,\mathrm{nm}$	$6.7\,\mathrm{nm}$	8
$\bar{\theta}$	Gaussian	$-0.1\,^\circ\mathrm{C}$	$0.2\,^\circ\mathrm{C}$	∞
Δ	arcsine	$0.0\,^\circ\mathrm{C}$	$0.5\,^\circ\mathrm{C}$	∞
α_s	uniform	$11.5 \times 10^{-6}\,^\circ\mathrm{C}^{-1}$	$2 \times 10^{-6}\,^\circ\mathrm{C}^{-1}$	∞
δ_α	uniform	$0.0\,^\circ\mathrm{C}^{-1}$	$1 \times 10^{-6}\,^\circ\mathrm{C}^{-1}$	50
δ_θ	uniform	$0.0\,^\circ\mathrm{C}$	$0.05\,^\circ\mathrm{C}$	2

Table 2: A summary of the information provided in the GUM for the end-gauge measurement. The uncertainty parameter depends on the type of distribution. For Gaussian distributions the parameter is the standard deviation, for uniform and arcsine distributions it is the half-width.

Now equations (2) and (3) pose some difficulties for a simulation study. They contain both terms representing physical quantities and terms representing estimates of physical quantities. Measurement equations should define the relationship between the quantities involved in a measurement, so we need to re-interpret (2) and (3) in this light. This could be done by adopting a notation to distinguish between quantities and estimates. However, the GUM expressly tries to avoid special notation, relying on context to help the reader decide whether a quantity or estimate is being referred to. We will endeavour to follow that policy. Nevertheless, it is important to be very clear about the distinction when developing a simulation model of the physical measurement system.

In equation (2), d, d_1 and d_2 are physical quantities (a length difference and two errors), however $\bar{d}$ is a numerical estimate (the mean of a series of comparator indications). If we re-interpret $\bar{d}$ as the quantity associated with operation of the comparator, we can obtain a quantity equation by writing

$$\bar{d} = d + d_0 + d_1 + d_2 \,, \tag{4}$$

where

- d, d_1 and d_2 are as defined above,

- d_0 is associated with a repeatability error,

We assume that $d_0 \approx 0$ and re-interpret $u(\bar{d})$, from the GUM, as the standard uncertainty $u(d_0)$. The quantity $\bar{d}$ is associated with the actual comparison measurement, which is designed to measure d but has some influences of its own. This interpretation seems compatible with the description in the GUM.

A similar issue arises in equation (3), where θ and Δ are quantities and $\bar{\theta}$ is described as an estimate. In this case, we simply choose to reinterpret $\bar{\theta}$ as a physical quantity (representing an actual mean temperature, rather than an estimate) in the defining equation (3). The same symbol will represent an estimate of that quantity, with an associated uncertainty $u(\bar{\theta})$, when the context requires it.

Equations (1), (2) and (3) are used for data processing and the evaluation of uncertainty in the GUM. In that context, all terms refer to quantity estimates. In particular, the measurand estimate is

$$ l = l_{\mathrm{s}} + \bar{d} $$

and the associated standard uncertainty can now be written as[3]

$$ u(l) = \sqrt{u^2(l_{\mathrm{s}}) + u^2(d_0) + u^2(d_1) + u^2(d_2) + l_{\mathrm{s}}^2\theta^2 u^2(\delta_\alpha) + l_{\mathrm{s}}^2\alpha_{\mathrm{s}}^2 u^2(\delta_\theta)} \ , $$

which takes into account the term d_0 introduced above.

3.1.2 Simulation

To develop a simulation model, we must consider the actual measurement, which is a comparison of two gauge blocks. This is expressed in equation H.1 of the GUM

$$ d = l(1 + \alpha\theta) - l_{\mathrm{s}}(1 + \alpha_{\mathrm{s}}\theta_{\mathrm{s}}) \ , $$

where α is the coefficient of thermal expansion of the gauge block being calibrated, θ_{s} is the temperature deviation from 20 °C of the standard gauge block and other terms have already been defined. The relationship of α and θ_{s} to other terms is

$$ \begin{aligned} \delta_\theta \ &= \theta - \theta_{\mathrm{s}} \\ \delta_\alpha \ &= \alpha - \alpha_{\mathrm{s}} \ . \end{aligned} $$

Combining this with equations (3) and (4) we obtain, for the measurement of the difference in length between two gauge blocks,

$$ \bar{d} \ = \ l \ - \ l_{\mathrm{s}} \ + \ l(\alpha_{\mathrm{s}} \ - \ \delta_\alpha)(\bar{\theta} \ + \ \Delta) \ - \ l_{\mathrm{s}}\alpha_{\mathrm{s}}(\bar{\theta} \ + \ \Delta \ - \ \delta_\theta) \ + \ d_0 \ + \ d_1 \ + \ d_2 \quad . \quad (5) $$

This is central to our simulation model, which must generate the estimates of $\bar{d}$. An algorithm based on the quantity equation (5) keeps the values of l, l_{s}, α_{s} and $\bar{\theta}$ constant and randomly generates values of d_0, d_1, d_2, Δ, δ_θ and δ_α for calculating $\bar{d}$. The distinction between the two sets of quantities recognizes that d_0, d_1, d_2, Δ, δ_θ and δ_α are essentially residual errors. On every occasion the best estimate of these quantities is zero, so it is necessary to simulate their underlying variability when generating values of $\bar{d}$. On the other hand, the quantities l_{s}, α_{s} and $\bar{\theta}$ are presumably measured. We take the view that these quantities are fixed physical quantities (including l, the measurand)

---

[3]This equation is the equivalent of (H.5) in [1, H.1.3] when our slight change of notation is taken into account. It is obtained using the standard GUM method for propagation of uncertainty and assuming that the different components of uncertainty are independent.

during the simulations that generate values of $\bar{d}$. Estimates of l_{s}, α_{s} and $\bar{\theta}$, representing measurements of these quantities, are generated independently for use in the data processing and uncertainty analysis.

Table 3 summarizes how each term in (5) is handled when generating values of $\bar{d}$ and gives the distribution parameters for those quantities. Table 4 summarizes the parameters used when generating independent estimates of l_{s}, α_{s} and $\bar{\theta}$. The type of distribution from which values are drawn in each case is taken from Table 2, as are the distribution parameters for the mean, standard deviation and degrees of freedom. In so doing, we expect the simulations will fairly represent the (unknown) conditions under which the data reported in the GUM were obtained.

Table 3: Summary of parameters used to simulate $\bar{d}$ using equation (5). For random quantites, the types of distribution and parameters for the associated random number generator are given.

quantity	random	distribution	value	uncertainty	dof
l_{s}	N	-	$50\,000\,623$ nm	-	-
d_0	Y	Gaussian	0	5.8 nm	24
d_1	Y	Gaussian	0	3.9 nm	5
d_2	Y	Gaussian	0	6.7 nm	8
$\bar{\theta}$	N	-	-0.1 °C	-	-
δ_θ	Y	uniform	0 °C	0.05 °C	2
Δ	Y	arcsine	$0\,\mathrm{°C^{-1}}$	$0.5\,\mathrm{°C^{-1}}$	∞
α_{s}	N	-	$11.5 \times 10^{-6}\,\mathrm{°C^{-1}}$	-	-
δ_α	Y	uniform	0	$10^{-6}\,\mathrm{°C^{-1}}$	∞

Table 4: Summary of parameters used to simulate estimates of l_{s}, α_{s} and $\bar{\theta}$. The distribution types and parameters for the associated random number generator are given.

quantity	distribution	value	uncertainty	dof
l_{s}	Gaussian	$50\,000\,623$ nm	25 nm	18
$\bar{\theta}$	Gaussian	-0.1 °C	0.2 °C	∞
α_{s}	uniform	$11.5 \times 10^{-6}\,\mathrm{°C^{-1}}$	$2 \times 10^{-6}\,\mathrm{°C^{-1}}$	∞

3.1.3 Data processing

Each simulated experiment yields estimates for $\bar{d}$, l_{s}, α_{s} and $\bar{\theta}$. In addition, there are five quantities in the measurement problem with finite degrees of freedom, so the standard uncertainties $u(l_{\mathrm{s}})$, $u(d_0)$, $u(d_1)$, $u(d_2)$ and $u(\delta_\alpha)$ are also generated randomly. As already mentioned, the estimates of d_0, d_1, d_2, Δ, δ_θ and δ_α are always zero. The standard uncertainties $u(\bar{\theta})$, $u(\Delta)$ and $u(\alpha_{\mathrm{s}})$, which have infinite degrees of freedom, are taken directly from Table 2. This data, together with the number of degrees of freedom reported in Table 2, can be used to obtain an estimate of l and apply the GUM method to evaluate $u(l)$.

The two second-order terms identified in the GUM as making significant contributions to the uncertainty are $l_{\mathrm{s}}u(\delta\alpha)u(\theta)$ and $l_{\mathrm{s}}u(\alpha_{\mathrm{s}})u(\delta\theta)$, where $u(\theta)$ is the combined standard uncertainty of $\bar{\theta} + \Delta$. These terms can be included in an extended GUM uncertainty calculation by following an approach used in an earlier study [6]. We treat the terms as independent influences and assign to each a number of degrees-of-freedom equal to the lesser of the degrees-of-freedom in their factors ($\nu_{\delta\alpha}$ and $\nu_{\delta\theta}$, respectively).

A listing of the R code used to simulate data, apply the GUM data processing and measure the performance of the two methods of uncertainty calculation is given in Appendix 6 with some additional comments.

3.1.4 Results

The results of testing are shown in Table 5. Coverage is found to be just less than nominal in the standard GUM calculation and slightly greater than nominal in the calculation with extra terms. The mean of the standard uncertainties generated during testing was 31.8 nm, increasing to 34.0 nm when the additional terms were included, which agrees with the GUM values (32 nm and 34 nm, respectively). The standard deviations of the standard uncertainties were 5 nm in both cases, indicating that a reasonable range of inputs to the uncertainty calculation was generated by the simulations.

method	coverage (%)
standard	94.4
extra terms	95.6

Table 5: Results from $N = 10^5$ simulations of the gauge block measurement. The standard uncertainty in the coverage values is approximately 0.07%.

The performance requirement for this uncertainty calculation is 95% coverage probability. The simulation results show a very slight difference between the two methods: the GUM calculation with additional terms meets this requirement and the standard GUM calculation falls short. The difference is very small and of questionable practical importance. It is interesting to note that the apparently significant difference in standard uncertainty (about 6%) obtained using the two methods of uncertainty calculation corresponds to a much smaller deviation from nominal coverage (about 0.5%). In view of the complexity of this measurement (nine influence quantities, three types of uncertainty distribution and non-linearity in the measurement equation), it is very difficult to anticipate how much coverage will be affected by departures from the basic assumptions. Simulation, is an effective method of checking the performance. Further investigation showed that the coverage of the standard GUM calculation is insensitive to the type of distribution used to simulate the influence quantities. For example, when every influence was modeled by a Gaussian distribution, the coverage of both calculations remained essentially the same.

3.2 Impedance measurement

Example H.2 of the GUM describes the measurement of several quantities associated with an electrical circuit component. The measurement design is artificially simple, but the example is interesting because it illustrates the GUM treatment of multiple measurands and the handling of uncertainty calculations when estimates are correlated.

Estimates of the resistance, reactance and the magnitude of the impedance are required. These can be obtained as a function of three other quantities measured in the circuit: the amplitude V of an AC potential difference across the component terminals, the amplitude I of the AC current through the component and the phase-shift ϕ of the potential difference with respect to the current. Five simultaneous measurements are made of V, I and ϕ, which provide a pool of data for the analysis. The three quantities of interest are related to in-circuit quantities by the following equations: the resistance

$$R = \frac{V}{I} \cos \phi , \qquad (6)$$

the reactance

$$X = \frac{V}{I} \sin \phi \qquad (7)$$

and the impedance-magnitude

$$Z = V/I . \qquad (8)$$

Appendix H.2 compares two ways of processing the measurement data. In the first, the means $\bar{V}$, $\bar{I}$ and $\bar{\phi}$ and the associated standard uncertainties $u(\bar{V})$, $u(\bar{I})$ and $u(\bar{\phi})$ are calculated, as are the sample correlation coefficients $r(\bar{V}, \bar{I})$, $r(\bar{V}, \bar{\phi})$ and $r(\bar{I}, \bar{\phi})$. Each of $u(\bar{V})$, $u(\bar{I})$ and $u(\bar{\phi})$ has four degrees-of-freedom, being based on five observations. Table 6 summarises data reported in the GUM. With equations (6), (7) and (8), and this data, the GUM method of uncertainty propagation can be applied to obtain the uncertainty of estimates of R, X and Z.

quantity	value	uncertainty
$\bar{V}$	4.9990 V	0.0032 V
$\bar{I}$	19.661 mA	0.0095 mA
$\bar{\phi}$	1.04446 rad	0.00075 rad
Correlation coefficients		
$r(\bar{V}, \bar{I}) = -0.36$		
$r(\bar{V}, \bar{\phi}) = +0.86$		
$r(\bar{I}, \bar{\phi}) = -0.65$		

Table 6: A summary of measurement data from the GUM for the resistance-reactance measurement.

The second way of processing data uses equations (6), (7) and (8) to calculate three samples of five values each for R, X and Z directly from measurements of V, I and ϕ. The sample means then estimate the corresponding measurands and standard uncertainties can be calculated from the sample standard deviation in the means, without any need for uncertainty propagation. We do not consider this method further here, but note that four degrees of freedom would naturally be associated with each of the standard uncertainties obtained.

Returning to the first method of data processing, the estimates of V, I and ϕ are obtained from a common pool of measurements, so they are correlated. The GUM method of calculating the effective degrees-of-freedom (the WS formula [1, G.2]) cannot be applied when estimates are correlated. So the GUM provides no guidance on how to construct uncertainty intervals for this problem.

In this section, we suppose that a final statement of uncertainty is required and consider three ways in which the analysis might be completed. The first attributes four degrees of freedom to the standard uncertainties of the results (the same as the degrees of freedom associated with the input quantity estimates), the second applies the WS formula (ignoring correlations) and the last method attributes an infinite number of degrees of freedom to the inputs (so the WS formula is not required – the degrees of freedom associated with the standard uncertainty in the results is also infinite). These are simplistic alternatives but they highlight a real issue: sometimes an assessment of the validity of an *ad hoc* uncertainty calculation procedure is required.

3.2.1 Simulation

The simulation model must generate estimates of three input quantities that are correlated and based on a small effective sample size. The function `sim.mvgaussian()`

provides a multivariate version of the routine used to simulate estimates of a Gaussian influence quantity (see §5.1.5). In this case, the parameters for `sim.mvgaussian()` are taken from the data in Table 6 and we are once again assuming that the conditions used in the simulation, for the purposes of evaluating uncertainty calculation coverage, will be close to those that prevailed during the actual experiment. The estimates of V, I and ϕ are returned as the elements of random vectors drawn from a multivariate Gaussian distribution. The distribution has mean vector

$$(\overline{V}, \overline{I}, \overline{\phi})$$

and covariance matrix

$$\begin{bmatrix} u(\overline{V})^2 & u(\overline{V})r(\overline{V},\overline{I})u(\overline{I}) & u(\overline{V})r(\overline{V},\overline{\phi})u(\overline{\phi}) \\ u(\overline{I})r(\overline{I},\overline{V})u(\overline{V}) & u(\overline{I})^2 & u(\overline{I})r(\overline{I},\overline{\phi})u(\overline{\phi}) \\ u(\overline{\phi})r(\overline{\phi},\overline{V})u(\overline{V}) & u(\overline{\phi})r(\overline{\phi},\overline{I})u(\overline{I}) & u(\overline{\phi})^2 \end{bmatrix} .$$

Each estimate is associated with an uncertainty matrix drawn from a Wishart distribution, which arises as the distribution of the sample covariance matrix for a sample from a multivariate Gaussian distribution [7, §4.4]. The elements of the list returned by `sim.mvgaussian()` consist of: N 3-element vectors of estimates, N 3×3 covariance matrices for the uncertainties and an N-element vector of degrees-of-freedom. The source code for this simulation is shown in Appendix 7.

3.2.2 Results

Coverage results for the three measurands in this problem are shown in Table 7. The column labelled $\nu = 4$ shows the observed coverage for the method that attributed four degrees-of-freedom to each output quantity. The column labeled $r = 0$ shows the results obtained when correlation is ignored in both the calculation of combined standard uncertainty and the WS formula.[4] The column labeled $\nu = \infty$ shows the observed coverage when infinite degrees-of-freedom are associated with the measurement results.

quantity	coverage (%)		
(Ω)	$\nu = 4$	$r = 0$	$\nu = \infty$
R	95.0	100.0	88.0
X	95.0	89.5	87.9
Z	95.0	95.2	87.8

Table 7: Coverage obtained from $N = 10^5$ simulations of the impedance measurement. See text for a description of the different cases. The standard uncertainty in coverage values is approximately 0.07%.

Without a doubt, the best method, suggested recently by Willink [8], assigns the same number of degrees-of-freedom to the output quantities as are attributed to the input quantities. The second method, in which correlation is ignored, generates uncertainty statements that are conservative for R and optimistic for X. The treatment in H.2 also evaluates the standard uncertainties under the assumption of independent inputs and obtains a standard uncertainty for R that is larger than the correlated result, a standard uncertainty for X that is smaller and a standard uncertainty for Z that is roughly the same. These findings are consistent with our coverage results. Finally, the third method, which ignores the effective sample size, generates uncertainty statements

---

[4]When inputs are correlated, the calculation of combined standard uncertainty should make use of the estimated correlation coefficients (see [1, §5.2.2]). However, here it is assumed that all correlation coefficients are zero, so the calculation of combined standard uncertainty is simplified (see [1, §5.1.2]).

that are uniformly too low for all outputs. The uniformity of coverage here suggests that raising the coverage factor[5] would obtain nominal coverage. That, of course, is what is achieved by the first method.

4 Discussion and Conclusion

A measurement is simply a procedure designed to estimate a quantity of interest. Every time it is applied, the result obtained differs from the quantity intended to be measured by some unknown error. A statement of uncertainty is an indication of how 'good' a measurement result is (how close it is likely to be to the measurand). However, with each new set of measurement data the uncertainty calculated may change. The method of calculating uncertainty, on the other hand, is fixed by the measurement procedure. It has a nominal coverage probability that can be interpreted in terms of the expected long-run success rate for many independent measurement problems.[6] So, while it is impossible to identify whether actual uncertainty intervals cover the measurand in real measurements, it is possible to assess the long-run rate of successful outcomes averaged over many independent measurements. This is an objective performance measure of the measurement procedure.

Numerical simulations can be used to test coverage, provided the simulation model gives a reasonable representation of the actual measurement system. We argue that a satisfactory simulation model can be developed from information about a measurement procedure, together with the data provided in an uncertainty budget, and have given examples in §3.1 and §3.2. This approach relies on the uncertainty budget being an adequate description of the main sources of measurement error. It is also important to clearly identify how other physical quantities may influence the measurement by stating a complete measurement equation. Both of these requirements are fundamental to the GUM recommendations on the reporting of measurement uncertainty.

A simulation model is not required to reproduce the finer details of an actual experimental set-up. In as much as the uncertainty budget is an adequate description of the main sources of measurement error and the measurement equation adequately describes how errors influence results, the model will be a suitable generator of data for testing uncertainty calculations. If desired, the parameters of the simulation model can be varied to explore a wider range of scenarios (this was illustrated in §2.3, but not done in §3.1 or §3.2). It is also possible to study the sensitivity of an uncertainty calculation procedure to the types of uncertainty distributions that are associated with influence quantity estimates. We alluded to this briefly at the end of §3.1.4, when stating that coverage remained the same when all every influence was modeled by a Gaussian distribution. It is certainly possible to investigate a particular choice of distributions in the uncertainty budget by testing the coverage of the procedure with simulated data.

Simulation offers the measurement scientist an important tool when there is any doubt about the validity of a GUM calculation. There is likely to be a number of specific 'niche' problems of this type, where it may not be appropriate to follow the general GUM guidelines exactly. Alternative procedures may offer intelligent, pragmatic, responses to the situation but there is a need to justify their use. A recent example of this occurred in the context of microwave measurements, where adaptive uncertainty calculations were being used to overcome perceived problems with the GUM method [12]. Expertise and common-sense suggested a 'better' course of action when the GUM method was thought

---

[5]The product of a coverage factor and the standard uncertainty is called the expanded uncertainty. It is the half-width of a GUM uncertainty interval [1, §2.3.6].

[6]More detailed discussion about using long-run success-rate to validate uncertainty calculations is given in [9, §2], [10, §2.2] and [11].

to be failing and the metrologist would use personal judgement to decide when to use the alternative procedure. By simulating this type of problem, it was possible to check the validity of the *ad hoc* method and optimise performance by using an objective criterion, rather than human judgement.

Recently [13], a simple radio-frequency measurement scenario was studied by simulation. The quantity of interest was the magnitude of a complex reflection coefficient that was approximately zero. The results highlighted differences between GUM uncertainty calculations and a Monte Carlo method proposed in the the first Supplement to the GUM (SUP) [14] and stimulated some discussion about the correctness of validating the SUP method in this way [9, 11, 15–17]. We note, however, that no objections have been raised regarding simulation as a way of testing of the GUM method. Indeed, at the time of writing, the most recent contribution to this discussion makes a compelling argument for its general use [11].

There may be good reasons to suspect that the GUM method is failing, however, when assessed objectively by simulation, its performance has turned out to be surprisingly good in difficult situations. It is hard to really identify measurement conditions in which uncertainty calculations will fail, because the randomness of measurement results leads to randomness in the associated uncertainty statements that is difficult to assess. It is necessary to consider the long-run behaviour of the calculation procedure to decide if there is really a problem. For example, when the lower bound of an uncertainty interval is negative for a quantity of interest that is strictly positive, failure of the GUM method is often assumed. Also, when a measurement function is strongly non-linear, such as the determination of a squared magnitude close to the origin in the complex plane, the GUM is expected to fail. The studies in [12, 13] deal with such cases and reveal that, while the GUM method certainly can fail, it performs rather better than might have been expected. More importantly, the conditions for which it is still safe to use the GUM can be assessed with some objectivity and a decision can be made about the whether alternative methods will perform better.

In conclusion, the use of long-run coverage success rate to objectively measure the performance of uncertainty calculations under simulated measurement conditions is an effective way of determining whether an uncertainty calculation is fit-for-purpose. The assessment of coverage of a statistical procedure by simulation is now taught at an introductory university level [18, p123]. However, it does not seem to be well-known within the measurement community.[7] There are only a few authors that report using the method to check the validity of methods of uncertainty calculation (see, e.g., [19, 20], also [21] describes the use of simulation to assess coverage performance in a slightly different context). The practicality of the approach does, of course, depend on the ease with which it can be applied to specific problems. That is why this article has described a recent general-purpose software implementation in some detail. Similar tools could be developed using a number of other readily available mathematical software packages.

Acknowledgement

---

[7]Note that, within the metrology community, the term 'simulation' has also been associated with the *propagation of distributions* method of evaluating uncertainty described in the SUP [14]. In that context, numerical Monte Carlo methods are used to simulate distributions-of-belief, not to simulate data and assess coverage as proposed here (for an application of the proposed method to uncertainty calculation using the SUP method, see [13]).

References

[1] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, and OIML. *Evaluation of measurement data Guide to the expression of uncertainty in measurement JCGM 100:2008 (GUM 1995 with minor corrections)*. BIPM Joint Committee for Guides in Metrology, Paris, Sèvres, 1 edition, 2008.

[2] R Development Core Team. *R: A language environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2009.

[3] B. D. Hall. Calculating uncertainty in r: `UncertainRealNumbers.R`. Manual 82, Industrial Research Ltd, March 2009.

[4] B. D. Hall. Computing with uncertain numbers. *Metrologia*, 43:L56–L61, 2006.

[5] Jennifer E. Decker and James R. Pekelsky. Uncertainty of gauge block calibration by mechnaical comparison: A worked example. Technical Report 39998, Institute for Mational Measurement Standards, NRC, Canada, 1996.

[6] B. D. Hall. Computer modelling of uncertainty calculations with finite degrees of freedom. *Measurement Science and Technology*, 11:1335–1341, 2000.

[7] Richard A Johnson and Dean W Wichern. *Applied multivariate statistical analysis*. Prentice-Hall International, 1998.

[8] R. Willink. A generalization of the Welch-Satterthwaite formula for use with correlated uncertainty components. *Metrologia*, 44:340–349, 2007.

[9] R. Willink. On the validity of methods of uncertainty evaluation. *Metrologia*, 47:80–89, 2010.

[10] R. Willink. Principles of probability and statistics for metrology. *Metrologia*, 43:S211–S219, 2006.

[11] R. Willink. Probability, belief and success rate: comments on 'On the meaning of coverage probabilities'. *Metrologia*, 47:343–346, 2010.

[12] B. D. Hall. Assessing the performance of uncertainty calculations by simulation. In *Proceedings of 74th ARFTG Microwave Measurement Symposium (Dec. 1-4, 2009, Bloomfield, Co. USA)*, 2009.

[13] B. D. Hall. Evaluating methods of calculating measurement uncertainty. *Metrologia*, 45:L5–L8, 2008.

[14] BIPM, IEC, IFCC, 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 JCGM 101:2008*. BIPM Joint Committee for Guides in Metrology, Paris, Sèvres, 1 edition, 2008.

[15] I. Lira. On the long-run success rate of coverage intervals. *Metrologia*, 45:L21–L23, 2008.

[16] I. Lira. On the meaning of coverage probabilities. *Metrologia*, 46:616–618, 2009.

[17] Blaza Toman Antonio Possolo and Tyler Estler. Contribution to a conversation about the Supplement 1 to the GUM. *Metrologia*, 46:L1–L7, 2009.

[18] Rand R. Wilcox. *Basic statistics: understanding conventional methods and modern insights.* Oxford University Press, New York, 2009.

[19] C. M. Wang and Hari K. Iyer. Uncertainty analysis for vector measurands using fiducial inference. *Metrologia*, 43:486–494, 2006.

[20] R. Willink and B. D. Hall. A classical method for uncertainty analysis with multidimensional data. *Metrologia*, 39:361–369, 2002.

[21] Gerd Wübbeler, Peter M Harris, Maurice G Cox, and Clemens Elster. A two-stage procedure for determining the number of trials in the application of a Monte Carlo method for uncertainty evaluation. *Metrologia*, 47:317–324, 2010.

[22] Joseph Adler. *R in a nutshell: a desktop quick reference.* O'Reilly, Sebastopol, California, USA, 2010.

[23] Luc Devroye. *Non-Uniform Random Variate Generation.* Springer-Verlag, New York, 1986.

5 Appendix: Simulation Software

Computations in this work used the R statistical computing language and environment [2]. The data processing and propagation of uncertainty were carried out using a package of R functions that is freely available (on-line at `http://mst.irl.cri.nz`) [3]. The next section documents some other functions that were used to simulate measurement results and assess coverage in this study. A useful reference to the R language is [22].

The following aspects of R syntax may be of interest.

- Assignment is indicated by an arrow `<-` (rather than an equals sign `=`)

- A function is defined by assigning a generic function to a name. For instance, `sim.standard.u <- function(N,sigma,df){....}`

- Names may include a dot, e.g., `sim.standard.u` (in many programming languages, dots are operators with special meanings)

- Vectors can be defined by the concatenation function `c()`, e.g., `c(1,2,3)` defines a 3-element vector with components 1, 2 and 3

- A tilde `~` indicates an R formula, which expresses a relationship among variables. For example, `~x1 + x2` appeared in the definition of the intermediate uncertain number `y` in §2.2.

5.1 Simulation functions

5.1.1 Simulating standard uncertainty

The function `sim.standard.u`, below, returns an N-element vector of random variates distributed as

$$\sigma \sqrt{\frac{\chi_\nu^2}{\nu}} \, ,$$

where χ_ν^2 is a chi-square random variable with ν degrees of freedom.

```
sim.standard.u<-function(N,sigma,df)
{
    if(!is.infinite(df)) {
        u <- sigma*sqrt(
            rchisq(N,df)/df
        )
    } else {
        u <- rep(sigma,N)
    }
    return(u)
}
```

This function is used to simulate values of standard uncertainty when an influence quantity has finite degrees-of-freedom. It is used in the functions `sim.gaussian`, `sim.uniform` and `sim.arcsine` described below. These functions all return a list containing: a vector of N estimates x_i, a vector of N standard uncertainties u_i and a vector of N degrees-of-freedom ν_i (although the degrees-of-freedom is a constant).

5.1.2 Gaussian influences

The function `sim.gaussian` is used to simulate measurements of an influence quantity with a Gaussian error. The estimates are generated by a normal random number generator `rnorm`.

```
sim.gaussian<-function(N,mu,sigma,df)
{
    x <- rnorm(N,mu,sigma)
    u <- sim.standard.u(N,sigma,df)
    return( list(x=x,u=u,df=df) )
}
```

5.1.3 Uniform influences

The function `sim.uniform` is used to simulate measurements of an influence quantity
with a uniformly distributed error. The estimates are obtained from a uniform random
number generator `runif`.

```
sim.uniform<-function(N,mu,a,df){
    x <- mu + runif(N,-a,a)
    u <- sim.standard.u(N,a/sqrt(3),df)
    return( list(x=x,u=u,df=df) )
}
```

We note that the sampling distribution of the standard uncertainty `u` for a uniform
population is not correctly modeled by `sim.standard.u`. However, our usage here is
consistent with the way that the GUM treats finite degrees of freedom in non-Gaussian
influence quantities.

5.1.4 Arcsine influences

The function `sim.arcsine` is used to simulate measurements of an influence quantity
with an arcsine error distribution. The estimates are obtained by an inversion method
[23, p 481]. The comment above in §5.1.3 about simulating the sampling distribution
of the standard uncertainty with `sim.standard.u` also applies.

```
sim.arcsine <-function(N,mu,a,df) {
    x <- mu + a * sin(
        runif(N,-pi/2,pi/2)
    )
    u <- sim.standard.u(N,a/sqrt(2),df)
    return( list(x=x,u=u,df=df) )
}
```

5.1.5 Multivariate Gaussian influences

The function `sim.mvgaussian` is used to simulate measurements of a vector of influence
quantities associated with correlated Gaussian errors. The estimates are obtained from
a multivariate normal random number generator `rmvnorm` and stored in an array of
N M-component vectors (assigned to `x`, in the list returned). The associated variance-
covariance matrices are generated by a Wishart random number generator `rwish`, which
returns an array of N $M \times M$ matrices (assigned to `xcv`, in the list returned) [7, §4.4].

```
sim.mvgaussian<-function(N,mu,cv,df)
{
    M <- length(mu)
    if(is.infinite(df)) {
        x <- t(
            sample <- rmvnorm(N,mu,cv)
        )
        xcv <- array(cv,dim=c(M,M,N))
    } else {
        x <- t( rmvnorm(N,mu,cv) )
        xcv <- array(dim=c(M,M,N))
        for(i in 1:N)
            xcv[,,i]<-rwish(df,cv) / df
```

```
    }
    return( list(x=x,cv=xcv,df=df) )
}
```

5.1.6 Assessing coverage

The function `sim.successes` is used to assess how many uncertainty intervals cover the
measurand. The function takes four arguments: the measurand `mu` and three vectors
`x`, `u` and `df` containing the N estimates, standard uncertainties and degrees-of-freedom.
It returns an N-element Boolean vector that can be summed to yield the number of
successful uncertainty intervals (see §2.2).

```
#------------------------------------
# OK[i] is TRUE if the 95% uncertainty
# interval [ x[i]-U[i], x[i]+U[i] ]
# contains 'mu'
#
sim.successes <-function(mu,x,u,df){
    k <- qt(0.975,df)
    U <- u * k
    OK <- (x-U <= mu) & (x+U >= mu)
    return(OK)
}
```

6   Appendix: Gauge block example

The code in this appendix was used to obtain the results presented in §3.1. The first
section defines a small number of constants and the measurand for the purposes of
testing. The constant `zero` is a vector of N zeros. A number of error terms are then
generated. These correspond to the quantities in equation (5) that will be estimated as
zero during data processing. Next, estimates of $\overline{d}$ are generated (`d.bar.sim.x`). The
simulation section of the code concludes with the generation of independent estimates
of l_{s}, $\overline{\theta}$ and α_{s} (`ls.sim`, `theta.bar.sim` and `alphas.s.sim`, respectively).

```
#------------------------------------
# Constants
#
ls.const <- 50.000623E-3
d.bar.const <- 215E-9
theta.bar.const <- -0.1
alpha.s.const <- 11.5E-6
zero <- rep.int(0,N)

l.measurand <- (ls.const + d.bar.const)

#------------------------------------
# Simulation of errors
#
d0.sim <- sim.gaussian(N,0,5.8E-9,24)
d1.sim <- sim.gaussian(N,0,3.9E-9,5)
d2.sim <- sim.gaussian(N,0,6.7E-9,8)
delta.sim <- sim.arcsine(N,0,0.5,Inf)
d.alpha.sim <- sim.uniform(N,0,1E-6,50)
d.theta.sim <- sim.uniform(N,0,0.05,2)

#------------------------------------
d.bar.sim.x <-  (
    l.measurand - ls.const
+ l.measurand * (
        (alpha.s.const + d.alpha.sim$x)
    *  (theta.bar.const + delta.sim$x)
```

```
      )
-ls.const * (
        alpha.s.const * (
              theta.bar.const +
              delta.sim$x -
              d.theta.sim$x
        )
    )
+ d0.sim$x + d1.sim$x + d2.sim$x
)


#-------------------------------------
# Simulation of other influences
#
ls.sim <- sim.gaussian(
    N,ls.const,25E-9,18
)
theta.bar.sim <- sim.gaussian(
    N,theta.bar.const,0.2,Inf
)
alpha.s.sim <- sim.uniform(
    N,alpha.s.const,2E-6,Inf
)
```

The calculation of uncertainty follows. This code performs data processing on the information generated in the simulation. The inputs to the GUM calculation are defined as elementary uncertain numbers. The values of $\overline{d}$ are defined as a constant uncertain numbers (`d.bar`), with no uncertainty. This allows the estimates of d to be defined as an intermediate uncertain number with components of uncertainty associated with d_0, d_1 and d_2. An expression for l, `l.expr` (H.3 in the GUM), is then used to define an intermediate uncertain number for the estimate of the measurand. Finally, the coverage is reported as a percentage of N.

Note that some simulated quantities contribute to variability to the values of $\overline{d}$ (`d.bar.sim.x`), but are not used in the uncertainty calculation. For example, the estimates of d_0, d_1 and d_2 in the uncertainty calculation are always zero, because they represent unobservable errors with a zero mean. On the other hand, some quantities are fixed during the simulation of $\overline{d}$. Estimates of these quantities are obtained separately for use in the uncertainty calculation, by assuming that the underlying physical quantity is fixed but that estimates (measurements) of the quantity are made for the purposes of data processing.

```
#-------------------------------------
# GUM uncertainty calculation
# input definitions
#
ls <- elementary(
        'ls',
        ls.sim$x,
        ls.sim$u,
        ls.sim$df
)
d0 <- elementary(
        'd0',
        zero,
        d0.sim$u,
        d0.sim$df
    )
d1 <- elementary(
        'd1',
```

```
            zero,
            d1.sim$u,
            d1.sim$df
        )
d2 <- elementary(
            'd2',
            zero,
            d2.sim$u,
            d2.sim$df
        )
theta.bar <- elementary(
            'theta_bar',
            theta.bar.sim$x,
            theta.bar.sim$u
        )
delta <- elementary(
            'delta',
            zero,
            delta.sim$u
        )
d.alpha <- elementary(
            'd_alpha',
            zero,
            d.alpha.sim$u,
            d.alpha.sim$df
        )
alpha.s <- elementary(
            'alpha_s',
            alpha.s.sim$x,
            alpha.s.sim$u
        )
d.theta <- elementary(
            'd_theta',
            zero,
            d.theta.sim$u,
            d.theta.sim$df
        )
#-------------------------------------
# Simulated estimate of dbar
d.bar <- constant('dbar',d.bar.sim.x)

#-------------------------------------
# GUM uncertainty calculation
# intermediate results
#
d <- intermediate(
            'd',
            ~dbar-d0-d1-d2,
            list(d0,d1,d2,d.bar)
        )
theta <- intermediate(
            'theta',
            ~ theta_bar - delta,
            list(theta.bar,delta)
        )
#-------------------------------------
# GUM uncertainty calculation
# giving the final result
#
l.expr <- ~ ls  + d - ls * (
            d_alpha * theta +
            alpha_s * d_theta
        )
l <- intermediate(
```

```
        'l',
        l.expr,
        list(d,ls,d.alpha,theta,
            alpha.s,d.theta)
    )

cover <- 100 * sum(
    sim.successes(
        l.measurand,
        value(l),
        uncertainty(l),
        dof(l)
    )
) / N
```

The following was appended to include the two additional second-order terms in the uncertainty calculation.

```
#-------------------------------------
# Nonlinear terms are added here as
# independent contributions
#
nl1.u <- ls.sim$x *
            alpha.s.sim$u *
            d.theta.sim$u
nl1 <- elementary(
        'nl1',
        zero,
        nl1.u,
        d.theta.sim$df
    )
nl2.u <- ls.sim$x *
            d.alpha.sim$u *
            uncertainty(theta)
nl2 <- elementary(
        'nl2',
        rep(0,N),
        nl2.u,
        d.alpha.sim$df
    )
nl <- intermediate(
        'nl',
        ~l + nl1 + nl2,
        list(nl1,nl2,l)
    )
cover <- 100 * sum(
sim.successes(
        l.measurand,
        value(nl),
        uncertainty(nl),
        dof(nl)
    )
) / N
```

7   Appendix: Impedance measurement example

The code below was used to obtain the results presented in §3.2 for the impedance measurement problem. The method of assessing coverage is shown for the resistance R only, to save space. Three different methods were used to assign a number of degrees-of-freedom to the final results, as described in §3.2. In the last section of code, different

lines can be uncommented to implement the different methods.

```
#-----------------------------------
# Simulation model parameters
#
V.val <- 4.999 # V
I.val <- 19.661E-3 # amp
Phi.val <- 1.04446 # radian

u1 <- 3.2E-3
u2 <- 9.5E-6
u3 <- 7.5E-4

r12 <- -0.36
r13 <- 0.86
r23 <- -0.65

df <- 4

#-----------------------------------
# The 3 measurands are in a vector
#
mu <- c(
     V.val * cos(Phi.val) / I.val,
     V.val * sin(Phi.val) / I.val,
     V.val / I.val
     )

#-----------------------------------
# Simulation model
#
x_exp <- c(V.val,I.val,Phi.val)

cv_exp <- matrix(
     c(
     u1^2,u1*r12*u2,u1*r13*u3,
     u1*r12*u2,u2^2,u2*r23*u3,
     u1*r13*u3,u2*r23*u3,u3^2
     ),
     ncol=3
)

sim.data <- sim.mvgaussian(N,
        x_exp,cv_exp,df
     )
#-----------------------------------
# GUM calculation inputs
#
V.bar <- elementary(
        'V_bar',
        sim.data$x[1,],
        sqrt(sim.data$cv[1,1,])
     )
I.bar <- elementary(
        'I_bar',
        sim.data$x[2,],
        sqrt(sim.data$cv[2,2,])
        )
phi.bar <- elementary(
        'phi_bar',
        sim.data$x[3,],
        sqrt(sim.data$cv[3,3,])
     )
set.correlation(
```

```
    V.bar,
    I.bar,
    sim.data$cv[1,2,]/sqrt(
      sim.data$cv[1,1,]*sim.data$cv[2,2,]
    )
)
set.correlation(
    V.bar,
    phi.bar,
    sim.data$cv[1,3,]/sqrt(
      sim.data$cv[1,1,]*sim.data$cv[3,3,]
    )
)
set.correlation(
    I.bar,
    phi.bar,
    sim.data$cv[2,3,]/sqrt(
      sim.data$cv[2,2,]*sim.data$cv[3,3,]
    )
)
inputs <- list(V.bar,I.bar,phi.bar)

#-------------------------------------
# Final results
#
R <- intermediate(
        'R',
        ~ V_bar * cos(phi_bar) / I_bar,
        inputs
    )
X <- intermediate(
        'X',
        ~ V_bar * sin(phi_bar) / I_bar,
        inputs
    )
Z <- intermediate(
        'Z',
        ~ V_bar / I_bar,
        inputs
    )
#-------------------------------------
# Alternative coverage calculations
# df <- 4
# df <- Inf
# ignore.corr <- TRUE
# df <- dof(R,ignore.corr)
rate <- 100 * sum(
    sim.successes(
        mu[1],
        value(R),
        uncertainty(R,ignore.corr),
        df
    )
) / N
```