XVIII IMEKO WORLD CONGRESS
Metrology for a Sustainable Development
September, 17 – 22, 2006, Rio de Janeiro, Brazil

# The GUM, *Guide to the expression of uncertainty in measurement*, and related documents. Present status and future developments

*Walter Bich*[1] *and Francesca Pennecchi*[1]

[1] Istituto nazionale di ricerca metrologica, INRIM, Torino, Italy, w.bich@imgc.cnr.it

**Abstract:** In this paper the recent developments in the uncertainty evaluation are outlined, by describing the activity of the Joint Committee for Guides in Metrology-Working Group 1 on the Expression of Uncertainty in Measurement, and the documents that this WG is preparing on uncertainty.

**Keywords:** Uncertainty matrix, coverage intervals, probability density functions.

## 1. INTRODUCTION

The Guide to the Expression of Uncertainty in Measurement (GUM) [1] was published in 1993 by seven international organizations in the field of measurement, i.e., BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML. It was reprinted in a slightly revised version in 1995 and since then it has been established as the authoritative document in the field of the uncertainty of measurement. Many Countries and several organizations have adopted it as a standard, or a law, or as a basis for other specific standards.

In 1997 a Joint Committee for Guides in Metrology (JCGM) was created by the same seven organizations that had prepared the GUM (and the International Vocabulary of Basic and General Terms in Metrology, VIM). This Joint Committee had the tasks "*to promote the use of ... the GUM; to prepare supplemental guides for its broad application; and to revise and promote the use of ... the VIM. The JCGM has taken over responsibility for these two documents from ISO TAG 4, who originally published them* [2].

JCGM has two Working Groups. Working Group 1, "Expression of Uncertainty in Measurement", has the task to promote the use of the GUM and to prepare supplements for its broad application. Working Group 2, "Working Group on International Vocabulary of Basic and General Terms in Metrology (VIM)", has the task to revise and promote the use of the VIM.

In this paper, the activity of WG1 is outlined and the documents in preparation are described.

## 2. ACTIVITY OF JCGM-WG1

To fulfill its task, the JCGM-WG1 is preparing, or has planned, a number of documents. The proposed titles of these documents, under the common banner *Evaluation of measurement data*, are given below.

— An introduction to the '*Guide to the Expression of Uncertainty in Measurement*' and related documents.
— Concepts and basic principles.
— Supplement 1 to the '*Guide to the Expression of Uncertainty in Measurement*'—Propagation of distributions using a Monte Carlo method.
— Supplement 2 to the '*Guide to the Expression of Uncertainty in Measurement*'—Models with any number of output quantities.
— Supplement 3 to the '*Guide to the Expression of Uncertainty in Measurement*'—Modelling.
— The role of measurement uncertainty in deciding conformance to specified requirements.
— Applications of the least-squares method.

These documents respond to as many identified needs in the task of supporting the GUM and its broad application. Three of them are Supplements to the GUM, and are intended to be used in conjunction with it. The others are supporting documents and address general issues in the field of evaluation of measurement data. All these documents, as well as the GUM (an electronic version is in preparation by ISO), will be also available on the web. For a deeper description of these documents, as well as of their motivation, see [3].

This paper focuses on Supplements 1 and 2, describing the motivations for the two documents (Section 3) and outlining the techniques proposed to address the issues mentioned there (Sections 5 and 6, respectively).

## 3. BACKGROUND

The GUM has been in use for more than a decade. During this period, its merits and drawbacks have been clearly identified. The main merit is to have proposed a unified method for treating in a comprehensive and logically sound framework both systematic and random effects. The main drawbacks are:

a) the assumptions implicit in the method, although sufficiently weak, are not fulfilled in a number of practical cases. This is especially true for the procedure concerning the expanded uncertainty at a prescribed coverage probability.

b) the case, frequent in metrology, in which more than one measurand are estimated, is only addressed marginally and not covered to sufficient detail.

In addition, a certain inconsistency exists, inherent in the symbiosis in the same document of frequentist and bayesian views of probability in the treatment of random and systematic effects, respectively.

To obviate these drawbacks, the JCGM/WG1 is preparing two specific supplements to the GUM addressing two specific cases:

a) when a coverage interval is required at a stipulated coverage probability;
b) when more than a measurand are involved in the measurement.

The first of the two documents, Supplement 1, is now at an advanced stage and should be issued within 2006. The second, Supplement 2, is at an earlier stage and should appear in 2007.

## 4. THE GUM: A BRIEF REMAINDER

### *4.1. The combined standard uncertainty*

The GUM framework is primarily intended to obtain a combined standard uncertainty[1] $u_c(y)$ for the measurand estimate $y$, given the standard uncertainties $u(x_i)$ of the input estimates $x_i$ for the input quantities $X_i$. To this purpose, the measurement model

$$Y = f(X_1, X_2, ... X_N) \quad (1)$$

is approximated by the first-order term of a power series expansion or, if model (1) is significantly non-linear, by including higher-order terms. However, these terms are not always easy to calculate, and anyway the involved input quantities must have Gaussian distributions and be independent. Even the first partial derivatives may be difficult to calculate, for example when the model is complicated.

In the case that the appropriate conditions are met, a meaningful standard uncertainty $u_c(y)$ is obtained for the measurand estimate $y$.

### *4.2. The expanded uncertainty*

An expanded uncertainty $U$ for the measurand estimate $y$ can be obtained by multiplying the standard uncertainty $u(y)$ by an appropriate coverage factor $k$, typically such that $2 \le k \le 3$. This measure of uncertainty "*may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand*" (GUM, 6.1.2). However, there is no increase in knowledge unless the qualifier "large" is quantified, that is, unless the coverage probability $p$ corresponding to $U$ is known. An expanded uncertainty having a known coverage probability $p$ is indicated by $U_p$. Its evaluation implies some knowledge about the probability distribution of the measurand. In the GUM it is suggested that in most cases this can be approximated by a Gaussian, or, for finite degrees of freedom, by a scaled-and-shifted Student-*t* distribution. In this case, the Welch-Satterthwaite formula can be used to evaluate the effective degrees of freedom $\nu_{eff}$ necessary to select the appropriate $k$ factor. This formula is a sort of weighted mean of the degrees of freedom of the input contributions to uncertainty. Therefore, one is obliged to attach a degrees of freedom not only to Type A, but also to Type B evaluations. Now, the notion of degrees of fredom is quite hard to associate to a subjective evaluation, despite the interpretation given in the GUM, Annex G. This is only one of the drawbacks inherent in the procedure. As a further issue, the Welch-Satterthwaite formula has been questioned [4, 5]. Last but not least, the conditions for the output distribution to be a Gaussian or a scaled-and-shifted Student-*t* distribution in practice are often not fulfilled, which limits the applicability of the procedure.

[1] Incidentally, the JCGM has decided that the qualifier "combined" and therefore the subscript "c" are superfluous and can be omitted.

## 5. SUPPLEMENT 1: THE CONSTRUCTION OF AN INTERVAL OF CONFIDENCE

To obviate these difficulties, both practical and conceptual, JCGM is preparing a Supplement to the GUM in which a method based on numerical simulation is used to construct an interval of confidence (or, better, a coverage interval). This method is more general than the GUM procedure and avoids the internal inconsistencies of the GUM. It is based on the notion of propagation of probability distributions, to be compared with the GUM notion of propagation of expectations and variances of these distributions. In this approach, to each of the input estimates is associated a probability density function (PDF) representing the level of knowledge available about the estimate. Guidance is given on how to assign a PDF to an input estimate in most practical cases.

Given the PDFs of the input estimates, the PDF of the output estimate can in principle be obtained analytically by using the theory of random variables. However, the calculations are difficult and in general cannot be carried out without resorting to numerical integration. Therefore, in the Supplement a different way is followed, that is, a numerical approximation to the distribution function for the output estimate is obtained by using the Monte Carlo method. From this numerical approximation, the relevant quantities can be obtained, i.e., the expectation, the standard deviation and a coverage interval having a prescribed coverage probability.

Innovative features of this document are:

The standard uncertainty is no longer the central issue, it is rather a byproduct of the procedure. Accordingly, the classification in Type A and B evaluations does not apply any longer.

The degrees of freedom for the input estimates, as well as the effective degrees of freedom for the output estimate, are no longer necessary. As a consequence, use of the Welch-Satterthwaite formula is avoided. An important issue involved by this innovation is that the questionable concept of uncertainty of the uncertainty can be avoided.

## 6. SUPPLEMENT 2: MODELS WITH ANY NUMBER OF OUTPUT QUANTITIES

In many measurement applications several measurands $Y_i$ depend on a common set of input quantities $X_i$. The GUM (GUM, 3.1.7) is not very informative on this issue, the only hint given there being that the scalar measurand and its variance are replaced by a vector measurand and its covariance matrix (GUM, 3.1.7). However, the situation occurs frequently. Just as an example, this is the case of a set of mass standards, calibrated by subdivision from a single reference kilogram standard using a common set of balances.

In these cases, matrix notation is convenient. With this notation, the measurand is a vector $\boldsymbol{Y}_{(M\text{x}1)}$, fonction of the input quantities $\boldsymbol{X}_{(N\text{x}1)}$ according to

$$\boldsymbol{Y} = f(\boldsymbol{X}) \ . \tag{2}$$

More complicated models are also encountered, especially in electrical metrology. These may involve complex input/output quantities, or may be implicit, that is, of the form

$$f(\boldsymbol{X},\boldsymbol{Y}) = \boldsymbol{0} \ . \tag{3}$$

Different models can be classified according to their level of complexity. From this broader viewpoint, the particular case covered in the GUM is the so-called univariate explicit model, which can be represented by

$$Y_{(1\text{x}1)} = f(\boldsymbol{X}_{(N\text{x}1)}) \ . \tag{4}$$

Supplement 2 will give guidance on the solution of these general model, with examples taken from metrological practice.

## REFERENCES


[1] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML, 1995 Guide to the Expression of Uncertainty in Measurement (Geneva, Switzerland: International Organisation for Standardisation) ISBN 92-67-10188-9

[2] http://www1.bipm.org/en/committees/jc/jcgm/.

[3] Bich W, Cox MG and Harris PM, Metrologia **43** (2006) S161–S166

[4] Ballico M, Metrologia, 2000, 37(1), 61-64

[5] Hall BD, Willink R, Metrologia, 2001, **38**, n°1, 9-15

[6] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, *Evaluation of measurement data*— Supplement 1 to the '*Guide to the Expression of Uncertainty in Measurement*'—Propagation of distributions using a Monte Carlo method, to be published