An Introduction to 'The Guide to the Expression of Uncertainty in Measurement'

David MARTIN, France

Key words: Quality, Traceability, Uncertainty, GUM

SUMMARY

Organisations are created to fulfil some need. They succeed when they satisfy the requirements and expectations of their stakeholders. Stakeholders are the people or organisations that can give or take something from an enterprise. They include government, suppliers, society, employees, and most particularly customers.

The customer is a special stakeholder. The customer is the person, or organisation that receives a product or service. The customer is the one who pays. And only the customer can decide if products or services are satisfactory. Customers require quality products and services delivered on time and at a price that reflects value for money.

Quality products are reliable, functional, durable, secure, available, and traceable - among many other things. Quality services reflect competence, responsiveness, integrity, reliability, and credibility. Quality is the degree to which a product or service fulfils a set of requirements: a requirement being a need or expectation.

Surveyors must provide legally accurate and precise information to their customers. Typically they will strive to do this in an optimal cost effective way and with the most appropriate instrumentation. Naturally this requires a good understanding and assurance in the instruments that are used.

All instruments are subject to measurement error or uncertainty. Measurement uncertainty can be an essential element in the professional decision making process. Additionally, as tolerances become more demanding, the role of measurement uncertainty is becoming more important in determining conformity. Measurement uncertainty almost always plays a central role in quality assessment and quality standards.

The Guide to the Uncertainty in Measurement - colloquially referred to as the GUM - is an internationally recognised guide that addresses uncertainty in measurement. Because Surveyors are so consummately involved in measurement, an understanding of the role of uncertainty is essential. This paper aims to provide an introductory discussion of the GUM in the context of the surveying profession.

An Introduction to 'The Guide to the Expression of Uncertainty in Measurement'

David MARTIN, France

1. INTRODUCTION

In their day to day work surveyors use a variety of measurements including distances, angles, height differences and Global Navigation Satellite Systems (GNSS) positions, both in Real Time Kinematic (RTK) and linked with Continuous Operating References Stations (CORS). Generally these measurements have some legal or economic value. Logically they should be traceable and provide some statement of the uncertainty associated with them.

The *Guide to the Expression of Uncertainty in Measurement* (GUM) is designed to promote the sound evaluation of measurement uncertainty through an agreed upon set of rules. The Joint Committee for Guides in Metrology (JCGM) is responsible for the GUM and its supplements.[1][1] The JCGM comprises eight member bodies - the International Organization for Standardization (ISO), the Bureau International des Poids et Mesures (BIPM), the International Electrotechnical Commission (IEC), the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC), and the International Union of Pure and Applied Chemistry (IUPAC), the International Union of Pure and Applied Physics (IUPAP), the International Organization of Legal Metrology (OIML) and the International Laboratory Accreditation Cooperation (ILAC). The GUM and the GUM Supplements are used in conjunction with the *International Vocabulary of Metrology Basic and general concepts and associated terms* [2] commonly referred to as the VIM.

2. BACKGROUND

2.1 International System of Units/ Le Système International d'Unités (SI)

There are seven official base units defined by the SI; the unit of length (metre), the unit of mass (kilogram), the unit of time (second), the unit of electric current (ampere), the unit of thermodynamic temperature (kelvin), the unit of amount of substance (mole), and the unit of luminous intensity (candela).[3]

In addition to the seven base units, there are a number of SI derived units which are defined uniquely only in terms of SI base units. For example, the coherent SI derived unit of resistance, the ohm, symbol Ω, is uniquely defined by the relation $\Omega = m^2 \cdot kg \cdot s^{-3} \cdot A^{-2}$, which follows from the definition of the quantity electrical resistance. Another example is the unit of plane angle (radian). One radian is the angle subtended at the centre of a circle by an arc of

[1] All of the GUM documents are available free of cost for download at http://www.bipm.org/en/publications/guides/gum.html.

circumference that is equal in length to the radius of the circle. There is 2π radian in a full circle. The radian is expressed in terms of the SI base unit the metre and is defined as $m \cdot m^{-1}$ (i.e. it is a dimensionless unit). (p. 118 [3])

It is important to distinguish between the definition of a unit and its realization. The definition of each base unit of the SI is carefully drawn up so that it is unique and provides a sound theoretical basis upon which the most accurate and reproducible measurements can be made. The realization of the definition of a unit is the procedure by which the definition may be used to establish the value and associated uncertainty of a quantity of the same kind as the unit. The realization of the value of a quantity of a SI unit or derived unit is typically made by National Metrology Institutes (NMIs). Examples of NMIs are; the National Physical Laboratory (NPL) in the UK, the Physikalisch-Technische Bundesanstalt (PTB) in Germany, and the National Institute of Standards and Technology (NIST) in the USA.

Traceability links measurements made by an instrument directly to quantities and units defined by the SI. Traceability is a method of ensuring that a measurement with its uncertainties is an accurate representation of what it is trying to measure. With traceability, it is possible to demonstrate an unbroken chain of comparisons that ends at a NMI and its realization of the definition of a unit.

2.2 Agreements Underpinning the SI

The basis of the SI system of units is the Convention of the Metre (Convention du Mètre).[4] The Convention of the Metre is a treaty that created the International Bureau of Weights and Measures (BIPM), an intergovernmental organization under the authority of the General Conference on Weights and Measures (CGPM) and the supervision of the International Committee for Weights and Measures (CIPM). The BIPM acts in matters of world metrology, particularly concerning the demand for measurement standards of ever increasing accuracy, range and diversity; and the need to demonstrate equivalence between national measurement standards.

The Convention was signed in Paris in 1875 by representatives of seventeen nations. As well as founding the BIPM and laying out the way in which the activities of the BIPM should be financed and managed, the Metre Convention established a permanent organizational structure for member governments to act in common accord on all matters relating to units of measurement. The Convention, modified slightly in 1921, remains the basis of international agreement on units of measurement. The BIPM now has 52 Member States, including all the major industrialized countries.

A Mutual Recognition Arrangement (CIPM MRA) for national measurement standards and for calibration and measurement certificates issued by NMIs was signed at a meeting held in Paris on 14 October 1999. The agreement has now been signed by the representatives of 93 institutes – from 52 Member States, 37 Associates of the CGPM, and 4 international organizations – and covers a further 151 institutes designated by the signatory bodies.

Through the MRA and a common statement between BIPM, OIML and ILAC, measurements made by different NMIs and accredited laboratories are recognized between signatories. This means that a calibration certificate issued by a COmité FRançais pour l'Accréditation (COFRAC) accredited laboratory in France is recognized in the UK and a calibration certificate issued by a United Kingdom Accreditation Service (UKAS) accredited laboratory is recognized in France. Similarly, a COFRAC calibration certificate is recognized by the Department of Standards Malaysia (DSM) and JUPEM, the Malaysian Department of Survey and Mapping. The GUM is used extensively by NMIs and accredited laboratories to evaluate measurement uncertainty and to establish traceability in measurement.

3. THE GUM AND IT'S SUPPLEMENTS

3.1 Guide to the Expression of Uncertainty in Measurement (GUM)

The GUM establishes general rules for evaluating and expressing uncertainty in measurement. These rules are intended to be applicable to a broad spectrum of measurements. The uncertainty of a measurement reflects the lack of exact knowledge of the value of the measurand. The measurand is well defined quantity intended to be measured. Even if a measurement is corrected for recognized systematic effects, it is still only an estimate of the value of the measurand. There always remain uncertainties due to random effects and imperfect corrections of the systematic effects. Uncertainty defined in the GUM is a parameter that describes the dispersion of values that could reasonably be attributed to the measurand.

The GUM and its underlying philosophy have been extensively adopted across metrology. The NMIs and industry have invested heavily in developing procedures that comply with it. For this reason, the JCGM is keen that the GUM remains unchanged in the foreseeable future and that any clarification and extensions are published as supplements.[5] Several supplements are now available:

- Evaluation of measurement data – *An introduction to the "Guide to the expression of uncertainty in measurement" and related documents* (JCGM 104:2009),
- Evaluation of measurement data – *Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method* (JCGM 101:2008)
- Evaluation of measurement data – *Supplement 2 to the "Guide to the expression of uncertainty in measurement" – Extension to any number of output quantities* (JCGM 102:2011)
- Evaluation of measurement data – *The role of measurement uncertainty in conformity assessment* (JCGM 106:2012)

Three additional documents (Concepts and basic principles, Supplement 3 to the "Guide to the expression of uncertainty in measurement" – Modelling, and Applications of the least-squares method) are being prepared.

The GUM and its supplements provide different approaches to uncertainty evaluation. The GUM is rooted in probability theory. It uses expectations and the standard deviations of the assumed Probability Density Functions (PDFs) – typically a Gaussian distribution or a scaled and shifted *t*-distribution - to make uncertainty estimates about the quantity being measured. *Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation* *of distributions using a Monte Carlo method* (referred to as GUM1 below) uses the richer information available in the PDFs for the values of the input quantities to a measurement model, rather than just the expectations and standard deviations of these PDFs, to determine the PDF for the value of the output quantity. GUM1 uses a Monte Carlo method (MCM), in which an approximation to the distribution function for measurement model is established numerically by making random draws from the probability distributions for the input quantities, and evaluating the model at the resulting values.[6] MCM has fewer conditions associated with its use than the GUM uncertainty framework. Whereas GUM and GUM1 are only concerned with models having a single scalar output quantity, *Supplement 2 to the "Guide to the expression of uncertainty in measurement" – Extension to any number of output quantities* extends the concepts of the GUM and GUM1 to multivariate measurement models with any number of output quantities. Such quantities are generally mutually correlated because they depend on common input quantities.

The GUM and its supplements propose two stages in the uncertainty evaluation; the formulation and calculation. The formulation stage is unique to each application and requires the development of a model of measurement as a basis for the evaluation of uncertainty. This model can have any number of input quantities. With the GUM and GUM1 uncertainty frameworks, the model produces a single output quantity referred to as the measurand. The approaches of the GUM and GUM1 diverge at the calculation stage.

There are (at least) three methods to calculate uncertainty. The analytical approach is the method of choice when it can be applied. It does not introduce any approximation. However, it is only applicable in relatively simple cases. The approach taken by the GUM uncertainty framework is an approximation to the analytical method. GUM1 uses the Monte Carlo simulation method.[6]

3.2 The GUM Uncertainty Evaluation Approach

The GUM uncertainty framework is founded on probability theory where information regarding measured quantities is characterized by probability distributions. At the time of its writing, the evaluation of uncertainty for general (linear or non-linear) models was considered too complex to form its basis. For this reason, the GUM provided a simplified approach, the so-called GUM uncertainty framework. This framework linearizes the measurement model about the best available estimates of the input quantities. Rather than working with the distributions themselves, the approach uses summarizing parameters of the distributions, namely, expectations (means), and standard deviations. The expectations and standard deviations are propagated through the linearized model. A Gaussian distribution (or a scaled and shifted *t*-distribution) is then used to characterize the output quantity in order to obtain a coverage interval corresponding to a stipulated coverage probability.[7] The justification for the use of the Gaussian distribution is the invocation of the well know central limit theorem.

The uncertainty of measurement generally consists of several components which the GUM approach groups into two categories according to the method used to estimate their numerical values:

- Type A: method of evaluation of uncertainty by the statistical analysis of series of observations,

– Type B: method of evaluation of uncertainty by means other than the analysis of series of observations.[8]

Broadly speaking, a Type A determination of uncertainty will be made with a series of measurements of the measurand. Typically, Type B determinations of uncertainty will incorporate uncertainties determined by manufacturer's specifications, calibration certificates, recognized handbooks or simply experience with, or general knowledge of the behaviour of materials or instruments.

Individual uncertainties whether they are Type A or Type B are combined together by applying the law of propagation of uncertainty (i.e. taking the square root of the sum of the variances). This combined uncertainty is then expressed in terms of an expanded uncertainty. The expanded uncertainty is obtained by multiplying the uncertainty by a coverage factor – typically 2.

There are no conditions for the valid application of the GUM uncertainty framework for linear models. However, there are conditions, outlined in the GUM1, to its validity when applied to non-linear models. (p. 13 [9]) This is considered one of its primary weaknesses that GUM1 aims to overcome with the more comprehensive Monte Carlo simulation approach.

3.3 The GUM1 approach

"The Monte Carlo simulation method provides a general approach to obtain an approximate numerical representation $\mathbf{G}$, say, of the distribution function $G_Y(\eta)$ for Y. The heart of the approach is repeated sampling from the PDFs for the X_i and the evaluation of the model in each case. Since $G_Y(\eta)$ encodes all the information known about Y, any property of Y such as expectation, variance and coverage intervals can be approximated using $\mathbf{G}$. The quality of these calculated results improves as the number of times the PDFs are sampled increases.

Expectations and variances (and higher moments) can be determined directly from the set of model values obtained. The determination of coverage intervals requires these model values to be ordered. If y_r, for $r = 1 \ldots M$, represent M model values sampled independently from a probability distribution for Y, then the expectation $E(Y)$ and variance $V(Y)$ can be approximated using the y_r. In general, the moments of Y (including $E(Y)$ and $V(Y)$) are approximated by those of the sampled model values. Let M_{y_0} denote the number of y_r that are no greater than y_0, any prescribed number. The probability $\Pr(Y \le y_0)$ is approximated by M_{y_0}/M. In this way, the y_r provide a step function (histogram-like) approximation to the distribution function $G_Y(\eta)$. Each y_r is obtained by sampling at random from each of the PDFs for the X_i and evaluating the model at the sampled values so obtained. $\mathbf{G}$, the primary output from MCM, constitutes the y_r arranged in strictly increasing order." (GUM1 p14 [9])

4. DISCUSSION

The key to establishing traceability with any type of measurement lies in somehow linking it to one of the SI base units. The first step is to define the measurand. If the measurand is not fully defined then its uncertainty and hence traceability cannot be established.

The GUM states that "The uncertainty of the result of a measurement reflects the lack of exact knowledge of the value of the measurand.... The result of a measurement after correction for recognized systematic effects is still only an estimate of the value of the measurand because of the uncertainty arising from random effects and from imperfect correction of the result for systematic effects.

NOTE The result of a measurement (after correction) can unknowably be very close to the value of the measurand (and hence have a negligible error) even though it may have a large uncertainty. Thus the uncertainty of the result of a measurement should not be confused with the remaining unknown error.

In practice, there are many possible sources of uncertainty in a measurement, including:

a) incomplete definition of the measurand;
b) imperfect realization of the definition of the measurand;
c) non-representative sampling — the sample measured may not represent the defined measurand;
d) inadequate knowledge of the effects of environmental conditions on the measurement or imperfect
e) measurement of environmental conditions;
f) personal bias in reading analogue instruments;
g) finite instrument resolution or discrimination threshold;
h) inexact values of measurement standards and reference materials;
i) inexact values of constants and other parameters obtained from external sources and used in the
j) data-reduction algorithm;
k) approximations and assumptions incorporated in the measurement method and procedure;
l) variations in repeated observations of the measurand under apparently identical conditions.

These sources are not necessarily independent, and some of sources a) to i) may contribute to source j). Of course, an unrecognized systematic effect cannot be taken into account in the evaluation of the uncertainty of the result of a measurement but contributes to its error." [8]

We will consider GNSS measurements as an example. GNSS measurements are inherently complex. If one considers the full measurement chain comprising the control, the space and the user segments and all of their associated error sources, the definition of the measurand becomes difficult or even impossible. Although it may feasible to establish traceability by including all of these elements, the actual realisation would be exceedingly complicated. Developing a consensus on the definition of this measurand would be difficult. On the other, if one considers the measurand is the baseline length, the problem becomes more tractable. Length is one of the base units of the SI.

It is recognized that there are many valid arguments against using this level of simplicity. Before continuing, however, it must be noted that this definition of the measurand does not mean that common errors are not considered at all. Corrections to known errors must be made if they are available and it is possible to make them. Models exist to correct for ionospheric and tropospheric errors. Similarly, satellite and antenna phase centre variations (PCV) can be modelled. [10, 11] When appropriate these corrections must be incorporated into the baseline length determination.

However, these corrections are made upstream of, and not specifically as part of the actual baseline length measurement. They are not explicitly considered as part of the measurand. This approach is in fact in accordance with the GUM, which stipulates that all recognized systematic effects must be a-priori corrected for. Similarly, the coordinates used in and the results of the network calculations issued from the measurements are not considered as part of the measurand; nor are the personnel that make the measurements.

Nevertheless, the effects of all of these different error sources must be incorporated into the uncertainty calculation. This is done with calibration certificates, uncertainty estimates, common knowledge, best practice etc... These effects are combined into what is referred to as the Type B contribution to the uncertainty.

If the measurand is defined as the baseline distance, then an uncertainty calculation can be made and traceability to the metre, a base unit of measurement established. Several different ways of establishing traceability can be envisaged. However, there is an existing ISO standard, ISO 17123 part 8,[12] that can be used as a starting point. The advantage of using the standard to provide basic guidelines upon which to build traceability in GNSS measurement is that it exists and as such has achieved consensus by a large number of concerned parties.

The implementation of ISO 17123 is discussed in [13]. Specifically the standard stipulates "The test field consists of a base point and two rover points. The location of the rover points shall be close to the area of the task concerned. The separation of two rover points shall be a minimum of 2 m and shall not exceed 20 m. The positions of two rover points may be selected at convenience in the field .... The horizontal distance and height difference between two rover points shall be determined by methods with precision better than 3 mm other than RTK. …"

A GNSS calibration field/facility could be established using the principles outlined in this standard and using calibrated instruments (e.g. total stations and/or levels) with established traceability to determine the distances and height differences between the base point and the two (or possibly more) *permanent* rover points. These instrument uncertainties as well as other contributions from refraction and possibly uncertainty in latitude, longitude and ellipsoidal height of a fixed pillar could be combined into a Type B uncertainty. The repeatability of the coordinate determinations the rover points should be made several times and incorporated into the Type B uncertainty as well. The Type A uncertainty must be established by repeated independent measures of the rover points using GNSS antennas. Different Type A uncertainties could be established using different instrument types.

The final uncertainty U of the calibration is determined by combining the Type A and Type B uncertainties and multiplying by a coverage factor (typically 2).

$$U = 2 \times \sqrt{(\text{Type A})^2 + (\text{Type B})^2}$$

These last paragraphs give a very broad outline to a possible method to establish traceability in GNSS measurements. Several GNSS test facilities exist. One well documented example is discussed in [14].

It is worth noting that a variant in this type of approach is used by JUPEM, the Malaysian Survey and Mapping Directorate to establish traceability in the Malaysian cadastral system. [15] Distance meters used by JUPEM to this end are calibrated at the ESRF accredited calibration bench.

5. SUMMARY

All measurements are subject to uncertainty. Measurement uncertainty is an essential element in the professional decision making process - particularly in assessing conformity. Measurement uncertainty almost always plays a central role in quality assessment and quality standards.

Surveyors are intimately concerned with measurement and the quality of the results of their measurements. The Guide to the Uncertainty in Measurement – commonly referred to as the GUM – and its supplements provide an internationally recognised way to assess uncertainty in measurement and provide an estimate of the quality of the surveyor's measurements.

REFERENCES

1. JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement and supplements* http://www.bipm.org/en/publications/guides/gum.html, J.C.f.G.i. Metrology, Editor. 2013, BIPM: Sèvres, France.
2. JCGM/WG2, *JCGM 200:2008: International vocabulary of metrology — Basic and general concepts and associated terms (VIM)* http://www.bipm.org/en/publications/guides/gum.html, J.C.f.G.i. Metrology, Editor. 2008, Joint Committee for Guides in Metrology: Sèvres, France.
3. BIPM, *The International System of Units (SI)* www.bipm.org. 2006, Comité International des Poids et Mesures
4. BIPM. *The Metre Convention* 2008 [cited 2008 November 2008]; Available from: http://www.bipm.org.
5. Cox, M.G., A.B. Forbes, and P.M. Harris. *Bayesian estimation methods in metrology*. 2004: AIP.
6. Cox, M. and P. Harris, *An outline of supplement 1 to the guide to the expression of uncertainty in measurement on numerical methods for the propagation of distributions.* Measurement Techniques, 2005. **48**(4): p. 336-345.
7. Bich, W., M.G. Cox, and P.M. Harris, *Evolution of the 'Guide to the Expression of Uncertainty in Measurement'.* Metrologia, 2006. **43**(4): p. S161-S166.
8. JCGM, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*, J.C.f.G.i. Metrology, Editor. 2008, BIPM: Sèvres.
9. JCGM, ed. *Evaluation of measurement data — Supplement 1 to the "Guide to the expression of uncertainty in measurement" — Propagation of distributions using a Monte Carlo method*. ed. J.C.f.G.i. Metrology. 2006, BIPM: Sèvres.
10. Wübbena, G., et al., *Automated Absolute Field Calibration of GPS Antennas in Real-*

Time, in *Proceedings of the 13th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GPS 2000*. 2000: Salt Lake City, Utah, USA.

11. Zeimetz, P. and H. Kuhlmann. *On the Accuracy of Absolute GNSS Antenna Calibration and the Conception of a New Anechoic Chamber* in *FIG Working Week 2008* 2008. Stockholm: FIG.
12. ISO, *ISO 17123 Optics and optical instruments - Field procedures for testing geodetic and surveying instruments Parts 1 to 8*. 2001-2007, ISO: Geneva Switzerland.
13. Heister, H. *The new ISO standard 17123-8 for checking GNSS field measuring systems* in *FIG Working Week 2008* 2008. Stockholm: FIG.
14. Featherstone, W.E. and M.P. Stewart, *Combined analysis of real-time kinematic GPS equipment and its users for height determination.* Journal of Surveying Engineering-Asce, 2001. **127**(2): p. 31-51.
15. Ses, S., et al., *Potential use of GPS for cadastral surveys in Malaysia*, in *40th Aust. & 6th S.E.Asian Surveyors Congress*. 1999: Fremantle, Australia. p. 176-184.

BIOGRAPHICAL NOTES

David Martin David Martin is head of the ESRF Alignment and Geodesy Group. He holds an MSc in Surveying from the University College London and a PhD in Engineering from the University of Warwick. His main interests include accelerator alignment, ground stability, hydrostatic levelling systems, and traceability and instrument calibration. He is the chair of International Federation of Surveyors (FIG) Standards Network and FIG Working Group 5.1 Standards, Quality Assurance and Calibration.

CONTACTS

David Martin
European Synchrotron Radiation Facility (ESRF)
71, avenue des Martyrs
CS 40220, 38043 Grenoble Cedex 9
FRANCE
Tel. +33 (0)4 76 88 22 45
Fax +33 (0)4 76 88 23 13
Email: martin@esrf.fr
Web site: http://www.esrf.fr