We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

5,900
Open access books available

145,000
International authors and editors

180M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS

BOOK
CITATION
INDEX

INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Monte Carlo Simulations Applied to Uncertainty in Measurement

Paulo Roberto Guimarães Couto,
Jailton Carreteiro Damasceno and
Sérgio Pinheiro de Oliveira

Additional information is available at the end of the chapter

1. Introduction

Metrology is the science that covers all theoretical and practical concepts involved in a measurement, which when applied are able to provide results with appropriate accuracy and metrological reliability to a given measurement process. In any area in which a decision is made from a measurement result, all attention is critical to the metrological concepts involved. For example, the control panels of an aircraft are composed by several instruments that must be calibrated to perform measurements with metrological traceability and reliability, influencing the decisions that the pilot will make during the flight. In this way, it is clear that concepts involving metrology and reliability of measurements must be well established and harmonized to provide reliability and quality for products and services.

In the last two decades, basic documents for the international harmonization of metrological and laboratorial aspects have been prepared by international organizations. Adoption of these documents helps the evolution and dynamics of the globalization of markets. The ISO IEC 17025:2005 standard [1], for example, describes harmonized policies and procedures for testing and calibration laboratories. The International Vocabulary of Metrology (VIM - JCGM 200:2012) presents all the terms and concepts involved in the field of metrology [2]. The JCGM 100:2008 guide (Evaluation of measurement data – Guide to the expression of uncertainty in measurement) provides guidelines on the estimation of uncertainty in measurement [3]. Finally, the JCGM 101:2008 guide (Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method) is responsible to give practical guidance on the application of Monte Carlo simulations to the estimation of uncertainty [4].

Measurement uncertainty is a quantitative indication of the quality of measurement results, without which they could not be compared between themselves, with specified reference values or to a standard. According to the context of globalization of markets, it is necessary to adopt a universal procedure for estimating uncertainty of measurements, in view of the need for comparability of results between nations and for a mutual recognition in metrology. The harmonization in this field is very well accomplished by the JCGM 100:2008. This document provides a full set of tools to treat different situations and processes of measurement. Estimation of uncertainty, as presented by the JCGM 100:2008, is based on the law of propagation of uncertainty (LPU). This methodology has been successfully applied for several years worldwide for a range of different measurements processes.

The LPU however do not represent the most complete methodology for the estimation of uncertainties in all cases and measurements systems. This is because LPU contains a few approximations and consequently propagates only the main parameters of the probability distributions of influence. Such limitations include for example the linearization of the measurement model and the approximation of the probability distribution of the resulting quantity (or measurand) by a Student's t-distribution using a calculated effective degrees of freedom.

Due to these limitations of the JCGM 100:2008, the use of Monte Carlo method for the propagation of the full probability distributions has been recently addressed in the supplement JCGM 101:2008. In this way, it is possible to cover a broader range of measurement problems that could not be handled by using the LPU alone. The JCGM 101:2008 provides especial guidance on the application of Monte Carlo simulations to metrological situations, recommending a few algorithms that best suit its use when estimating uncertainties in metrology.

2. Terminology and basic concepts

In order to advance in the field of metrology, a few important concepts should be presented. These are basic concepts that can be found on the International Vocabulary of Metrology (VIM) and are explained below.

Quantity. "Property of a phenomenon, body, or substance, where the property has a magnitude that can be expressed as a number and a reference". For example, when a cube is observed, some of its properties such as its volume and mass are quantities which can be expressed by a number and a measurement unit.

Measurand. "Quantity intended to be measured". In the example given above, the volume or mass of the cube can be considered as measurands.

True quantity value. "Quantity value consistent with the definition of a quantity". In practice, a true quantity value is considered unknowable, unless in the special case of a fundamental quantity. In the case of the cube example, its exact (or true) volume or mass cannot be determined in practice.

Measured quantity value. "Quantity value representing a measurement result". This is the quantity value that is measured in practice, being represented as a measurement result. The volume or mass of a cube can be measured by available measurement techniques.

Measurement result. "Set of quantity values being attributed to a measurand together with any other available relevant information". A measurement result is generally expressed as a single measured quantity value and an associated measurement uncertainty. The result of measuring the mass of a cube is represented by a measurement result: 131.0 g ± 0.2 g, for example.

Measurement uncertainty. "Non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used". Since the true value of a measurement result cannot be determined, any result of a measurement is only an approximation (or estimate) of the value of a measurand. Thus, the complete representation of the value of such a measurement must include this factor of doubt, which is translated by its measurement uncertainty. In the example given above, the measurement uncertainty associated with the measured quantity value of 131.0 g for the mass of the cube is 0.2 g.

Coverage interval. "Interval containing the set of true quantity values of a measurand with a stated probability, based on the information available". This parameter provides limits within which the true quantity values may be found with a determined probability (coverage probability). So for the cube example, there could be 95% probability of finding the true value of the mass within the interval of 130.8 g to 131.2 g.

3. The GUM approach on estimation of uncertainties

As a conclusion from the definitions and discussion presented above, it is clear that the estimation of measurement uncertainties is a fundamental process for the quality of every measurement. In order to harmonize this process for every laboratory, ISO (International Organization for Standardization) and BIPM (Bureau International des Poids et Mesures) gathered efforts to create a guide on the expression of uncertainty in measurement. This guide was published as an ISO standard – ISO/IEC Guide 98-3 "Uncertainty of measurement - Part 3: Guide to the expression of uncertainty in measurement" (GUM) – and as a JCGM (Joint Committee for Guides in Metrology) guide (JCGM 100:2008). This document provides complete guidance and references on how to treat common situations on metrology and how to deal with uncertainties.

The methodology presented by the GUM can be summarized in the following main steps:

a.   Definition of the measurand and input sources.

It must be clear to the experimenter what exactly is the measurand, that is, which quantity will be the final object of the measurement. In addition, one must identify all the variables that directly or indirectly influence the determination of the measurand. These variables are

known as the input sources. For example, Equation 1 shows a measurand y as a function of four different input sources: x_1, x_2, x_3 and x_4.

$$y = f\left(x_1,\ x_2,\ x_3,\ x_4\right) \tag{1}$$

b.   Modeling.

In this step, the measurement procedure should be modeled in order to have the measurand as a result of all the input sources. For example, the measurand y in Equation 1 could be modeled as in Equation 2.

$$y = \frac{x_1(x_2 + x_3)}{x_4^2} \tag{2}$$

Construction of a cause-effect diagram helps the experimenter to visualize the modeling process. This is a critical phase, as it defines how the input sources impact the measurand. A well defined model certainly allows a more realistic estimation of uncertainty, which will in-clude all the sources that impact the measurand.

c.   Estimation of the uncertainties of input sources.

This phase is also of great importance. Here, uncertainties for all the input sources will be estimated. According to the GUM, uncertainties can be classified in two main types: Type A, which deals with sources of uncertainties from statistical analysis, such as the standard devi-ation obtained in a repeatability study; and Type B, which are determined from any other source of information, such as a calibration certificate or obtained from limits deduced from personal experience.

Type A uncertainties from repeatability studies are estimated by the GUM as the standard deviation of the mean obtained from the repeated measurements. For example, the uncer-tainty $u(x)$ due to repeatability of a set of n measurements of the quantity x can be ex-pressed by $s(\bar{x})$ as follows:

$$u(x) = s\left(\bar{x}\right) = \frac{s(x)}{\sqrt{n}} \tag{3}$$

where $\bar{x}$ is the mean value of the repeated measurements, $s(x)$ is its standard deviation and $s(\bar{x})$ is the standard deviation of the mean.

Also, it is important to note that the estimation of uncertainties of the Type B input sources must be based on careful analysis of observations or in an accurate scientific judgment, us-ing all available information about the measurement procedure.

d.   Propagation of uncertainties.

The GUM uncertainty framework is based on the law of propagation of uncertainties (LPU). This methodology is derived from a set of approximations to simplify the calculations and is valid for a wide range of models.

According to the LPU approach, propagation of uncertainties is made by expanding the measurand model in a Taylor series and simplifying the expression by considering only the first order terms. This approximation is viable as uncertainties are very small numbers compared with the values of their corresponding quantities. In this way, treatment of a model where the measurand y is expressed as a function of N variables $x_1,..., x_N$ (Equation 4), leads to a general expression for propagation of uncertainties (Equation 5).

$$y = f(x_1, \ldots, x_N) \tag{4}$$

$$u_y^2 = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \right)^2 u_{x_i}^2 + 2 \sum_{i=1}^{N\text{-}1} \sum_{j=i+1}^{N} \left(\frac{\partial f}{\partial x_i} \right) \left(\frac{\partial f}{\partial x_j} \right) \text{cov}(x_i, x_j) \tag{5}$$

where u_y is the combined standard uncertainty for the measurand y and u_{x_i} is the uncertainty for the ith input quantity. The second term of Equation 5 is related to the correlation between the input quantities. If there is no supposed correlation between them, Equation 5 can be further simplified as:

$$u_y^2 = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \right)^2 u_{x_i}^2 \tag{6}$$

e.   Evaluation of the expanded uncertainty.

The result provided by Equation 6 corresponds to an interval that contains only one standard deviation (or approx. 68.2% of the measurements). In order to have a better level of confidence for the result, the GUM approach expands this interval by assuming a Student's t-distribution for the measurand. The effective degrees of freedom v_{eff} for the t-distribution can be estimated by using the Welch-Satterthwaite formula (Equation 7).

$$v_{\text{eff}} = \frac{u_y^4}{\sum_{i=1}^{N} \frac{u_{x_i}^4}{v_{x_i}}} \tag{7}$$

where v_{x_i} is the degrees of freedom for the ith input quantity.

The expanded uncertainty is then evaluated by multiplying the combined standard uncertainty by a coverage factor k that expands it to a coverage interval delimited by a t-distribution with a chosen level of confidence (Equation 8).

$$U_y = k u_y \tag{8}$$

4. The GUM limitations

As mentioned before, the approach to estimate measurement uncertainties using the law of propagation of uncertainties presented by the GUM is based on some assumptions, that are not always valid. These assumptions are:

• The model used for calculating the measurand must have insignificant non-linearity. When the model presents strong elements of non-linearity, the approximation made by truncation of the first term in the Taylor series used by the GUM approach may not be enough to correctly estimate the uncertainty output.

• Validity of the central limit theorem, which states that the convolution of a large number of distributions has a resulting normal distribution. Thus, it is assumed that the probability distribution of the output is approximately normal and can be represented by a t-distribution. In some real cases, this resulting distribution may have an asymmetric behavior or does not tend to a normal distribution, invalidating the approach of the central limit theorem.

• After obtaining the standard uncertainty by using the law of propagation of uncertainties, the GUM approach uses the Welch-Satterthwaite formula to obtain the effective degrees of freedom, necessary to calculate the expanded uncertainty. The analytical evaluation of the effective degrees of freedom is still an unsolved problem [5], and therefore not always adequate.

In addition, the GUM approach may not be valid when one or more of the input sources are much larger than the others, or when the distributions of the input quantities are not symmetric. The GUM methodology may also not be appropriate when the order of magnitude of the estimate of the output quantity and the associated standard uncertainty are approximately the same.

In order to overcome these limitations, methods relying on the propagation of distributions have been applied to metrology. This methodology carries more information than the simple propagation of uncertainties and generally provides results closer to reality. Propagation of distributions involves the convolution of the probability distributions of the input quantities, which can be accomplished in three ways: a) analytical integration, b) numerical integration or c) by numerical simulation using Monte Carlo methods. The GUM Supplement 1 (or JCGM 101:2008) provides basic guidelines for using the Monte Carlo simulation for the propagation of distributions in metrology. It is presented as a fast and robust alternative method for cases where the GUM approach fails. This method provides reliable results for a wider range of measurement models as compared to the GUM approach.

5. Monte Carlo simulation applied to metrology

The Monte Carlo methodology as presented by the GUM Supplement 1 involves the propagation of the distributions of the input sources of uncertainty by using the model to provide the distribution of the output. This process is illustrated in Figure 1 in comparison with the propagation of uncertainties used by the GUM.
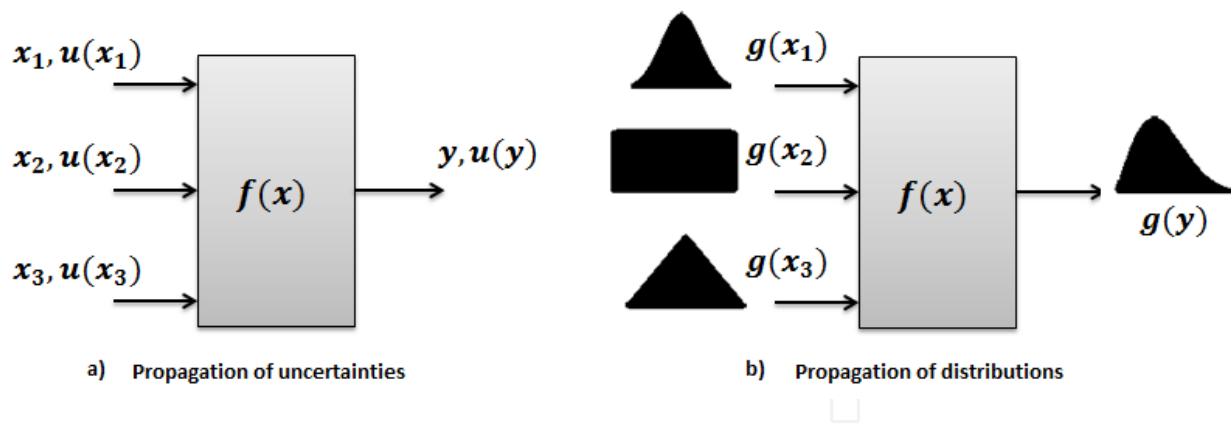
Figure 1. Illustrations of the methodologies of a) propagation of uncertainties, where x_1, x_2 and x_3 are input quantities, $u(x_1)$, $u(x_2)$ and $u(x_3)$ are their respective uncertainties and y and $u(y)$ are the measurand and its uncertainty, respectively; and b) propagation of distributions, where $g(x_1)$, $g(x_2)$ and $g(x_3)$ are the distribution functions of the input quantities and $g(y)$ is the distribution function of the measurand

Figure 1a) shows an illustration representing the propagation of uncertainties. In this case, three input quantities are presented x_1, x_2 and x_3 along with their respective uncertainties $u(x_1)$, $u(x_2)$ and $u(x_3)$. As can be noted, only the main moments (expectation and standard deviation) of the input quantities are used in the propagation and thus a certain amount of information is lost. When propagating distributions however, see Figure 1b), no approximations are made and the whole information contained on the input distributions are propagated to the output.

The GUM Supplement 1 provides a sequence of steps to be followed similarly as to what is done in the GUM:

a. definition of the measurand and input quantities;

b. modeling;

c. estimation of the probability density functions (PDFs) for the input quantities;

d. setup and run the Monte Carlo simulation;

e. summarizing and expression of the results.

The steps (a) and (b) are exactly the same as described in the GUM. Step (c) now involves the selection of the most appropriate probability density functions (or PDFs) for each of the input quantities. In this case, the maximum entropy principle used in the Bayesian theory can be applied in the sense one should consider the most generic distribution for the level of information that is known about the input source. In other words, one should select a PDF that does not transmit more information than that which is known. As an example, if the only information available on an input source is a maximum and a minimum limits, a uniform PDF should be used.

After all the input PDFs have been defined, a number of Monte Carlo trials should be selected – step (d). Generally, the greater the number of simulation trials, the greater the conver-

gence of the results. This number can be chosen *a priori* or by using an adaptive methodology. When choosing *a priori* trials, the GUM Supplement 1 recommends the selection of a number M of trials, according to the following general rule, in order to provide a reasonable representation of the expected result:

$$M > \frac{10^4}{1 - p} \tag{9}$$

where $100p\%$ is the selected coverage probability. So for example, when the chosen coverage probability is 95%, $p = 0.95$ and M should be at least higher than 200,000.

The adaptive methodology involves the selection of a condition to check after each trial for the stabilization of the results of interest. The results of interest in this case are the expectation (or mean) and the standard deviation of the output quantity and the endpoints of the chosen interval. According to the GUM Supplement 1, a result is considered to be stabilized if twice the standard deviation associated with it is less than the numerical tolerance associated with the standard deviation of the output quantity.

The numerical tolerance of an uncertainty, or standard deviation, can be obtained by expressing the standard uncertainty as $c \times 10^l$, where c is an integer with a number of digits equal to the number of significant digits of the standard uncertainty and l is an integer. Then the numerical tolerance δ is expressed as:

$$\delta = \frac{1}{2}10^l \tag{10}$$

The next step after setting M is to run the simulation itself. Despite the advantages discussed for the Monte Carlo numerical methodology for estimating measurement uncertainties, one of the main requirements for a reliable simulation is to have a good pseudo-random number generator. In this way, the GUM Supplement 1 recommends the use of the enhanced Wichmann-Hill algorithm [6].

Simulations can easily be setup to run even on low cost personal computers. Generally, a simulation for an average model with 200,000 iterations, which would generate reasonable results for a coverage probability of 95%, runs in a few minutes only, depending on the software and hardware used. In this way, computational costs are usually not a major issue.

The last stage is to summarize and express the results. According to the GUM Supplement 1, the following parameters should be reported as results: a) an estimate of the output quantity, taken as the average of the values generated for it; b) the standard uncertainty, taken as the standard deviation of these generated values; c) the chosen coverage probability (usually 95%); and d) the endpoints corresponding to the selected coverage interval.

The selection of this coverage interval should be done by determining: i) the probabilistically symmetric coverage interval, in the case of a resulting symmetric PDF for the output quantity; ii) the shortest $100p\%$ coverage interval, when the output PDF is asymmetrical.

6. Case studies: Fuel cell efficiency

In order to better understand the application of Monte Carlo simulations on the estimation of measurement uncertainty, some case studies will be presented and discussed along this chapter. The first example to be shown concerns the estimation of the real efficiency of a fuel cell. As discussed before, the first steps are to define the measurand and the input sources, as well as a model associated with them.

Fuel cells are electrochemical devices that produce electrical energy using hydrogen gas as fuel [7]. The energy production is a consequence of the chemical reaction of a proton with oxygen gas yielding water as output. There is also the generation of heat as a byproduct which could be used in cogeneration energy processes enhancing the overall energy efficiency. Two kinds of fuel cells are most used currently: PEMFC (proton exchange membrane fuel cell) and SOFC (oxide solid fuel cell). The former is used in low temperature applications (around 80 °C) and the last is used in the high temperature range (near 1000 °C).

One of the most important parameters to be controlled and measured in a fuel cell is its energy efficiency. To do so, it is necessary to know both the energy produced by the cell and the energy generated by the chemical reaction. The thermodynamic efficiency of a fuel cell can be calculated by Equation 11 [8].

$$\eta_t = \frac{\Delta G}{\Delta H} \tag{11}$$

where η_t is the thermodynamic efficiency, ΔG is the maximum energy produced by the fuel cell (Gibb's free energy in kJ/mol) and ΔH is the energy generated by the global reaction (or the enthalpy of formation in kJ/mol). However, in order to calculate the real efficiency of a fuel cell Equation 12 is necessary [8].

$$\eta_R = \eta_t \frac{E_R}{E_I} = \frac{\Delta G}{\Delta H} \frac{E_R}{E_I} \tag{12}$$

where η_R is the real efficiency, E_R is the real electric voltage produced by the fuel cell (V) and E_I is the ideal electric voltage of the chemical reaction (V).

In this case study, the real efficiency will be considered as the measurand, using Equation 12 as its model. The values and sources of uncertainty of inputs for a fuel cell operating with pure oxygen and hydrogen at standard conditions have been estimated from data cited in the literature [8, 9]. They are as follows:

Gibbs free energy (ΔG). The maximum free energy available for useful work is of 237.1 kJ/mol. In this example, one can suppose an uncertainty of 0.1 kJ/mol as a poor source of information, i.e. no probability information is available within the interval ranging from 237.0 kJ/mol to 237.2 kJ/mol. Thus, a uniform PDF can be associated with this input source using these values as minimum and maximum limits, respectively.

Enthalpy of formation (ΔH). The chemical energy, or enthalpy of formation, for the oxygen/hydrogen reaction at standard conditions is given as 285.8 kJ/mol. Again, considering

an uncertainty of 0.1 kJ/mol, a uniform PDF can be associated with this input source using 285.7 kJ/mol and 285.9 kJ/mol as minimum and maximum limits, respectively.

Ideal voltage (E_I). The ideal voltage of a fuel cell operating reversibly with pure hydrogen and oxygen in standard conditions is 1.229 V (Nernst equation). It is possible to suppose an uncertainty of ± 0.001 V as a poor source of information, and have a uniform PDF associated with that input source in this interval.

Real voltage (E_R). The real voltage was measured as 0.732 V with a voltmeter that has a digital resolution of ± 0.001 V. The GUM recommends that half the digital resolution can be used as limits of a uniform distribution. Then, a uniform PDF can be associated with this input source in the interval of 0.7315 V to 0.7325 V.

Figure 2 shows the cause-effect diagram for the evaluation of the real efficiency and Table 1 summarizes the input sources and values. All input sources are considered to be type B sources of uncertainty since they do not come from statistical analysis. In addition, they are supposed to be non-correlated.
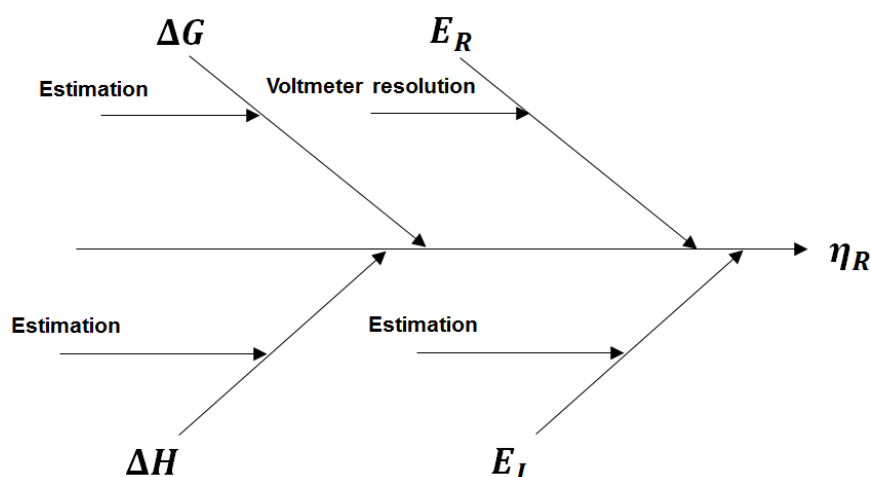
Figure 2. Cause-effect diagram for the fuel cell thermodynamic efficiency uncertainty estimation

Input source	Type	PDF	PDF parameters
Gibbs free energy (ΔG)	B	Uniform	Min: 237.0 kJ/mol; Max: 237.2 kJ/mol
Enthalpy of formation (ΔH)	B	Uniform	Min: 285.7 kJ/mol; Max: 285.9 kJ/mol
Ideal voltage (E_I)	B	Uniform	Min: 1.228 V; Max: 1.230 V
Real voltage (E_R)	B	Uniform	Min: 0.7315 V; Max: 0.7325 V

Table 1. Input sources and associated PDFs with their respective parameters for the estimation of uncertainty of the real efficiency of a fuel cell (Min stands for minimum value and Max stands for maximum value)

Monte Carlo simulation was set to run $M = 2 \times 10^5$ trials of the proposed model, using the described input sources. The final histogram representing the possible values for the real efficiency of the cell is shown on Figure 3. Table 2 shows the statistical parameters obtained

for the final PDF corresponding to the histogram. The low and high endpoints represent the 95% coverage interval for the final efficiency result of 0.49412.
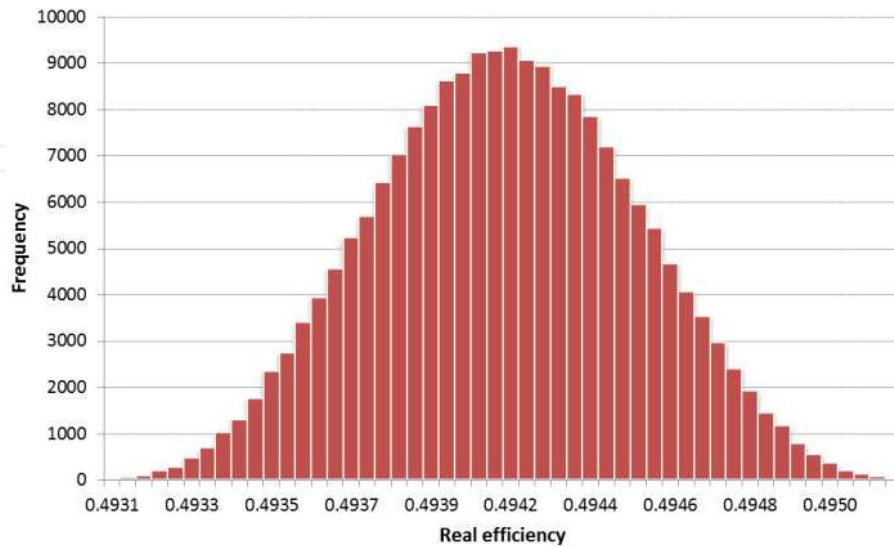

Figure 3. Histogram representing the resulting PDF for the real efficiency of a fuel cell estimated by Monte Carlo simulation

Parameter	Value
Mean	0.49412
Standard deviation	0.00034
Low endpoint for 95%	0.49346
High endpoint for 95%	0.49477

Table 2. Statistical parameters obtained for the Monte Carlo simulation of the real efficiency estimation model

In order to have a comparison with the traditional GUM methodology, Table 3 is shown with the results obtained by using the LPU. The number of effective degrees of freedom is infinite because all the input sources are uncertainties of type B. Consequently, for a coverage probability of 95%, the coverage factor obtained from a t-distribution is 1.96. It can be noted that the values obtained for the standard deviation of the resulting PDF (from the Monte Carlo simulation) and for the standard uncertainty (from the LPU methodology) are practically the same.

Parameter	Value
Combined standard uncertainty	0.00034
Effective degrees of freedom	∞
Coverage factor (k)	1.96
Expanded uncertainty	0.00067

Table 3. Results obtained for the real efficiency of a fuel cell uncertainty estimation using the GUM uncertainty approach, with a coverage probability of 95%

Even though results from both methodologies are practically the same, the GUM Supplement 1 provides a practical way to validate the GUM methodology with the Monte Carlo simulation results. This will be shown in detail in the next case studies.

7. Case studies: Measurement of torque

Torque is by definition a quantity that represents the tendency of a force to rotate an object about an axis. It can be mathematically expressed as the product of a force and the lever-arm distance. In metrology, a practical way to measure it is by loading the end of a horizontal arm with a known mass while keeping the other end fixed (Figure 4).
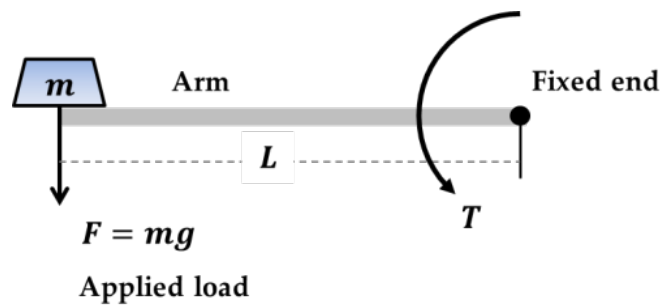
Figure 4. Conceptual illustration of the experimental setup for a measurement of torque (T), where F is the applied force, m is the mass of the load, g is the local gravity acceleration, L is the length of the arm

The model to describe the experiment can be expressed as follows:

$$T = mgL \tag{13}$$

where T is the torque, m is the mass of the applied load (kg), g is the local gravity acceleration (m/s²) and L is the total length of the arm (m). In practice, one can imagine several more sources of uncertainty for the experiment, like for example the thermal dilatation of the arm as the room temperature changes. However, the objective here is not to exhaust all the possibilities, but instead, to provide basic notions of how to use the Monte Carlo methodology for uncertainty estimation based on a simple model. In this way, only the following sources will be considered:

Mass (m). In the example, the mass m was repeatedly measured for ten times in a calibrated balance, with a capacity of 60 kg. The average mass was 35.7653 kg, with a standard deviation of 0.3 g. This source of uncertainty is purely statistical and is classified as being of type A according to the GUM. The PDF that best represents this case is a Gaussian distribution, with mean of 35.7653 kg and standard deviation equal to the standard deviation of the mean, i.e. $0.3\,\mathrm{g}/\sqrt{10} = 9.49 \times 10^{-5}$ kg.

In addition, the balance used for the measurement has a certificate stating an expanded uncertainty for this range of mass of 0.1 g, with a coverage factor $k = 2$ and a level of confidence of 95%. The uncertainty of the mass due to the calibration of the balance constitutes another source of uncertainty involving the same input quantity (mass). In this case, a Gaussian distribution can also be used as PDF to represent the input uncertainty, with mean of zero and standard deviation of $U / k = 0.1$ g $/ 2 = 0.00005$ kg, i.e. the expanded uncertainty divided by the coverage factor, resulting in the standard uncertainty. The use of zero as the mean value is a mathematical artifice to take into account the variability due to this source of uncertainty without changing the value of the quantity (mass) used in the model. More on this will be discussed later.

Local gravity acceleration (g). The value for the local gravity acceleration is stated in a certificate of measurement as 9.7874867 m/s², as well as its expanded uncertainty of 0.0000004 m/s², for $k = 2$ and 95% of confidence. Again, a Gaussian distribution is used as the PDF representing this input source, with mean of 9.7874867 m/s² and standard deviation of $U / k = (0.0000004$ m/s²$) / 2 = 0.0000002$ m/s².

Length of the arm (L). The arm used in the experiment has a certified value for its total length of 1999.9955 mm, and its calibration certificate states an expanded uncertainty of 0.0080 mm, for $k = 2$ and 95% of confidence. The best PDF in this case is a Gaussian distribution with mean 1999.9955 mm and standard deviation of $U / k = 0.0080$ mm $/ 2 = 0.000004$ m.

Figure 5 illustrates the cause-effect diagram for the model of torque measurement. Note that there are three input quantities in the model, but four input sources of uncertainty, being one of the input quantities, the mass, split in two sources: one due to the certificate of the balance and other due to the measurement repeatability.

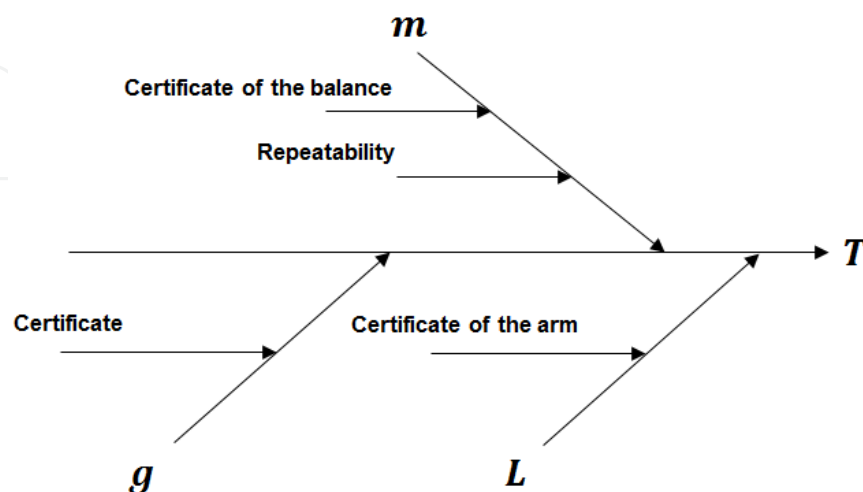Table 4 summarizes all the input sources and their respective associated PDFs.

Figure 5. Cause-effect diagram illustrating the proposed model for the measurement of torque

Input source	Type	PDF	PDF parameters
Mass (m)			
– due to repeatability	A	Gaussian	Mean: 35.7653 kg; SD: 9.49×10^{-5} kg
– due to certificate	B	Gaussian	Mean: 0 kg; SD: 0.00005 kg
Local gravity acceleration (g)	B	Gaussian	Mean: 9.7874867 m/s²; SD: 0.0000002 m/s²
Length of the arm (L)	B	Gaussian	Mean: 1999.9955 m; SD: 0.000004 m

Table 4. Input sources and associated PDFs with their respective parameters for the estimation of uncertainty of torque measurement (SD stands for standard deviation)

Running the Monte Carlo simulation using $M = 2 \times 10^5$ trials leads to results shown on Figure 6 and Table 5. It is important to note that the two input sources due to the mass are added together in the model, in order to account for the variability of both of them. Figure 6 shows the histogram constructed from the values of torque obtained by the trials. Table 5 contains the statistical parameters corresponding to the histogram. The low and high endpoints represent the 95% coverage interval for the final torque result of 700.1034 N.m.
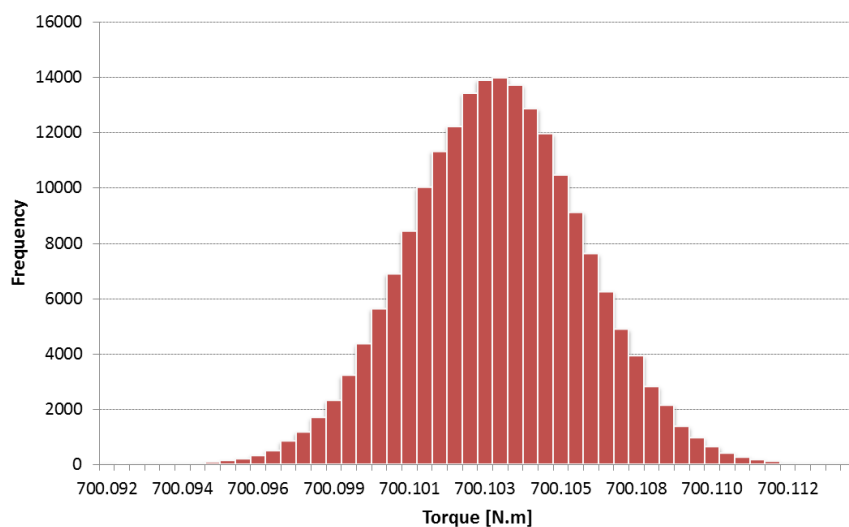

Figure 6. Histogram representing the resulting PDF for the torque measurement estimated by Monte Carlo simulation

Parameter	Value
Mean	700.1032 N.m
Standard deviation	0.0025 N.m
Low endpoint for 95%	700.0983 N.m
High endpoint for 95%	700.1082 N.m

Table 5. Statistical parameters obtained for the Monte Carlo simulation of the torque measurement model

Once more a comparison with the GUM approach is done and the results obtained by this methodology are shown on Table 6, for a coverage probability of 95%.

Parameter	Value
Combined standard uncertainty	0.0025 N.m
Effective degrees of freedom	30
Coverage factor (k)	1.96
Expanded uncertainty	0.0050 N.m

Table 6. Results obtained for the torque model using the GUM uncertainty approach, with a coverage probability of 95%

As commented before, the GUM Supplement 1 presents a procedure on how to validate the LPU approach addressed by de GUM with the results from Monte Carlo simulation. This is accomplished by comparing the low and high endpoints obtained from both methods. Thus, the absolute differences d_{low} and d_{high} of the respective endpoints of the two coverage intervals are calculated (Equations 14 and 15) and compared with the numerical tolerance δ of the standard uncertainty defined by Equation 10. If both d_{low} and d_{high} are lesser than δ, the GUM approach can be validated in this instance.

$$d_{low} = |\, y - U(y) - y_{low} \,| \tag{14}$$

$$d_{high} = |\, y + U(y) - y_{high} \,| \tag{15}$$

where y is the measurand estimate, $U(y)$ is the expanded uncertainty obtained by the GUM approach and y_{low} and y_{high} are the low and high endpoints of the PDF obtained by the Monte Carlo simulation for a given coverage probability, respectively.

In the case of the torque example, d_{low} and d_{high} are respectively calculated as 9.93×10^{-6} N.m and 4.12×10^{-6} N.m. Also, to obtain δ, the standard uncertainty 0.0025 N.m can be written as 25×10^{-4} N.m, considering two significant digits, then $\delta = 1/2 \times 10^{-4}$ N.m $= 5 \times 10^{-5}$ N.m. As both d_{low} and d_{high} are lesser than δ, the GUM approach is validated in this case.

When working with cases where the GUM approach is valid, like the example given for the measurement of torque, the laboratory can easily continue to use it for its daily uncertainty estimations. The advantages of the GUM traditional approach is that it is a popular widespread and recognized method, that does not necessarily require a computer or a specific software to be used. In addition, several small laboratories have been using this method since its publication as an ISO guide. It would be recommended however that at least one Monte Carlo run could be made to verify its validity, according to the criteria established (numerical tolerance). On the other hand, Monte Carlo simulations can provide reliable results on a wider range of cases, including those where the GUM approach fails. Thus, if their use for the laboratory would not increase the overall efforts or costs, then it would be recommended.

Now, extending the torque measurement case further, one can suppose that the arm used in the experiment has no certificate of calibration, indicating its length value and uncertainty, and that the only measuring method available for the arm's length is by the use of a ruler

with a minimum division of 1 mm. The use of the ruler leads then to a measurement value of 2000.0 mm for the length of the arm. Though, in this new situation very poor information about the measurement uncertainty of the arm's length is available. As the minimum division of the ruler is 1 mm, one can assume that the reading can be done with a maximum accuracy of up to 0.5 mm, which can be thought as an interval of ± 0.5 mm as limits for the measurement. However, no information of probabilities within this interval is available, and therefore the only PDF that can be assumed in this case is a uniform distribution, on which there is equal probability for the values within the whole interval. The uniform PDF then has 1999.5 mm as lower limit and 2000.5 mm as higher limit.

After running the Monte Carlo simulation with $M = 2 \times 10^5$ trials in this new situation, new results are obtained, as shown in Figure 7 (histogram) and Table 7 (statistical parameters).
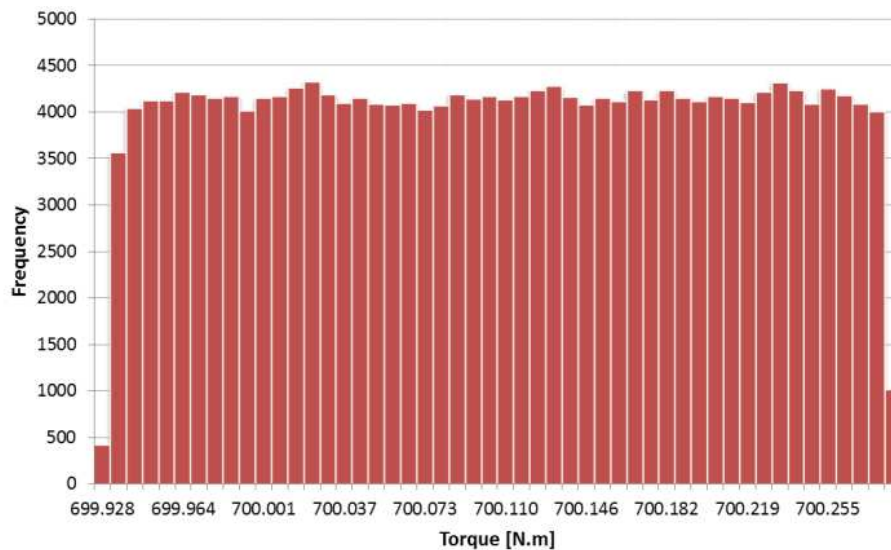

Figure 7. Histogram representing the resulting PDF for the torque measurement estimated by Monte Carlo simulation (measuring the non-certified arm with a ruler)

As can be noted, the resulting PDF changed completely, from a Gaussian-like shape (Figure 6) to an almost uniform shape (Figure 7). This is a consequence of the relatively higher uncertainty associated with the arm length in the new situation, as well as the fact that the PDF used for it was a uniform distribution. Thus, the strong influence of this uniform source was predominant in the final PDF. It is important to note that in the GUM methodology this new PDF would be approximated to a t-distribution, which has a very different shape.

Also as a result, the new standard uncertainty, represented by the standard deviation on Table 7 (0.1011 N.m), is higher than the one found in the former simulation, as shown in Table 5 (0.0025 N.m).

Estimating the uncertainty in this new situation by the traditional GUM approach, i.e. using the LPU and the Welch-Satterthwaite formula, one can obtain the results shown on Table 8.

Parameter	Value
Mean	700.1035 N.m
Standard deviation	0.1011 N.m
Low endpoint for 95%	699.9370 N.m
High endpoint for 95%	700.2695 N.m

Table 7. Statistical parameters obtained for the Monte Carlo simulation of the torque measurement model (measuring the non-certified arm with a ruler)

Parameter	Value
Combined standard uncertainty	0.1011 N.m
Effective degrees of freedom	∞
Coverage factor (k)	1.96
Expanded uncertainty	0.1981 N.m

Table 8. Results obtained for the torque model by using the GUM uncertainty approach and measuring the non-certified arm with a ruler, with a coverage probability of 95%

In this new situation, $d_{low} = 3.18 \times 10^{-2}$ N.m and $d_{high} = 3.18 \times 10^{-2}$ N.m, and the standard uncertainty 0.1011 N.m can be written as 10×10^{-2} N.m, considering two significant digits, then $\delta = 1/2 \times 10^{-2}$ N.m $= 5 \times 10^{-3}$ N.m. Thus, as both d_{low} and d_{high} are higher than δ, the GUM approach is not validated in this case. Note that considering only one significant digit, i.e. using a less rigid criterion, $\delta = 1/2 \times 10^{-1}$ N.m $= 5 \times 10^{-2}$ N.m and the GUM approach is validated.

8. Case studies: Preparation of a standard cadmium solution

This example is quoted from the EURACHEM/CITAC Guide [10] (Example A1) and refers to the preparation of a calibration solution of cadmium. In this problem, a high purity metal (Cd) is weighted and dissolved in a certain volume of liquid solvent. The proposed model for this case is shown on Equation 16.

$$c_{Cd} = \frac{1000mP}{V} \tag{16}$$

where c_{Cd} is the cadmium concentration (mg/L), m is the mass of the high purity metal (mg), P is its purity and V is the volume of the final solution (mL). The factor 1000 is used to convert milliliter to liter.

The sources of uncertainty in this case are identified as follows:

Purity (P). The purity of cadmium is quoted in the supplier's certificate as being 99.99% ± 0.01%. Thus, the value of P is 0.9999 and its uncertainty can be only be assumed to be in a uniform PDF as there is no extra information from the manufacturer concerning the probabilities within the interval. In this case the uniform PDF would have maximum and minimum limits of ± 0.0001, i.e. P would range from 0.9998 to 1.0000.

Mass (m). The mass of metal is obtained from its weighting in a certified balance. The value of mass for Cd is obtained as $m = 0.10028$ g. The uncertainty associated with the mass of the cadmium is estimated, using the data from the calibration certificate and the manufacturer's recommendations on uncertainty estimation, as 0.05 mg. As this is provided as a standard uncertainty in this example, a Gaussian PDF can be assumed with mean 0.10028 g and standard deviation of 0.05 g.

Volume (V). The total volume of solution is measured by filling a flask of 100 mL and has three major influences: calibration of the flask, repeatability and room temperature.

The first input source is due to filling of the flask, which is quoted by the manufacturer to have a volume of 100 ml ± 0.1 mL measured at a temperature of 20 °C. Again, poor information about this interval is available. In this particular case, the EURACHEM guide considers that it would be more realistic to expect that values near the bounds are less likely than those near the midpoint, and thus assumes a triangular PDF for this input source, ranging from 99.9 mL to 100.1 mL, with an expected value of 100.0 mL.

The uncertainty due to repeatability can be estimated as a result of variations in filling the flask. This experiment has been done and a standard uncertainty has been obtained as 0.02 mL. A Gaussian PDF is then assumed to represent this input source, with mean equal to zero and standard deviation of 0.02 mL.

The last input source for the volume is due to the room temperature. The manufacturer of the flask stated that it is calibrated for a room temperature of 20 °C. However the temperature of the laboratory in which the solution was prepared varies between the limits of ± 4 °C. The volume expansion of the liquid due to temperature is considerably larger than that for the flask, thus only the former is considered. The coefficient of volume expansion of the solvent is 2.1×10^{-4} °C^{-1}, which leads to a volume variation of ± (100 mL × 4 °C × 2.1×10^{-4} °C^{-1}) = ± 0.084 mL. So, as this is also a source with poor information, a uniform PDF is assumed in this interval, ranging from -0.084 mL to 0.084 mL.

A simple cause-effect diagram for this case study is presented in Figure 8 and the input sources are summarized in Table 9.
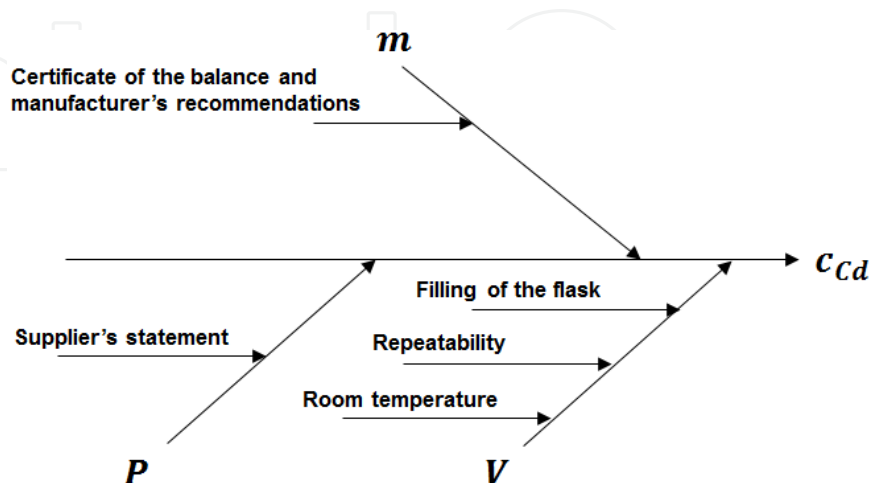

Figure 8. Cause-effect diagram for the preparation of a standard cadmium solution

Input source	Type	PDF	PDF parameters
Purity (P)	B	Uniform	Min: 0.9998; Max: 1.0000
Mass (m)	B	Gaussian	Mean: 100.28 mg; SD: 0.05 mg
Volume (V)			
– due to filling	B	Triangular	Mean: 100 mL; Min: 99.9 mL; Max: 100.1 mL
– due to repeatability	A	Gaussian	Mean: 0 mL; SD: 0.02 mL
– due to temperature	B	Uniform	Min: -0.084 mL; Max: 0.084 mL

Table 9. Input sources and associated PDFs with their respective parameters for the estimation of uncertainty of the concentration of a Cd standard solution (SD means standard deviation, Min stands for minimum value and Max stands for maximum value)

Monte Carlo simulation was done using $M = 2 \times 10^5$ trials and the results are shown in Figure 9 (histogram) and Table 10 (statistical parameters).
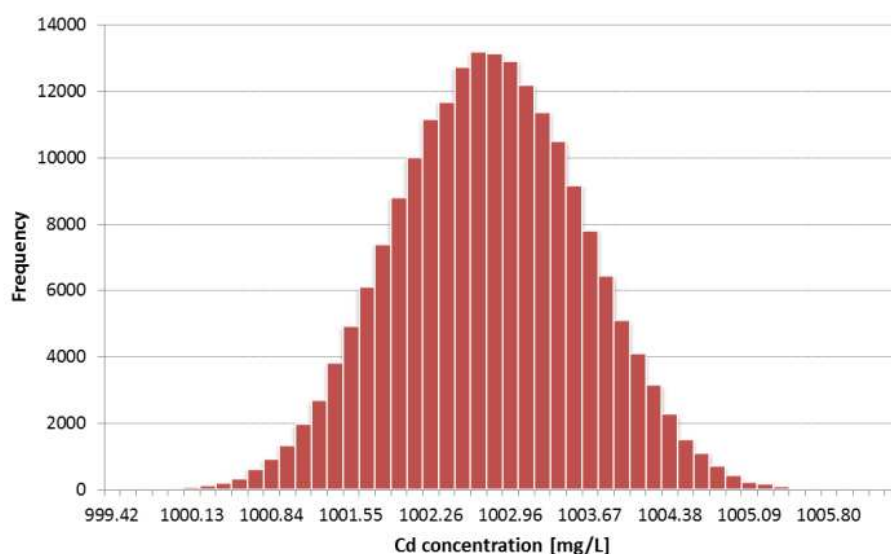

Figure 9. Histogram representing the resulting PDF for the concentration of a Cd standard solution estimated by Monte Carlo simulation

Parameter	Value
Mean	1002.705 mg/L
Standard deviation	0.835 mg/L
Low endpoint for 95%	1001.092 mg/L
High endpoint for 95%	1004.330 mg/L

Table 10. Statistical parameters obtained for the Monte Carlo simulation of the concentration of a Cd standard solution

Again, a comparison is made to the results found when using the GUM approach for a coverage probability of 95% (Table 11). Combined standard uncertainty (GUM approach) and standard deviation (Monte Carlo simulation) have practically the same value.

Parameter	Value
Combined standard uncertainty	0.835 mg/L
Effective degrees of freedom	1203
Coverage factor (k)	1.96
Expanded uncertainty	1.639 mg/L

Table 11. Results obtained for the concentration of a Cd standard solution model by using the GUM uncertainty approach, with a coverage probability of 95%

The endpoints obtained from both methods were compared using the numerical tolerance method proposed in the GUM Supplement 1. In this case, $d_{low} = 3.13 \times 10^{-2}$ mg/L and $d_{high} = 7.90 \times 10^{-3}$ mg/L, and writing the standard uncertainty as 84×10^{-2} mg/L (using two significant digits), $\delta = \frac{1}{2} \times 10^{-2}$ mg/L $= 5 \times 10^{-3}$ mg/L, which is lower than both d_{low} and d_{high}, and then the GUM approach is not validated.

9. Case studies: Measurement of Brinell hardness

The last example to be presented in this chapter will show a simple model for the measurement of Brinell hardness. This test is executed by applying a load on a sphere made of a hard material over the surface of the test sample (Figure 10).
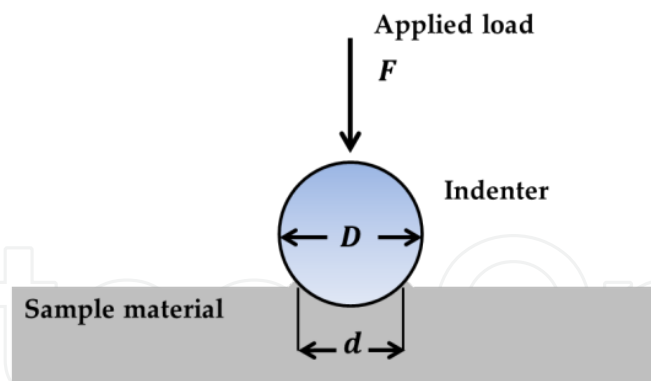

Figure 10. Schematic representation of a Brinell hardness test

During the test the sphere will penetrate through the sample leaving an indented mark upon unloading. The diameter of this mark is inversely proportional to the hardness of the material of the sample.

The model used here for the Brinell hardness (HB) is represented in Equation 17:

$$HB = \frac{0.204F}{\pi D \left(D - \sqrt{D^2 - d^2} \right)} \tag{17}$$

where F is the applied load (N), D is the indenter diameter (mm) and d is the diameter of the indentation mark (mm).

The input sources of uncertainty for this case study are:

Load (F). A fixed load is applied by the hardness test machine and is indicated as being 29400 N. The certificate of the machine indicates an expanded uncertainty of 2%, with $k = 2$ and a coverage probability of 95%. The best distribution to use in this case is a Gaussian PDF with mean 29400 N and standard deviation of $U / k = (0.02 \times 29400 \text{ N}) / 2 = 294$ N.

Indenter diameter (D). The sphere used as indenter has a certificate of measurement for its diameter with the value of 10 mm. Its expanded uncertainty, as indicated in the certificate, is 0.01 mm, for $k = 2$ and a coverage probability of 95%. Again, a Gaussian PDF should be used, with a mean of 10 mm and standard deviation of $U / k = 0.01 \text{ mm} / 2 = 0.005$ mm.

Diameter of the mark (d). The diameter of the indented mark was measured 5 times with the help of an optical microscope and a stage micrometer. The mean value was 3 mm, with a standard deviation of 0.079 mm. Besides the contribution due to repeatability, one could also consider the influence from the calibrated stage micrometer, but for the sake of simplicity, this source will be neglected. Thus, in this case a Gaussian PDF with mean 3 mm and standard deviation of $0.079 \text{ mm} / \sqrt{10} = 0.035$ mm would best represent the diameter of the mark.

Figure 11 shows the cause-effect diagram for this case study and Table 12 summarizes all the input sources of uncertainty.
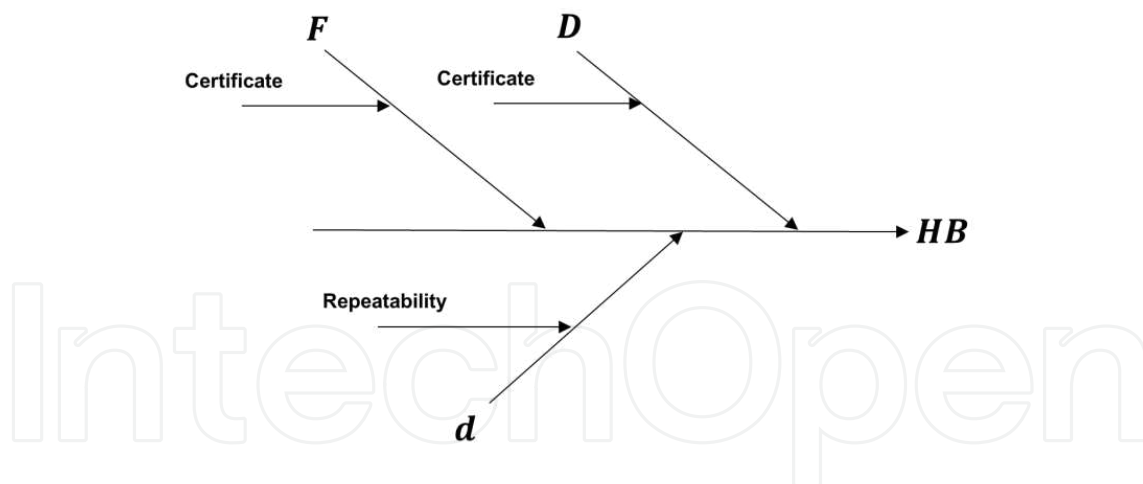

Figure 11. Cause-effect diagram for the measurement of Brinell hardness

Input source	Type	PDF	PDF parameters
Load (F)	B	Gaussian	Mean 29400 N; SD: 294 N
Indenter diameter (D)	B	Gaussian	Mean: 10 mm; SD: 0.005 mm
Diameter of the mark (d)	A	Gaussian	Mean: 3 mm; SD: 0.035 mm

Table 12. Input sources and associated PDFs with their respective parameters for the estimation of uncertainty of a Brinell hardness test (SD means standard deviation)

Monte Carlo simulation results for $M = 2 \times 10^5$ trials are shown in Table 13, as well as the resulting PDF for the Brinell hardness values in Figure 12.
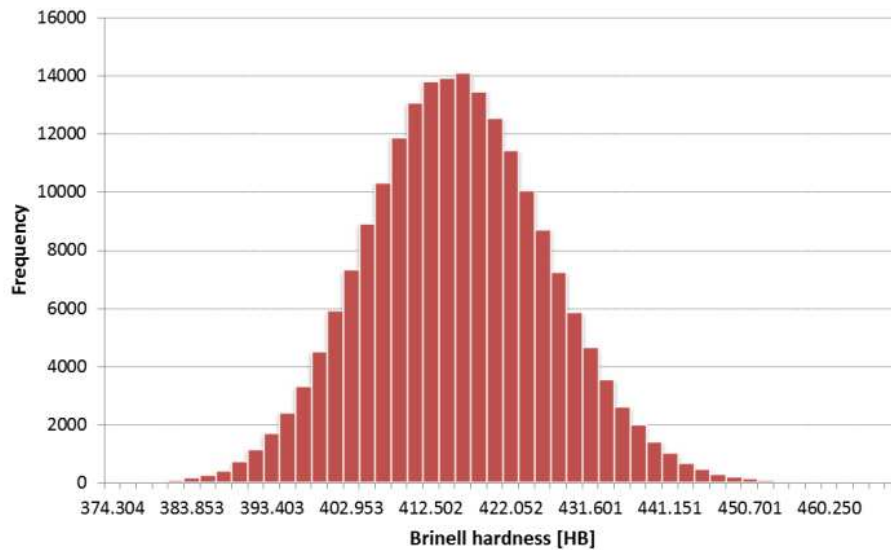

Figure 12. Histogram representing the resulting PDF for the Brinell hardness estimated by Monte Carlo simulation

Parameter	Value
Mean	415 HB
Standard deviation	11 HB
Low endpoint for 95%	394 HB
High endpoint for 95%	436 HB

Table 13. Statistical parameters obtained for the Monte Carlo simulation of the Brinell hardness measurement model

Results obtained by using the GUM approach in this case are show on Table 14.

Parameter	Value
Combined standard uncertainty	11 HB
Effective degrees of freedom	5
Coverage factor (k)	2.57
Expanded uncertainty	28 HB

Table 14. Results obtained for the Brinell hardness model by using the GUM uncertainty approach, with a coverage probability of 95%

Although the values of combined standard uncertainty (from the GUM approach) and standard deviation (from Monte Carlo simulation) are practically the same, the GUM approach is not validated using the numerical tolerance methodology. In this case, $\delta = 0.5$ HB, and the differences $d_{low} = 7.3$ HB and $d_{high} = 5.9$ HB, which are higher than δ.

The main difference of this case study to the others presented in this chapter is due to the non-linear character of the Brinell hardness model, which can lead to strong deviations from the GUM traditional approach of propagation of uncertainties. In fact, Monte Carlo simulation methodology is able to mimic reality better in such cases, providing richer information about the measurand than the GUM traditional approach and its approximations. In order to demonstrate this effect, one can suppose that the standard deviation found for the values of diameter of the indenter mark were 10 times higher, i.e. 0.35 mm instead of 0.035 mm. Then, Monte Carlo simulation would have given the results shown on the histogram of Figure 13, with statistical parameters shown on Table 15, for $M = 2 \times 10^5$ trials.
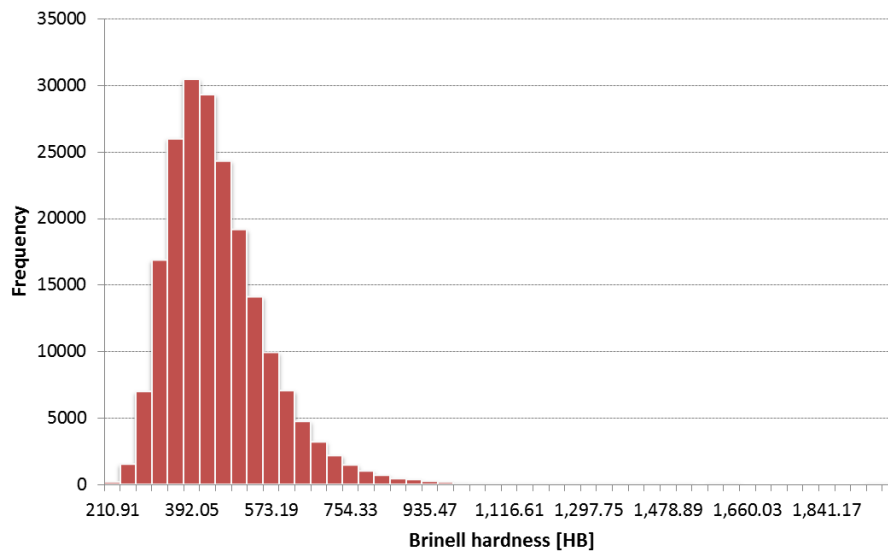

Figure 13. Histogram representing the resulting PDF for the Brinell hardness estimated by Monte Carlo simulation (with 0.35 mm instead of 0.035 mm as standard deviation of d)

Parameter	Value
Mean	433 HB
Median	414 HB
Standard deviation	114 HB
Low endpoint for 95%	270 HB
High endpoint for 95%	708 HB

Table 15. Statistical parameters obtained for the Monte Carlo simulation of the Brinell hardness measurement model (with 0.35 mm instead of 0.035 mm as standard deviation of d)

As can be observed, the resulting PDF for the Brinell hardness is strongly skewed to the left (lower values of HB). This behavior is a consequence of having a predominant uncertainty component (or input source) in a model that has non-linear characteristics. The unusual shape of this PDF gives an idea of how much the GUM approach can misestimate the coverage interval of the measurand when supposing a t-distribution as its final distribution.

It also can be noted that the expected mean calculated for the PDF (433 HB) is shifted to higher values in this case if compared with the expected value of 415 HB found in the former simulation (or by direct calculation of *HB*). Table 15 also shows the value of the median (the value that equally divides the sample values in two halves) as 414 HB, which is much closer to the expected value of 415 HB. In fact, for skewed distributions the median is generally considered to be the best representative of the central location of the data.

Table 16 shows the results obtained for this new situation by using the GUM approach.

Parameter	Value
Combined standard uncertainty	100 HB
Effective degrees of freedom	4
Coverage factor (*k*)	2.78
Expanded uncertainty	278 HB

Table 16. Results obtained for the Brinell hardness model by using the GUM uncertainty approach (with 0.35 mm instead of 0.035 mm as standard deviation of *d*), for a coverage probability of 95%

Calculating the differences of the endpoints between the two methods yields d_{low}= 133 HB and d_{high} =16 HB, which are much higher than δ = 0.5 HB, invalidating the GUM approach, as expected.

10. Conclusions

The GUM uncertainty framework is currently still the most extensive method used on estimation of measurement uncertainty in metrology. Despite its approximations, it suits very well on a wide range of measurement systems and models.

However, the use of numerical methods like Monte Carlo simulations has been increasingly encouraged by the Joint Committee for Guide in Metrology (JCGM) of the Bureau International des Poids et Mesures (BIPM) as a valuable alternative to the GUM approach. The simulations rely on the propagation of distributions, instead of propagation of uncertainties like the GUM, and thus are not subjected to its approximations. Consequently, Monte Carlo simulations provide results for a wider range of models, including situations in which the GUM approximations may not be adequate (see section 4 for GUM limitations), like for example when the models contain non-linear terms or when there is a large non-Gaussian input source that predominates over the others.

The practical use of Monte Carlo simulations on the estimation of uncertainties is still gaining ground on the metrology area, being limited to National Institutes of Metrology and some research groups, as it still needs more dissemination to third party laboratories and institutes. Nevertheless, it has proven to be a fundamental tool in this area, being able to address more complex measurement problems that were limited by the GUM approximations.

Author details

Paulo Roberto Guimarães Couto[1], Jailton Carreteiro Damasceno[2] and
Sérgio Pinheiro de Oliveira[1]

*Address all correspondence to: prcouto@inmetro.gov.br

1 Mechanical Metrology Division, Inmetro, Duque de Caxias, Brazil

2 Materials Metrology Division, Inmetro, Duque de Caxias, Brazil

References

[1] ISO IEC 17025:2005 – General requirements for the competence of testing and calibration laboratories. International Organization for Standardization, Geneva; 2005.

[2] JCGM 200:2012 – International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM). Joint Committee for Guides in Metrology; 2012.

[3] JCGM 100:2008 – Evaluation of measurement data – Guide to the expression of uncertainty in measurement. Joint Committee for Guides in Metrology; 2008.

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

[5] Lepek A. A computer program for a general case evaluation of the expanded uncertainty. Accreditation and Quality Assurance 2003; 8(6) 296-299.

[6] Wichmann BA, Hill ID. Generating good pseudo-random numbers. Computational Statistics & Data Analysis 2006; 51(3) 1614–1622.

[7] U.S. Department of Energy. Energy Efficiency & Renewable Energy. http://www.eere.energy.gov/hydrogenandfuelcells/fuelcells/index.html (accessed 1 August 2012).

[8] Pilatowsky I, Romero RJ, Isaza CA, Gamboa SA, Sebastian PJ, Rivera W. Cogeneration Fuel Cell-Sorption Air Conditioning Systems. Springer; 2011.

[9] Oliveira SP, Rocha AC, Filho JT, Couto PRG. Uncertainty of Measurement by Monte-Carlo Simulation and Metrological Reliability in the Evaluation of Electric Variables of PEMFC and SOFC Fuel Cells. Measurement 2009, 42(10) 1497-1501.

[10] EURACHEM/CITAC Guide CG4. Quantifying Uncertainty in Analytical Measurement. EURACHEM/CITAC; 2012.