GUM – Guide to the Expression of Uncertainty in Measurement

– and its possible use in geodata quality assessment

C.-G. PERSSON

Lantmäteriet, SE-801 82 Gävle, Sweden, clas-goran.persson@lm.se

Presented at Q-KEN, Riga, Latvia, 25th of October 2011

1 Introduction

The objective of the paper and the lecture is to give an introduction to GUM – Guide to the Expression of Uncertainty in Measurement.

It is also to convince the readers and the audience that this is something one needs to know a little about, and make a decision on what attitude one should take to it.

And finally, to strengthen the quality assessment concept within the "geodata industry", regarding accuracy-/uncertainty.

The author and lecturer is senior geodesist at Lantmäteriet (The Swedish Mapping, Cadastral and Land Registration Authority) in Gävle <u>and</u> Adjoint Professor in Applied Geodesy at the Royal Institute of Technology (KTH) in Stockholm.

2 GUM and JCGM

The first GUM embryo was published in 1980. This was an initiative of BIPM, the Bureau International des Poids et Mesures. The first official GUM document was produced in 1992.

In 1997 a Joint Committee for Guides in Metrology (JCGM) was created by seven international organizations:

- **BIPM**

- **IEC,** the International Electrotechnical Commission

- **IFCC,** the International Federation of Clinical Chemistry and Laboratory Medicine

- **ILAC,** the International Laboratory Accreditation Cooperation

- **ISO,** the International Organization for Standardization

- **IUPAC,** the International Union of Pure and Applied Chemistry

- **IUPAP,** the International Union of Pure and Applied Physics

- **OIML,** the International Organization of Legal Metrology).

Since then JCGM has been responsible for GUM. – So GUM is "more" than an ISO standard!

GUM and GUM related publications:

- **ISO/IEC Guide 98-3:2008** *"Uncertainty of Measurement -- Part 3: Guide to the Expression of Uncertainty in Measurement (GUM: 1995)"*.

(Is today maintained by JCGM as *JCGM 100:2008*.)

JCGM also maintains the document

- ***JCGM 200*** *"International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM)".*

These documents can be found on the web site:

www.iso.org/sites/JCGM/JCGM-introduction.htm

3 The Theoretical Fundament of GUM

Statements like "the accuracy is 2 meters" are abundant – even among geodata professionals – but what does this mean?

Is it a standard deviation (1σ), is it a maximum error (3σ) or is it a 95 % confidence interval (2σ)? Who knows, the measure is not defined! GUM is more precise in that aspect.

In classical statistics and in geodetic-/photogrammetric "theory of errors" much of the theoretic platform is built on "true errors". The measurements are related to these, but the problem is that the true errors seldom are available.

In GUM true errors are not needed because the concept of measurement uncertainty relates only to the observed data themselves (the *observables* or *measurands*).

From the GUM document we quote:

"*Uncertainty (of measurement) is a parameter associated with the result of a measurement that characterizes the dispersion of the values that could reasonably be attributed to the measurand*".

Therefore, according to GUM, *error in measurement* and *error analysis* can be replaced by *uncertainty in measurement* and *analysis of uncertainty*.

4 A Measurement

Let's start with another quote from GUM:

*"The objective of a **measurement** is to determine the **value** of the **measurand**, that is, the value of the **particular quantity** to be measured. A measurement therefore begins with an appropriate specification of the measurand, the **method of measurement**, and the **measurement procedure**.*

*In general, the **result of a measurement** is only an approximation or **estimate** of the value of the measurand and thus is complete only when accompanied by a statement of the **uncertainty** of that estimate."*

In Figure 1 the input-output model for propagation of uncertainty is shown. It defines the measurand – the output quantity – as a function $Y = f(X)$ of the *input quantities* X.
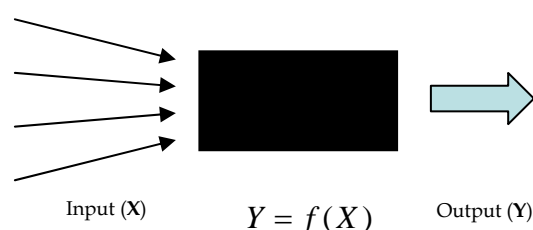

Input (**X**) $Y = f(X)$ Output (**Y**)

Figure 1: *The Input-Output Model for Propagation of Uncertainty.*

5 Type A and Type B Evaluation of Uncertainty

GUM distinguishes between two types of uncertainty evaluation:

Type A: Evaluation of uncertainty by statistical analysis of series of observa-

tions, including quite complex least squares adjustments.

Type B: Evaluation of uncertainty by means other than statistical analysis, e.g. with the use of estimates from previous measurements, specifications from the manufacturer, hand-books, calibration certificates etc.

NB: The classification refers to the way uncertainty is determined – there is no difference in nature or quality and neither type is better than the other. Mixed types occur quite frequently.

6   Standard Uncertainty

6.1   Definition and Notation

Standard uncertainty is usually expressed in terms of the usual *standard deviation*, *root mean square error (RMS)* etc.; use <u>two</u> significant digits.

It is denoted $u(y)$, where y is a result of a measurement or an estimation from several measurements; $u^2(y)$ is used to denote its square, the *variance*.

<u>Examples</u>: "The standard uncertainty in a single measurement" or "the standard uncertainty of the mean" (of repeated measurements).

Standard uncertainty is usually determined with the use of the observed data (Type A), but this is not a necessity; Type B may also be applied.

6.2   Correlation   and   Systematic effects

Although the standard deviation is frequently used for estimating standard uncertainty, one should make a distinction between *uncertainty in measurement* and *precision*.

By contrast with precision, GUM estimates of uncertainty include both

an analysis of *correlations* between the input quantities and an analysis of *systematic effects* in data. Correlations and incompleteness in the modeling of systematic effects have to be taken into consideration – and included in the standard uncertainty estimate.

So once and for all: Precision is not equal to Uncertainty!

6.3   Reporting Standard Uncertainty

GUM gives several alternatives for the reporting of standard uncertainty:

- "L = 2,499 m with a standard uncertainty of 0,0014 m".
- "L = 2,4990(14) m", where the numbers in brackets refers to the standard uncertainty in the last digit of the measurement result.
- "L = 2,499(0,0014) m", where the numbers in brackets is the standard uncertainty in meters.

NB: Do not use the expression $L \pm u$ in connection with standard uncertainty; it should be reserved for *expanded uncertainty* (see below).

7   Combined Uncertainty

Combined uncertainty is an application of the law of propagation of uncertainty in measurement on the function $Y = f(X) = f(X_1, X_2, X_3,...)$:

$$u_c^2(y) = \sum c_i^2 u^2(x_i) = \sum \frac{\delta Y}{\delta X_i} u^2(x_i)$$

where c denotes "combined" and y estimates Y using the estimates $x_1, x_2, x_3,...$ of X.

The partial derivatives c_i (the *sensitivity coefficients*) are determined through *analytical* or *numerical differentiation*. Alternatively, $u_c(y)$ is

directly computed using *Monte Carlo simulation.*

Combined Uncertainty tends to be of Type B, but could be of Type A if <u>all</u> quantities are estimated from the observations.

NB: Sometimes the *c* in $u_c(y)$ and the word "combined" is omitted (and taken for obvious).

8 Expanded Uncertainty

8.1 Definition and Notation

Expanded uncertainty is a quantity defining an interval about the result of a measurement.

This *confidence interval* is expected to encompass a large fraction *p* of the probability distribution characterized by that result and its standard uncertainty.

The fraction *p* is denoted *coverage probability* or *level of confidence.*

To create the interval, the standard uncertainty (or the combined standard uncertainty) is multiplied by a *coverage factor k.*

The expanded uncertainty is denoted $U(y) = k \cdot u(y)$ or $U(y) = k \cdot u_c(y)$.

8.2 Reporting Expanded Uncertainty

Report the *standard uncertainty* and the *coverage factor* as well as the resulting *expanded uncertainty.* Use, with advantage, the $L \pm U$ mode of expression.

Also report the estimated *level of confidence* (in %), which could be expressed in text or as suffixes, e.g:

$$U_{95}(y) = k_{95} \cdot u_c(y)$$

8.3 Examples of Coverage Factors

Here are some examples of coverage factors ($k_{\%}$) for expanded uncertainty:

- Normal distribution, $\lambda_{95} = 1{,}96$.

- t-distribution, $t_{95}(10) = 2{,}23$ (10 degrees of freedom).

- t-approximation using the "effective number of degrees of freedom".

- Monte Carlo simulation and computation of percentiles.

k = 2 could be regarded as the standard GUM k-value. It gives an approximate coverage probability of 95%; deviations should be reported, that is if *k* ≠ *2* or if *k = 2* gives another level of confidence than 95 %.

9 Complete Reports of Uncertainty in Measurement

The following are two examples of complete reports of uncertainty in measurement:

- *m* = (100,02147±0,00079) g, where the number following the symbol $\pm$ is the numerical value of an expanded uncertainty $U = k \cdot u_c$, with *U* determined from a combined standard uncertainty u_c = *0,35* mg and a coverage factor *k = 2,26* based on the t-distribution for 9 degrees of freedom, and defines an interval estimated to have a level of confidence of 95%.

- The positions (p_i) have been determined with the use of Network RTK with an estimated two-dimensional standard un-

certainty $u(p_i)$ = 10 mm. RTK observations is known to have a distribution close to normal and, therefore, a coverage factor k = 2 gives a level of confidence $\geq$ 95%. Thus, the expanded uncertainty of the positions is $U_{95}(y)$ = 20 mm.

10 GUM Tools

And here are some examples of methods and software:

Methods; means, regression analysis, analysis of variance, general least squares adjustments, variance component estimation, Fourier analysis, numerical methods (differentiation etc.), quantitative methods (e.g. Monte Carlo simulation).

Software; Excel, MatLab, specially designed software (for example @Risk).

One example from the simulation software package @Risk is shown in Figure 2.
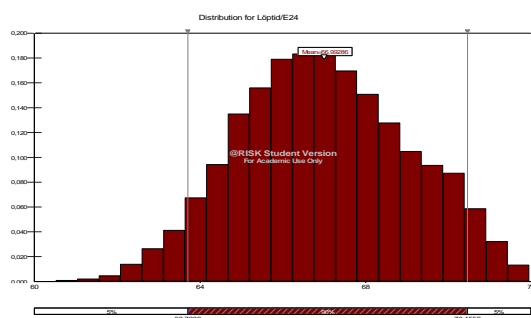

Figure 2: *Monte Carlo-Simulated Probability Distribution from @Risk.*

11 A "GUM Cookbook"

The GUM concept could be summarized in the following "cookbook":

1. Define the relation between the output quantity and all input quantities that can influence on it.

2. Estimate the values of the input quantities.

3. Estimate the standard uncertainties of the input quantities – through statistical analysis or by other means.

4. Determine the sensitivity coefficient that belongs to each input quantity.

5. Calculate the combined uncertainty of the output quantity.

6. Determine a coverage factor that corresponds to the chosen coverage level.

7. Calculate the expanded uncertainty of the output quantity.

8. Report the measurement result together with the expanded measurement uncertainty.

12 Strength, News and Weaknesses

The properties of GUM could be stated in the following way:

Good; a strict terminology and standardized reporting, flexible (Type A and Type B), emphasizes common sense, many examples.

New; numerical methods and Monte Carlo simulations as standard procedures.

Shortcomings; underestimates the impact, and the need for analysis, of correlation.

GUM has many users, is close to practice and provides a basis for the comparison of measurement results through standardized uncertainty statements.

The use of explicit coverage factors makes the quality assessments more precise, and 2σ-intervals (95%) are closer to the intuitive understanding of accuracy/uncertainty than the usual 1σ-expressions. Such intervals are used in for example navigation.

However, in the GUM document you can find the following remainder, reflecting a rather humble attitude:

"Although this Guide provides a framework for assessing uncertainty, it cannot substitute for critical thinking, intellectual honesty, and professional skill. The evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement. The quality and utility of the uncertainty quoted for the result of a measurement therefore ultimately depend on the understanding, critical analysis, and the integrity of those who contribute to the assignment of its value."

13 GUM vs. Geodesy and Geodata

13.1 Today's Situation

Activities dealing with geographic information (*geodata*) have had an appropriate concept for data quality statements and reports for more than 200 years – through geodesy and the work of C F Gauss and others.

So we did not "jump on to the GUM train", and here we are "alone on the platform".

13.2 Why We Need to Know about GUM

We are responsible for a lot of surveying work and capture of geodata, but we speak a different language compared to many of our – existing or potential - users and customers; and these users and customers are in the majority.

There is no real necessity to change the concept, but we need insight and some practical GUM attainments – together with a translation table between GUM and our terminology. This is especially important if we want to include data from new applications into our data bases – and combine them with geographic information.

13.3 What is natural?

As shown in Figure 3, there is a more natural relation between *uncertainty* and standard deviation than between *accuracy* and standard deviation.
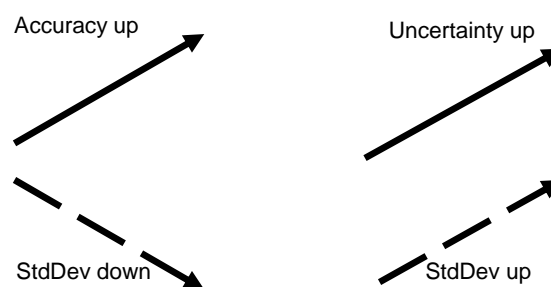

Figure 3: *Higher accuracy gives a lower standard deviation but higher uncertainty gives a higher standard deviation!*

And: In the classical concept we introduce the term *accuracy*, but then we only talk about "errors": Mean errors, gross errors, systematic errors etc. Is that natural?

We also note that 2σ-expressions can be used directly as tolerances for control measurements.

13.4 The Key to Success

We should not see GUM as a problem but as a possibility – to broaden and improve our own concept regarding quality assessment.

The most important action, in the author's opinion, is to express the accuracy/uncertainty part of metadata in terms of GUM – in addition to today's mode of expression; that is, primarily *positional accuracy* and *attribute accuracy* – sorry, *positional uncertainty* and *attribute uncertainty.*

14 References

ISO/IEC Guide 98-3:2008: *"Uncertainty of Measurement -- Part 3: Guide to the Expression of Uncertainty in Measurement (GUM: 1995)"*. (Maintained by JCGM as JCGM 100:2008.)

JCGM 200: *"International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM)"*.

On Lantmäteriet's home page, with address

http://www.lantmateriet.se/templates/LMV_Page.aspx?id=15968

the following, Swedish GUM documents are presented:

Persson C-G (2010a): *"Några vanliga fördelningar (Some Common Distributions)"*.

Persson C-G (2010b): *"Sammanlagd mätosäkerhet och kvantitativa metoder (Combined Uncertainty and Quantitative Methods)"*.

Persson C-G (2010c): *"Exempel – Typ A-bestämning (Examples of Type A Evaluation)"*.

Persson C-G (2010d): *"Exempel – Typ B-bestämning (Examples of Type B Evaluation)"*.

Persson C-G (2010e): *"Korrelerade mätningar (Correlated Measurements)"*.

These form the basis of this presentation.