

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

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



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



Methods for Evaluation of Measurement Uncertainty

Jailton Carreteiro Damasceno and Paulo R.G. Couto

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.74873>

Abstract

This chapter presents and explains the most used methodologies for the evaluation of measurement uncertainty in metrology with practical examples. The main topics are basic concepts and importance, existing documentation, the harmonized methodology of the Guide to the Expression of Uncertainty in Measurement, types of uncertainty, modeling of measurement systems, use of alternative methods (including the GUM supplement 1 Monte Carlo numerical method), evaluation of uncertainty for calibration curves, correlated uncertainties, uncertainties arising from the calibration of instruments, and the main proposals for the new revised GUM. The chapter also discusses the GUM as a tool for the technical management of measurement processes.

Keywords: metrology, measurement, uncertainty, GUM, Monte Carlo

1. Introduction

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. Uncertainty evaluation is essential to guarantee the metrological traceability of measurement results and to ensure that they are accurate and reliable. In addition, measurement uncertainty must be considered whenever a decision has to be taken based on measurement results, such as in accept/reject or pass/fail processes.

Considering the context of globalization of markets, it is necessary to adopt a universal procedure for evaluating uncertainty of measurements, in view of the need for comparability of results between nations and for mutual recognition in metrology. As an example, laboratories accredited under the ISO/IEC 17025:2017 standard [1] need to demonstrate their technical

competence and the ability to properly operate their management systems, and so they are required to evaluate the uncertainty for their measurement results.

In addition, the use of uncertainty evaluation methods as a tool for technical management of measurement processes is extremely important to reduce, for example, the large number of losses that occurs in the industry, which can be highly significant in relation to the gross domestic product (GDP) of some countries. One of the probable causes of the waste can be attributed to instruments whose accuracy is inadequate to the tolerance of a certain measurement process.

In this chapter, detailed steps for uncertainty evaluation are given.

2. Main references for uncertainty evaluation

In order to harmonize the uncertainty evaluation process for every laboratory, the *Bureau International des Poids et Mesures* (BIPM) published in 1980 the Recommendation INC-1 [2] on how to express uncertainty in measurement. This document was further developed and originated the “Guide to the Expression of Uncertainty in Measurement”—GUM in 1993, which was revised in 1995 and lastly in 2008. This document provides complete guidance and references on how to treat common situations on metrology and how to deal with uncertainties in metrology. Currently, it is published by International Organization for Standardization (ISO) as the ISO/IEC Guide 98-3 “Uncertainty of measurement—Part 3: Guide to the expression of uncertainty in measurement” (GUM), and by the Joint Committee for Guides in Metrology (JCGM) as the JCGM 100:2008 guide [3]. The JCGM was established by BIPM to maintain and further develop the GUM. They are in fact currently producing a series of documents and supplements to accompany the GUM, four of which are already published [4–7].

Evaluation of uncertainty, as presented by the JCGM 100:2008, is based on the law of propagation of uncertainties (LPU). This methodology has been successfully applied for several years worldwide for a range of different measurement systems and is currently the most used procedure for uncertainty evaluation in metrology. However, since its twentieth anniversary in 2013, JCGM decided to revise it again [8–10]. In this new revision, uncertainty terms and concepts [11] will be aligned with the current International Vocabulary of Metrology (VIM) [12] and with the new GUM supplements [5, 6]. Aspects such as a new Bayesian approach, the redefinition of coverage intervals and the elimination of the Welch-Satterthwaite formula to evaluate the effective degrees of freedom will be covered [9]. In late 2014, a first draft of the newly revised version of the GUM was circulated among National Metrology Institutes. Remarkable changes were made that could affect the way laboratories deal with the measurement uncertainty results. This revision is still being discussed, and some information about it has also been released elsewhere [10].

In the field of analytical chemistry, there is also another document worth mentioning that is the “Quantifying Uncertainty in Analytical Measurement” guide [13], produced by a joint

EURACHEM/CITAC Measurement Uncertainty Working Group. This document was first published in 1995 and further revised in 2000 [14]. This last edition had a widespread implementation and is among the most highly cited publications in chemical metrology area [14]. Recently, a new revised edition was published in 2012 with improved content and added information on developments in uncertainty evaluation [14]. This document basically presents the uncertainty evaluation process following the suggestions of the GUM, but also contains several examples in the analytical chemistry area.

3. Using the GUM approach on uncertainty evaluation

The following main steps summarize the methodology presented by the GUM.

3.1. Definition of the measurand and of input quantities

It must be clear to the analyst which quantity will be the final object of the measurement in question. This quantity is known as the measurand. In addition, it is important to identify all the variables that directly or indirectly influence the measurand. These variables are known as the input quantities. As an example, Eq. (1) shows a measurand y as a function of three different input quantities: x_1 , x_2 , and x_3 .

$$y = f(x_1, x_2, x_3) \quad (1)$$

3.2. Modeling the measurement process

In this step, the measurement procedure should be modeled in order to have a functional relationship expressing the measurand as a result of all the input quantities. The measurand y in Eq. (1) could be modeled, for example, as in Eq. (2)

$$y = \frac{x_1 x_2}{x_3^2} \quad (2)$$

The modeling step is critical for the uncertainty evaluation process as it defines how the input quantities impact the measurand. The better the model is defined, the better its representation of reality will be, including all the sources that impact the measurand on the uncertainty evaluation. The modeling process can be easily visualized by using a cause-effect diagram (Figure 1).

Example: To illustrate these steps, let us consider a measurement model for a torque test. Torque is 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 a known mass to the end of a horizontal arm while keeping the other end fixed (Figure 2).

Note: This example is also presented, with a few adaptations, in other publications by the same authors [15].

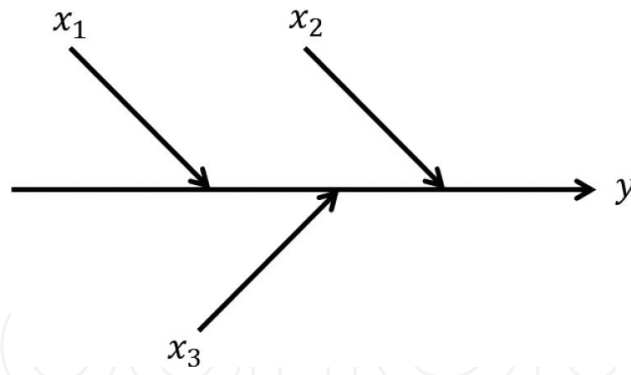


Figure 1. A cause-effect diagram representing the model from Eq. (2).

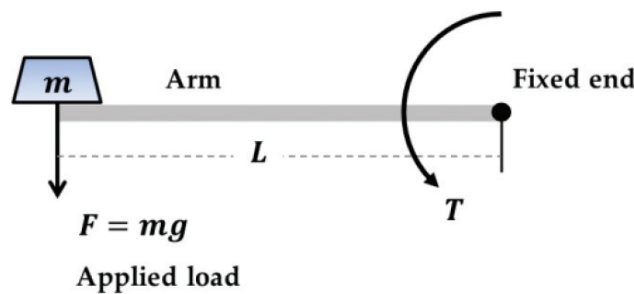


Figure 2. A conceptual illustration of the experimental setup for a measurement of torque (T), where F is the applied force, m is the mass of the load, g is the local gravity acceleration, and L is the length of the arm.

A simple model that describes this experiment can be expressed as follows:

$$T = mgL \quad (3)$$

where T is the torque (N.m), m is the mass of the applied load (kg), g is the local gravity acceleration (m/s^2), and L is the total length of the arm (m). Thus, m , g , and L are the input quantities for this model. This example will be further discussed in the subsections ahead.

3.3. Evaluating the uncertainties of the input quantities

This step is also of great importance. Here, uncertainties for all the input quantities are individually evaluated. The GUM classifies uncertainty sources as being of two main types: Type A, which usually originates from some statistical analysis, such as the standard deviation obtained in a repeatability study; and Type B, which is determined from any other source of information, such as a calibration certificate or deduced from personal experience.

Type A uncertainties from repeatability studies are evaluated by the GUM as the standard deviation of the mean obtained from the repeated measurements. For example, if a set of n indications x_i about a quantity x are available, the uncertainty u_x due to repeatability of the measurements can be expressed by $s(\bar{x})$ as follows in Eq. (4):

$$u_x = s(\bar{x}) = \frac{s(x_i)}{\sqrt{n}} \quad (4)$$

where \bar{x} is the mean value of the repeated measurements, $s(x_i)$ is its standard deviation, and $s(\bar{x})$ is the standard deviation of the mean. As such, the statistical distribution associated with this input source is considered to be normal or Gaussian.

Note: This evaluation is not consistent with the GUM supplement 1 [5], where repeated indications are treated as Student's t -distributions to account for the lack of degrees of freedom or a low number of indications. In this way, the new proposal for the draft GUM is to consider repeated indications as t -distributions, just like in supplement 1. Therefore, its uncertainty would be evaluated as in Eq. (5). This equation takes the degrees of freedom for the indications ($n - 1$) into account, raising the uncertainty for a low number of indications. This correction would then be in accordance with the approach suggested by the other GUM supplements for this type of uncertainty

$$u_x = \left(\frac{n-1}{n-3} \right)^{1/2} \frac{s(x_i)}{\sqrt{n}} \quad (5)$$

It is important to note that the evaluation of uncertainties of Type B input sources must be based on careful analysis of observations or in an accurate scientific judgment, using all available information about the measurement procedure. This uncertainty type is generally used when repeated experiments would not be possible, not available, or would be too costly or time-consuming. In this case, the GUM also suggests the use of two more types of statistical distributions: the uniform and the triangular distributions.

The uniform distribution should be used when only a range of values are available, that is, an interval with the minimum and maximum values, and no detailed information about the probability of values within this interval is available. The standard uncertainty associated with such an interval is given by Eq. (6):

$$u_x = \frac{b-a}{\sqrt{12}} \quad (6)$$

where b is the maximum and a is the minimum values for the range. For example, if the only information about the room temperature of a laboratory is known to be $(20 \pm 2)^\circ\text{C}$, then $b - a = 22 - 18 = 4^\circ\text{C}$ and the standard uncertainty associated with the room temperature would be evaluated as $u_\theta = 4/\sqrt{12}^\circ\text{C} = 1.15^\circ\text{C}$.

The triangular distribution can be used when there is a strong evidence that the most probable value lies in the middle of a given interval, but still without knowing exactly how this probability behave within the interval. In chemistry, for example, the uncertainty associated with the volume of a measuring flask could be evaluated by a triangular distribution. The standard uncertainty for a triangular distribution is given by Eq. (7):

$$u_x = \frac{a}{\sqrt{6}} \quad (7)$$

where a is the semi-interval for the total range of the triangular distribution.

Another common Type B source of uncertainty is due to calibration certificates, related to a standard or to a calibrated instrument. In this case, the standard uncertainty to be used is normally obtained by dividing the expanded uncertainty U by the coverage factor k , both provided by the calibration certificate (Eq. (8))

$$u_x = \frac{U}{k} \quad (8)$$

Several good examples on how to treat some of the most common uncertainty sources can be found on the GUM [3], the EURACHEM/CITAC guide [13], and elsewhere [16].

Example: Returning to the example of torque measurement and considering the model defined in Eq. (3), the following sources of uncertainty are considered:

Mass (m). The mass m was repeatedly measured 10 times in a calibrated balance. 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 standard uncertainty (u_{m_R}) that applies in this case is obtained by Eq. (4), that is, $u_{m_R} = 0.3 \text{ g}/\sqrt{10} = 9.49 \times 10^{-5} \text{ kg}$.

In addition, the balance used for the measurement has a certificate stating an expanded uncertainty for this range of mass of $U_m = 0.1 \text{ g}$, with a coverage factor $k = 2$ and a coverage probability 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, the standard uncertainty (u_{m_C}) is calculated by using Eq. (8), that is, $u_{m_C} = U_m/k = 0.1 \text{ g}/2 = 0.00005 \text{ kg}$.

Local gravity acceleration (g). The value for the local gravity acceleration is stated in a certificate of measurement as 9.80665 m/s^2 , as well as its expanded uncertainty of $U_g = 0.00002 \text{ m/s}^2$, for $k = 2$ and $p = 95\%$. Again, Eq. (8) is used to calculate the standard uncertainty (u_g), that is, $u_g = U_g/k = (0.00002 \text{ m/s}^2)/2 = 0.00001 \text{ m/s}^2$.

Length of the arm (L). Let us suppose that in this hypothetical case, 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. However, in this 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 $\pm 0.5 \text{ mm}$ as limits for the measurement. As no information of probabilities within this interval is available, the assumption of a uniform distribution is the best option, on which there is equal probability for the values within the whole interval. Thus, Eq. (6) is used to determine the standard uncertainty (u_L), that is, $u_L = (2000.5 - 1999.5) \text{ mm}/\sqrt{12} = 0.000289 \text{ m}$.

In practice, one can imagine several more sources of uncertainty for this 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 implement the methodology of the GUM on a simple model.

3.4. Propagation of uncertainties

3.4.1. The law of propagation of uncertainties

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

According to the LPU, the propagation of uncertainties is accomplished 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, the treatment of a model where the measurand y is expressed as a function of N variables x_1, \dots, x_N (Eq. (9)) leads to the general expression for the propagation of uncertainties shown in Eq. (10)

$$y = f(x_1, \dots, x_N) \quad (9)$$

$$u_y^2 = \sum_{i=1}^N \left(\frac{\partial y}{\partial x_i} \right)^2 u_{x_i}^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left(\frac{\partial y}{\partial x_i} \right) \left(\frac{\partial y}{\partial x_j} \right) \text{COV}(x_i, x_j) \quad (10)$$

where u_y is the combined standard uncertainty for the measurand y and u_{x_i} is the uncertainty for the i th input quantity. The second term of Eq. (10) is due to the correlation between the input quantities. If there is no supposed correlation between them, Eq. (10) can be further simplified as

$$u_y^2 = \sum_{i=1}^N \left(\frac{\partial y}{\partial x_i} \right)^2 u_{x_i}^2 \quad (11)$$

The partial derivatives of Eq. (11) are known as sensitivity coefficients and describe how the output estimate y varies with changes in the values of the input estimates x_1, x_2, \dots, x_N . It also converts the units of the inputs to the unit of the measurand.

Another important observation regarding the sensitivity coefficient occurs when the mathematical model that defines the measurand does not contemplate a given quantity, known as influence quantity. In this case, the determination of the sensitivity coefficient of the measurand in relation to the input quantity must be done experimentally. For example, biodiesel is susceptible to oxidation when exposed to air, and this oxidation process affects fuel quality. The oxidation time is determined by measuring the conductivity of an oil sample when inflated with air at a given flow rate. There are a number of influence quantities that impact the oxidation time of biodiesel such as temperature, air flow, conductivity, sample mass, and so on. In this case, the sensitivity coefficients for oxidation time with respect to each of these influence quantities are determined from an interpolation function obtained with experimental data. For example, **Figure 3** presents the table and its resulting graph, which shows the model of the function that relates the oxidation time to the temperature of a biofuel sample (case study of the authors).

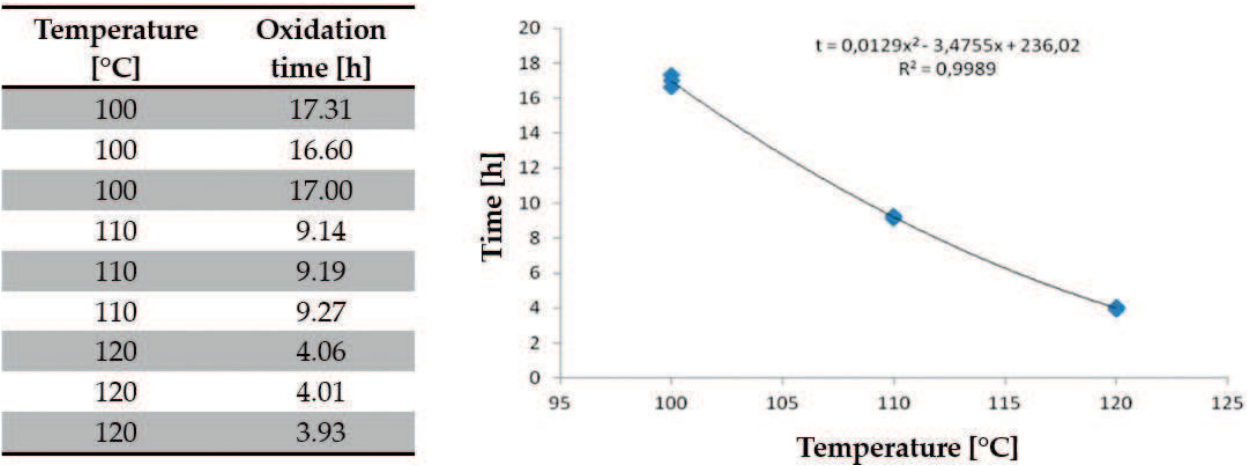


Figure 3. A table and a graph representing the variation of the oxidation time of a biofuel sample as a function of temperature.

Example: On returning to the torque measurement example, assuming that all the input quantities are independent, the combined standard uncertainty for the torque is calculated by using the LPU (Eq. (11)). The final expression is then

$$u_T = \sqrt{\left(\frac{\partial T}{\partial m}\right)^2 u_{m_R}^2 + \left(\frac{\partial T}{\partial m}\right)^2 u_{m_C}^2 + \left(\frac{\partial T}{\partial g}\right)^2 u_g^2 + \left(\frac{\partial T}{\partial L}\right)^2 u_L^2} = 0.096 \text{ N m} \tag{12}$$

It is important to note that the terms (not squared) of Eq. (12), that is, each sensitivity coefficient multiplied by its corresponding uncertainty, are known as uncertainty components. These components can be compared to each other as they are in the same units of the measurand. **Figure 4** shows the comparison between the uncertainty components for the torque measurement model.

As can be noted, the dominant uncertainty component is due to the uncertainty associated with the measurement of the arm length, which was taken as the resolution of the non-calibrated

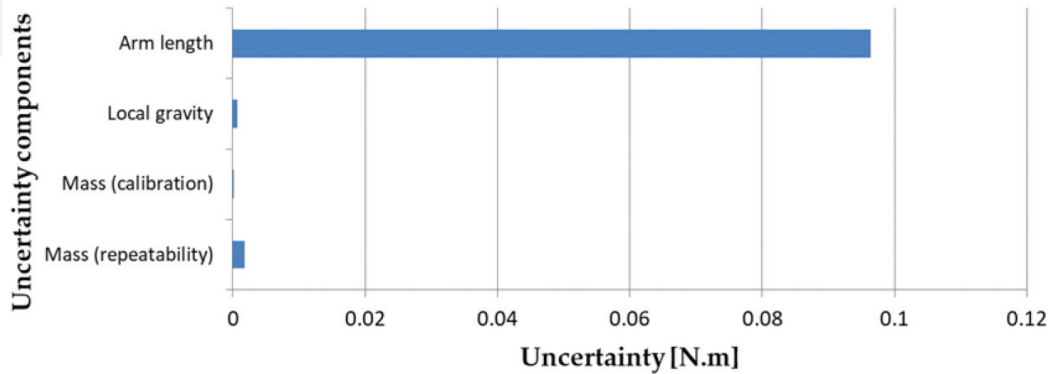


Figure 4. Uncertainty component balance for the input quantities in the torque measurement model.

ruler used in the measurement. This analysis shows to the analyst that, to reduce the final uncertainty and improve the measurement system, a calibrated ruler, with a better uncertainty, should be used. This represents the importance of the GUM as a management tool to the measurement process.

3.4.2. The Kragten method

The Kragten method is an approximation that facilitates the calculation of the combined uncertainty using finite differences in place of the derivatives [13]. This approximation is valid when the uncertainties of the inputs are relatively small compared to the respective values of the input quantities, generating discrepancies in the final result in relation to the LPU that occur in decimals that can be ignored.

Assuming a measurand y , which is calculated from the input quantities x_1 , x_2 and x_3 according to the mathematical model of Eq. (2), the uncertainties u_{x_1} , u_{x_2} and u_{x_3} for the input quantities are evaluated normally, according to methodologies already explained previously. From there, the calculations of the measurand are performed individually for each input magnitude (y_{x_1} , y_{x_2} and y_{x_3}) so that each time their respective values are added with their uncertainties, as shown in Eqs. (13)–(15)

$$y_{x_1} = \frac{(x_1 + u_{x_1})x_2}{x_3^2} \quad (13)$$

$$y_{x_2} = \frac{x_1(x_2 + u_{x_2})}{x_3^2} \quad (14)$$

$$y_{x_3} = \frac{x_1x_2}{(x_3 + u_{x_3})^2} \quad (15)$$

The value of the measurand y varies for y_{x_i} due to the addition of the uncertainty u_{x_i} to the value of its respective input quantity. Thus, the uncertainty component of each input source in the unit of the measurand y is defined by the difference $|y_{x_i} - y|$, according to Eqs. (16)–(18)

$$u_y(x_1) = |y_{x_1} - y| \quad (16)$$

$$u_y(x_2) = |y_{x_2} - y| \quad (17)$$

$$u_y(x_3) = |y_{x_3} - y| \quad (18)$$

Thus, the combined standard uncertainty of y is finally evaluated as

$$u_y = \sqrt{\sum_{i=1}^N u_y^2(x_i)} \quad (19)$$

or by Eq. (20), if there are correlated uncertainties

$$u_y = \sqrt{\sum_{i=1}^N u_y^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N u_y(x_i) u_y(x_j) r(x_i, x_j)} \quad (20)$$

where $r(x_i, x_j)$ is the correlation coefficient between x_i and x_j .

3.5. Evaluation of the expanded uncertainty

The result provided by Eqs. (10) and (11) corresponds to an interval that contains only one standard deviation (or approx. 68.2% of the measurements for a normal distribution). In order to have a better coverage probability for the result, the GUM approach expands this interval by assuming that the measurand follows the behavior of a Student's t -distribution. An effective degrees of freedom v_{eff} for the t -distribution can be obtained by using the Welch-Satterthwaite formula (Eq. (21))

$$v_{eff} = \frac{u_y^4}{\sum_{i=1}^N \frac{\left(\frac{\partial y}{\partial x_i}\right)^4 u_{x_i}^4}{v_{x_i}}} \quad (21)$$

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

The effective degrees of freedom is used to obtain a coverage factor k that depends also of a chosen coverage probability p , which is often 95%. The expanded uncertainty U_y is then evaluated by multiplying the combined standard uncertainty by the coverage factor k that finally expands it to a coverage interval delimited by a t -distribution with a coverage probability p (Eq. (22))

$$U_y = k u_y \quad (22)$$

Note: The draft for the new GUM proposal suggests that the final coverage interval cannot be reliably determined if only an expectation y and a standard deviation u_y are known, mainly if the final distribution deviates significantly from a normal or a t -distribution. Thus, they propose distribution-free coverage intervals in the form of $y \pm U_p$, with $U_p = k_p u_y$: (a) if no information is known about the final distribution, then a coverage interval for the measurand Y for coverage probability of at least p is determined using $k_p = 1/(1-p)^{1/2}$. If $p = 0.95$, a coverage interval of $y \pm 4.47 u_y$ is evaluated. (b) If it is known that the distribution is unimodal and symmetric about y , then $k_p = 2/[3(1-p)^{1/2}]$ and the coverage interval $y \pm 2.98 u_y$ would correspond to a coverage probability of at least $p = 0.95$.

Example: The effective degrees of freedom for the torque measurement example is calculated using Eq. (21). As the number of degrees of freedom for Type B uncertainties is considered infinite, only Type A uncertainties are accounted. In this case,

$$v_{eff} = \frac{u_T^4}{\frac{\left(\frac{\partial T}{\partial m_R}\right)^4 u_{m_R}^4}{v_{m_R}}} = 6.5 \times 10^7 \quad (23)$$

Using t -distribution tables, the coverage factor for this value of v_{eff} and $p = 95\%$ is $k = 1.96$. Therefore, the expanded uncertainty is calculated as $U = ku_T = 1.96 \times 0.096 = 0.2 \text{ N m}$, and the measurement result is expressed as $668.0 \pm 0.2 \text{ N m}$. The GUM recommends that the final uncertainty should be expressed with one or at most two significant digits.

4. Calibration curve and correlated uncertainties

One of the most valuable tools for the metrologist is the calibration curve. It is widely used in measurement systems on which one cannot directly obtain the property value to be measured of an object. Instead, a response from the system is measured. In this way, a calibration curve is used to correlate the response from the system with well-known property values, usually calibration standards (see **Figure 5**).

With a calibration curve in hands, the property value for a new unknown sample can be directly determined by using the equation for the fitted curve, which is usually adjusted by a linear regression. However, the calibration curve contains errors due to the lack of fit for the experimental data, causing an uncertainty source to arise. Thus, when evaluating the uncertainty for a predicted property value of x_o corresponding to a new observation y_o (for a new unknown sample, for example), the LPU with correlation terms is applied on the linear regression model in the form of Eq. (24). Eq. (25) represents the model for a predicted value y_o corresponding to a new observed value x_o , in the case of the inverse process

$$x_0 = \frac{y_o - a}{b} \quad (24)$$

$$y_o = a + bx_0 \quad (25)$$

where a and b are, respectively, the intercept and the slope parameters of the linear regression.

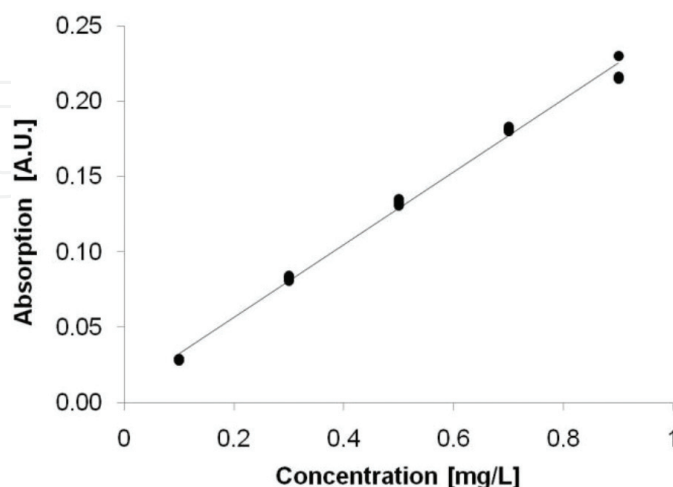


Figure 5. An example of a linear calibration curve for atomic absorption spectroscopy: the absorption signals (instrument responses) are plotted against the concentrations for a specific analyte.

The application of the LPU with the correlation term to Eqs. (24) and (25) leads to Eqs. (26) and (27), respectively, for both cases:

$$u_{x_o} = \sqrt{\left(\frac{\partial x_o}{\partial y_o}\right)^2 u_{y_o}^2 + \left(\frac{\partial x_o}{\partial a}\right)^2 u_a^2 + \left(\frac{\partial x_o}{\partial b}\right)^2 u_b^2 + 2\left(\frac{\partial x_o}{\partial a}\right)\left(\frac{\partial x_o}{\partial b}\right) u_a u_b r_{a,b}} \quad (26)$$

$$u_{y_o} = \sqrt{\left(\frac{\partial y_o}{\partial x_o}\right)^2 u_{x_o}^2 + \left(\frac{\partial y_o}{\partial a}\right)^2 u_a^2 + \left(\frac{\partial y_o}{\partial b}\right)^2 u_b^2 + 2\left(\frac{\partial y_o}{\partial a}\right)\left(\frac{\partial y_o}{\partial b}\right) u_a u_b r_{a,b}} \quad (27)$$

For Eq. (26), u_{x_o} is the combined uncertainty for the predicted value x_o and u_{y_o} is the uncertainty for the new observed response y_o . For Eq. (27), u_{y_o} is the combined uncertainty for the predicted value y_o and u_{x_o} is the uncertainty for the new observed response x_o . In both cases, u_a and u_b are, respectively, the uncertainties for the intercept and the slope, and $r_{a,b}$ is the correlation coefficient between a and b . These last equations can also be found in the ISO/TS 28037 [17], concerning the use of straight-line calibration functions.

The uncertainties for a and b can be obtained by Eqs. (28) and (29), respectively, while the correlation coefficient $r_{a,b}$ is given by Eq. (30)

$$u_a = S_e \sqrt{\frac{\sum x_i^2}{n \sum x_i^2 - (\sum x_i)^2}} \quad (28)$$

$$u_b = S_e \sqrt{\frac{n}{n \sum x_i^2 - (\sum x_i)^2}} \quad (29)$$

$$r_{a,b} = -\frac{\sum x_i}{\sqrt{n \sum x_i^2}} \quad (30)$$

where n is the number of points used to construct the curve, x_i are the values for the independent variable of the linear equation for each y_i , and S_e^2 is the residual variance of the fitted curve, obtained by Eq. (31)

$$S_e^2 = \frac{\sum (y_i - \hat{y}_i)^2}{n - 2} \quad (31)$$

where \hat{y}_i are the interpolated values in the fitted curve for each x_i , that is, $\hat{y}_i = a + bx_i$.

Example: This time, let us consider that the calibration certificate of a thermometer presents the results shown in **Table 1**.

For the data shown in **Table 1**, the calibration curve of the thermometer is expressed by $\hat{y}_o = 1.1484 + 0.9578x_o$. For a temperature value indicated by the thermometer of $x_o = 22^\circ\text{C}$, applying the equation of the calibration curve yields a reference value of $\hat{y}_o = 22.22^\circ\text{C}$.

Indication (x_i) (°C)	Reference value (y_i) (°C)
20	20.3
21	21.3
22	22.2
23	23.1
24	24.2
25	25.1
27	27.0

Table 1. Values of the calibration certificate of a thermometer.

Using Eqs. (28)–(31), it is possible to calculate the values of **Table 2** that shows the statistical data for the thermometer calibration curve.

Considering that there is no uncertainty for the observed point $x_o = 22^\circ\text{C}$, that is, $u_{x_o} = 0$, the uncertainty of y_o arising from the interpolation process of the point $x_o = 22^\circ\text{C}$ can then be calculated by applying Eq. (27) and the data from **Table 2**, resulting in the following:

$$u_{y_o} = \sqrt{1^2 \cdot 0.1943^2 + 22^2 \cdot 0.0084^2 + 2 \cdot 1 \cdot 22 \cdot 0.1943 \cdot 0.0084 \cdot (-0.995)} = 0.021^\circ\text{C}.$$

Another frequently used expression for the standard uncertainty of the predicted value u_{x_o} is given by Eq. (32) [13, 18]:

$$u_{x_o} = \frac{S_e}{b} \sqrt{\frac{1}{m} + \frac{1}{n} + \frac{(\bar{y}_o - \bar{y})^2}{b^2 \sum (x_i - \bar{x})^2}} \quad (32)$$

where S_e is the residual standard deviation of the fitted line, m is the number of observations of y_o , n is the number of points composing the calibration curve, and \bar{y}_o is the average value obtained from the observations of y_o . In this expression, the uncertainty component due to the observations of y_o is evaluated by [19]

$$u_{y_o} = \frac{S_e}{\sqrt{m}} \quad (33)$$

However, Hibbert [19] suggests that if the standard deviation of the indications is known from consistent data, u_{y_o} can be better evaluated by

Data	Value	Unit
S_e^2	0.0024	°C ²
u_a	0.1943	°C
u_b	0.0084	
$r_{a,b}$	−0.995	

Table 2. Statistical data for the calibration curve of a thermometer.

$$u_{y_o} = \frac{S_{y_o}}{\sqrt{m}} \quad (34)$$

where S_{y_o} is the standard deviation of the observations of y_o , and Eq. (32) is then expressed as Eq. (35) [18, 19]:

$$u_{x_o} = \frac{1}{b} \sqrt{\frac{S_{y_o}^2}{m} + \frac{S_e^2}{n} + \frac{S_e^2 (\bar{y}_o - \bar{y})^2}{b^2 \sum (x_i - \bar{x})^2}} \quad (35)$$

5. Assessment of uncertainty in instrument calibration

The methodology presented in the GUM can also be used to evaluate the uncertainty in the calibration of a measuring instrument. Following the steps of the GUM, the measurand for the model used in the calibration must be defined by the quantity that evaluates the systematic error of an instrument over its entire measurement range. Thus, Eq. (36) can be generally used for the evaluation of uncertainty in a calibration process:

$$e = V_{ind} - V_{ref} \quad (36)$$

where e is the systematic error of the instrument for a fixed range, V_{ind} is the value indicated by the instrument, and V_{ref} is the reference value corresponding to the indicated value.

From Eq. (36), a basic cause-and-effect diagram can be assembled for the calibration uncertainty assessment of an instrument, as shown in **Figure 6**.

The sources of uncertainty in this case involve the repeatability of indicated values, the resolution of the instrument in calibration, and the certificate of calibration of the reference values. Thus, an evaluation of the uncertainty about the systematic error should be done for each nominal value of the instrument in calibration. The combined standard uncertainties u_{e_i} for each calibrated nominal value are obtained by applying the LPU, as shown in Eq. (37)

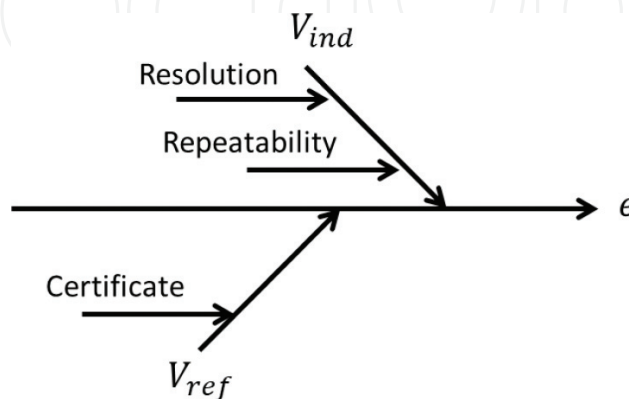


Figure 6. A general cause-and-effect diagram for the calibration of an instrument.

$$u_{e_i} = \sqrt{\left(\frac{\partial e_i}{\partial V_{ind}}\right)^2 u_{V_{indRes}}^2 + \left(\frac{\partial e_i}{\partial V_{ind}}\right)^2 u_{V_{indRep}}^2 + \left(\frac{\partial e_i}{\partial V_{ref}}\right)^2 u_{V_{ref}}^2} \quad (37)$$

where $u_{V_{indRes}}$, $u_{V_{indRep}}$, and $u_{V_{ref}}$ are, respectively, standard uncertainties due to resolution of the instrument, repeatability of indication values, and certificate of calibration of the reference. These standard uncertainties are obtained as described in Section 3.

The final calibration result can then be presented according to **Table 3**. In addition, correction values or systematic errors can also be reported.

6. Monte Carlo simulation applied to metrology

This section presents the limitations of the GUM and shows an alternative methodology based on the propagation of distributions that overcome those limitations. For further details, please refer to the authors' publication that addresses the use of the Monte Carlo methodology applied to uncertainty in measurement [15] or to the JCGM 101:2008 guide [5]. Also, in the field of analytical chemistry, the latest version of EURACHEM/CITAC guide (2012) was updated with procedures to use Monte Carlo simulations [13].

6.1. Limitations of the GUM approach

As mentioned earlier, the approach to evaluate measurement uncertainties using the LPU as presented by the GUM is based on some approximations that are not valid for every measurement model [5, 20–22]. These approximations comprise (1) the linearization of the measurement model made by the truncation of the Taylor series, (2) the use of a t -distribution as the distribution for the measurand, and (3) the calculation of an effective degrees of freedom for the measurement model based on the Welch-Satterthwaite formula, which is still an unsolved problem [23]. Moreover, the GUM approach usually presents deviated results when one or more of the input uncertainties are relatively much larger than others, or when they have the same order of magnitude than its quantity.

The limitations and approximations of the LPU are overcome when using a methodology that relies on the propagation of distributions. This methodology carries more information than the simple propagation of uncertainties and generally provides results closer to reality. It is

Range	Indicated value	Reference value	Expanded uncertainty	Coverage factor
Range 1	V_{ind1}	V_{ref1}	U_1	k_1
Range 2	V_{ind2}	V_{ref2}	U_2	k_2
...
Range N	V_{indN}	V_{refN}	U_N	k_N

Table 3. A typical format for the result of calibration of an instrument.

described in detail by 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) [5], providing basic guidelines for using Monte Carlo numerical simulations for the propagation of distributions in metrology. This method provides reliable results for a wider range of measurement models as compared to the GUM approach and is presented as a fast and robust alternative method for cases where the GUM approach does not present good results.

6.2. Running Monte Carlo simulations

The propagation of distributions as presented by the JCGM 101:2008 involves the convolution of the probability distributions for the input sources of uncertainty through the measurement model to generate a distribution for the output (the measurand). In this process, no information is lost due to approximations, and the result is much more consistent with reality.

The main steps of this methodology are similar to those presented in the GUM. The measurand must be defined as a function of the input quantities through a model. Then, for each input, a probability density function (PDF) must be assigned. In this step, the concept of maximum entropy used in the Bayesian statistics should be used to assign a PDF that does not contain more information than that which is known by the analyst. A number of Monte Carlo trials are then chosen and the simulation can be set to run.

Results are expressed in terms of the average value for the output PDF, its standard deviation, and the end points that cover a chosen probability p .

Example: Returning once more to the torque measurement example, one can consider the following PDFs for the input sources:

Mass (m). For repeated indications, the JCGM 101:2008 suggests the use of a scaled and shifted t -distribution. Thus, the distribution should use 35.7653 kg as its average, a scale value of $s/\sqrt{n} = 0.3 \text{ g}/\sqrt{10} = 9.49 \times 10^{-5} \text{ kg}$, and $n - 1 = 9$ degrees of freedom.

For the calibration component, the supplement 1 recommends the use of a normal distribution if the number of degrees of freedom is not available. In this case, the mass value of 35.7653 kg is taken as the mean and a standard deviation of $U_m/k = 0.1 \text{ g}/2 = 0.00005 \text{ kg}$ should be used. However, to facilitate the calculation of the final mean value of the measurand, the mean should be shifted to zero, since both values for the mass sources will be added together.

Local gravity acceleration (g). This case is similar to the case of the balance certificate, for which we have values of expanded uncertainty and coverage factor without information on the number of effective degrees of freedom. Thus, a normal distribution with a mean of 9.80665 m/s^2 and a standard deviation of $U_g/k = (0.00002 \text{ m/s}^2)/2 = 0.00001 \text{ m/s}^2$ are assumed.

Length of the arm (L). In this case, as poor information about the interval is available ($\pm 0.5 \text{ mm}$), an uniform distribution is assumed with a minimum value of 1999.5 mm and a maximum value of 2000.5 mm.

Uncertainty source	Type	PDF	PDF parameters
Mass (repeatability)	A	<i>t</i> -distribution	Mean: 35.7653 kg; scale: 9.49×10^{-5} kg; degrees of freedom: 9
Mass (certificate)	B	Normal	Mean: 0 kg; standard deviation: 0.00005 kg
Local gravity	B	Normal	Mean: 9.80665 m/s ² ; standard deviation: 0.00001 m/s ²
Arm length	B	Uniform	Minimum: 1999.5 mm; maximum: 2000.5 mm

Table 4. A summary of sources of uncertainty and their associated distributions for the measurement of torque.

Table 4 resumes the input information for the simulation, which was executed for $M = 200,000$ trials, generating the output distribution shown in **Figure 7**.

Table 5 summarizes the statistical data of the output distribution, including the upper and lower limits of a probabilistically symmetric range for a 95% coverage probability.

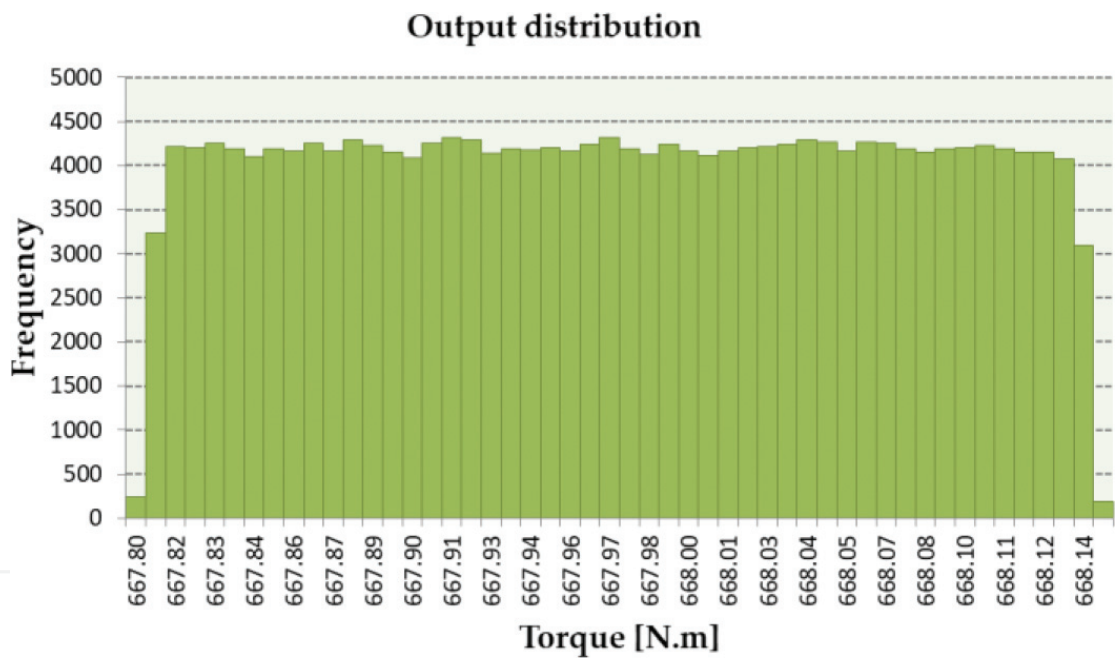


Figure 7. Output distribution resulting from the Monte Carlo simulation for the evaluation of uncertainty of measurement of torque.

Statistical data	Value (N m)
Mean	667.970
Standard deviation	0.096
Lower limit for $p = 95\%$	667.812
Upper limit for $p = 95\%$	668.129

Table 5. A summary of the statistical data for the output distribution for the measurement of torque.

7. Conclusions

Measurement uncertainty and metrological traceability are interdependent concepts. The evaluation of uncertainties of measurement results is essential to ensure that they are reliable and comparable. Moreover, the process that involves the modeling of measurement systems and evaluation of their uncertainties is of great importance for the metrologist as it constitutes a tool for the management of the measurement laboratory, since it can indicate exactly where to invest to get better, more qualified results.

The GUM and the application of the LPU continue to be the most used and widespread methodology for bottom-up uncertainty evaluation in metrology. It is adopted worldwide and provides a strong base for comparability of measurement results between laboratories. On the other hand, a new version for the GUM is currently under revision. This version should be aligned with its supplements in a more harmonized way, incorporating concepts from Bayesian statistics and resolving some inconsistencies. As a consequence, if the mentioned distribution-free coverage intervals are maintained, results for the expanded uncertainty will be greatly overestimated compared to the current version of the GUM.

In this way, the best alternative for a more realistic and lean uncertainty assessment would be through a numerical simulation using the Monte Carlo method, which should lead to a smaller and more reliable uncertainty result.

Author details

Jailton Carreteiro Damasceno* and Paulo R.G. Couto

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

National Institute of Metrology, Quality and Technology (Inmetro), Rio de Janeiro, Brazil

References

- [1] ISO/IEC 17025. 2005. General Requirements for the Competence of Testing and Calibration Laboratories. ISO: Geneva
- [2] Kaarls R. BIPM Proc.-Verb. Com. Int. Poids et Mesures 49:A1–12 (in French), Giacomo P (1981). Metrologia. 1981;17:73-74 (in English)
- [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 104:2009. Evaluation of measurement data—An introduction to the “Guide to the expression of uncertainty in measurement” and related documents. Joint Committee for Guides in Metrology. 2009

- [5] 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
- [6] JCGM 102:2011. Evaluation of measurement data—Supplement 2 to the “Guide to the expression of uncertainty in measurement” —Extension to any number of output quantities. Joint Committee for Guides in Metrology. 2011
- [7] JCGM 106:2012. Evaluation of measurement data—The role of measurement uncertainty in conformity assessment. Joint Committee for Guides in Metrology. 2012
- [8] Bich W, Cox MG, Dybkaer R, Elster C, Estler WT, Hibbert B, Imai H, Kool W, Michotte C, Nielsen L, Pendrill L, Sidney S, van der Veen AMH, Woger W. Revision of the ‘guide to the expression of uncertainty in measurement’. *Metrologia*. 2012;**49**:702-705
- [9] Bich W. Revision of the ‘guide to the expression of uncertainty in measurement’ —Why and how. *Metrologia*. 2014;**51**:S155-S158
- [10] Bich W, Cox M, Michotte C. Towards a new GUM—An update. *Metrologia*. 2016;**53**:S149-S159
- [11] Ehrlich C. Terminological aspects of the guide to the expression of uncertainty in measurement (GUM). *Metrologia*. 2014;**51**:S145-S154
- [12] JCGM 200:2012. International vocabulary of metrology—basic and general concepts and associated terms (VIM). Joint Committee for Guides in Metrology. 2012
- [13] EURACHEM/CITAC Guide CG4. Quantifying uncertainty in analytical measurement. EURACHEM/CITAC. 2012
- [14] Ellison SLR. Implementing measurement uncertainty for analytical chemistry: The Eurachem guide for measurement uncertainty. *Metrologia*. 2014;**51**:S199-S205
- [15] Couto PRG, Damasceno JC, Oliveira SP. Chapter 2—Monte Carlo simulations applied to uncertainty in measurement. In: Chan V, editor. *InTech: Theory and Applications of Monte Carlo Simulations*; 2013
- [16] Meyer VR. Measurement uncertainty. *Journal of Chromatography. A*. 2007;**1158**:15-24
- [17] ISO/TS 28037. Determination and Use of Straight-line Calibration Functions. Geneva: ISO; 2010
- [18] Massart DL, Vandeginste BGM, Buydens LMC, Jong S, Lewi PJ, Smeyers-Verbeke J. Data Handling in Science and Technology. v. 20. *Handbook of Chemometrics and Qualimetrics: Part A*. Amsterdam: Elsevier; 1997
- [19] Hibbert DB. The uncertainty of a result from a linear calibration. *Analyst*. 2006;**131**:1273-1278
- [20] Harris PM, Cox MG. On a Monte Carlo method for measurement uncertainty evaluation and its implementation. *Metrologia*. 2014;**51**:S176-S182

- [21] Possolo A. Statistical models and computation to evaluate measurement uncertainty. *Metrologia*. 2014;**51**:S228-S236
- [22] Gonzalez AG, Herrador MA, Asuero AG. Uncertainty evaluation from Monte-Carlo simulations by using Crystal-Ball software. *Accreditation and Quality Assurance*. 2005; **10**:149-154
- [23] Lepek A. A computer program for a general case evaluation of the expanded uncertainty. *Accreditation and Quality Assurance*. 2003;**8**:296-299