
MEASUREMENT UNCERTAINTY GUIDELINES

A supplement to the MSL Quality Manual

April 10, 2024

Text in this colour is awaiting approval by the Quality Council.

Text in this colour is deprecated and awaiting approval to be deleted.

Measurement Quality Council

Measurement Standards Laboratory of New Zealand

Contents

1	The GUM.....	1
1.1	GUM Supplements	1
1.1.1	Multidimensional measurands	2
2	Reporting expanded uncertainty.....	3
2.1	Rounding expanded uncertainty	3
2.1.1	Changing units	3
2.2	Unphysical limits and expanded uncertainty.....	4
3	Fixed-digit indications.....	5
3.1	Rounding	5
3.1.1	Single readings	5
3.1.2	Multiple readings	6
4	Uncertainty style	7
4.1	Background.....	7
4.2	Recommendations.....	7
4.2.1	Only the measured value is an estimate	7
4.2.2	Use the term 'expanded uncertainty'	8
4.2.3	The level of confidence is assumed known	8
A	References.....	9

1 The GUM

The Guide to the Expression of Uncertainty in Measurement (GUM) is our basic guidance document for the calculation of measurement uncertainty.

The current GUM edition was released in 2008 and is the first edition published by the BIPM JCGM. It has very minor corrections to earlier ISO editions published in 1995 and 1993.

Previous editions had to be purchased from the ISO, but the GUM is now available free of charge. A PDF copy can be downloaded:

http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf

and an HTML version is available:

<http://www.iso.org/sites/JCGM/GUM/JCGM100/C045315e-html/C045315e.html?csnumber=50461>

Bibliographic reference the following (from Metrologia) is suitable for academic publications:

BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML 2008, “*Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement JCGM 100:2008 (GUM 1995 with minor corrections)*”, 1st edition (Sèvres: BIPM Joint Committee for Guides in Metrology).

GUM reference for calibration reports the following is sufficient:

See the “*Guide to the expression of uncertainty in measurement JCGM 100:2008*” (BIPM, 1st edition, 2008) for an explanation of terms.

While the title is incomplete in this shorter reference, a Google search on the text in parentheses, or the quoted text, returns a link to the on-line versions immediately.

1.1 GUM Supplements

Two GUM Supplements have also been released: GUM Supplement 1 (S1) [2] and GUM Supplement 2 (S2) [3]. Use of S1, and the (second) part of S2 that deals with Monte Carlo methods, is not recommended until serious issues have been resolved relating to the meaning of probability in

uncertainty statements [8]. The interpretation of probability used in the supplements is Bayesian. So, a physically meaningful interpretation cannot be given to uncertainty statements obtained from using the supplement Monte Carlo methods.

1.1.1 Multidimensional measurands

GUM Supplement 2 addresses the evaluation of measurement uncertainty for multidimensional measurands, which is not covered in the GUM or S1. The first part of S2 describes extensions to the analytical techniques described in the GUM. These extensions are compatible with the conventional interpretation of probability, as a “relative frequency of events”, and so they may safely be used. A more extensive review of multidimensional extensions to the GUM has been published [4].

2 Reporting expanded uncertainty

2.1 Rounding expanded uncertainty

The expanded uncertainty in a measurement result should be reported to two significant figures and the associated numerical result should be reported using the same precision. Reporting fewer than two significant figures in the expanded uncertainty can significantly affect the coverage probability (level of confidence).

Reporting more than two significant figures is acceptable when there is good reason to do so. The associated result should be reported using the same numerical precision as the uncertainty.

There are situations where it may be desirable to round a value of expanded uncertainty to one significant figure. This is not recommended, but is acceptable when the least significant digit remaining is 5 or above. In such cases, rounding will introduce an error that is no more than 10 % of the unrounded value. There is no significant impact to the coverage probability in this case [5].

The practice of consistently rounding the uncertainty up is not recommended (the reason for doing this might be to 'err on the safe side' by ensuring that a conservative interval is reported). For example, 0.1234 should be not rounded to 0.13. Values should be rounded to the nearest value in the chosen number of significant figures. For example, 0.1234 rounded to two significant figures is 0.12, and 0.1251 is 0.13.

There is a NIST Good Laboratory Practice Guide [6] with examples reported to two significant figures. (The additional step of even/odd rounding in GLP 9 (2.3) is optional.)

2.1.1 Changing units

In some technical areas, a client may prefer results to be reported in units that are different from those best suited to the evaluation of measurement uncertainty. An example of this is the logarithmic units used in optical and electrical measurements.

In such cases, the usual data processing should be followed, producing a result and an expanded uncertainty in the most appropriate metrological units. Then, as a final step, the result and the limits of the expanded uncertainty interval can be transformed into the units requested by the client.

This procedure does not alter the level of confidence (coverage probability) of the expanded uncertainty. However, when a non-linear transformation is involved in the unit change, the transformed result may not be in the centre of the transformed expanded uncertainty interval, and one or other of the interval limits may seem strange (for example, a limit of infinity is possible).

The best way of reporting this information to the client needs to be considered, as does the process

by which the client can use information in the report to extract values for the standard uncertainty and degrees of freedom.

2.2 Unphysical limits and expanded uncertainty

On occasions, the expanded uncertainty calculated for a measurement result can lead to an uncertainty interval that covers unphysical values. For example, negative values for a quantity such as mass or optical density which must not be less than zero. This may happen, for example, if a measurement is sensitive to the variability in a sample of data.

Unphysical values should not generally be reported to a client. Such cases need to be carefully considered and the reason for the 'unphysical' result should be clear in terms of statistical fluctuations of influence quantities in the measurement process.

A procedure for adjusting the expanded uncertainty by removing unphysical values should be envisaged, with a clear understanding of how the 'unphysical' result arose.

In the simplest case, when a measurement model is linear, it may be acceptable to reset one of the expanded uncertainty interval limits. For instance, if an expanded uncertainty from -0.01 to 0.10 is obtained from a measurement of a quantity that must be greater than zero, the negative values may be removed to form an interval from 0.00 to 0.10. The level of confidence (coverage probability) is not changed by this operation.

A more sophisticated approach is possible for linear models when there are infinite degrees of freedom [9]. Table X in that paper describes limits of 95 % confidence intervals (expanded uncertainties), given the sample mean and sample standard deviation.

In any case, the best format for reporting to the client needs to be considered, as does the process by which the client can extract a value for the standard uncertainty and for the degrees of freedom from the information given in the report.

3 Fixed-digit indications

3.1 Rounding

The rounding guidelines in NIST Good Laboratory Practice Guide [6] apply when instrument readings are formatted in a fixed number of digits.

3.1.1 Single readings

The expression of uncertainty in a single indication should not be constrained by the format of the data.

For example, suppose an indication is limited to integer values and that the number 101 is observed. It is not recommended to report the result as 101, $U = 1$. The standard uncertainty (of a uniform distribution of error between 0 and 1) is 0.29. So with a coverage factor $k = 2$, this result would be reported as 101.00, $U = 0.58$ (or perhaps as 101.0, $U = 0.6$).

Why report more digits than the instrument can display? The extra digits are needed because the uncertainty of the reading, as an estimate of the quantity being measured, determines the number of significant figures required.

For the case of an instrument that displays truncated readings, a simple measurement model is

$$Y = w + E ,$$

where w is the reading, which is known, and E is the truncation error, which is unknown. The error can be described statistically, using a uniform distribution with zero mean and a range equal to the least significant digit displayed.

The reported value is y , an estimate of Y . This is the sum of w and the best estimate of E , which is zero. So,

$$y = w + 0 .$$

The uncertainty of y , as an estimate of Y , is found by combining the uncertainty associated with w , which is zero (we know exactly what the reading is), with the uncertainty associated with taking zero as an estimate of E , which is 0.29. So,

$$u(y) = \sqrt{(0^2 + 0.29^2)} = 0.29 .$$

3.1.2 Multiple readings

The uncertainty associated with the mean of a series of readings with a fixed number of decimal places can be evaluated by a simple method that takes advantage of variations in the data to reduce the size of the uncertainty [7]. Section 3 of Willink's paper describes the best way of processing data in this situation (a view).

This method has been independently corroborated as best-practice by NIST. It is available in GTC software as the function `type_a.estimate_digitised()`. The following example of use is in the GTC documentation:

```
# LSD = 0.0001, data varies between -0.0055 and -0.0057
>>> seq = (-0.0056,-0.0055,-0.0056,-0.0056,-0.0056,
...        -0.0057,-0.0057,-0.0056,-0.0056,-0.0057,-0.0057)
>>> type_a.estimate_digitized(seq,0.0001)
ureal(-0.005627272727272727, 1.9497827808661157e-05, 10)
```

A slightly simplified version of the result given in [7], in the case where a set of differences between a reference instrument and a device under test (DUT) is averaged to determine a correction, is to assign the uncertainty in this correction as either s/\sqrt{N} or $\delta/\sqrt{12}$, whichever is larger, where s is the standard deviation of the N differences and δ is the resolution of the DUT. This is appropriate advice for second-tier calibration laboratories provided that N is at least 4.

4 Uncertainty style

4.1 Background

A measured value is an estimate of the measurand. The measurement is subject to error, and it is that unknown error that leads to uncertainty: the measured value is only an approximation of the measurand. So, for a measurand Y and some measured value of the measurand y , the standard uncertainty $u(y)$ is, strictly speaking, *the uncertainty of y as an estimate of Y* .

A combined standard uncertainty is calculated using the measurement model and the GUM rules for uncertainty propagation, with information about influence quantities.

Expanded uncertainty is of a different nature to standard uncertainty: an expanded uncertainty is used to make an inference about the measurand, whereas a standard uncertainty describes the variability of something that is inherently unpredictable, like the measurement error.

An expanded uncertainty is also qualified by a level of confidence, or coverage probability, whereas a standard uncertainty is not. There is no right or wrong value for the expanded uncertainty at a particular level of confidence.

To calculate an expanded uncertainty, one needs to know the distribution of measurement error (the default GUM assumption is Gaussian), the combined standard uncertainty associated with the result, the degrees of freedom associated with the combined standard uncertainty and the required level of confidence.

In principle, we wish to report sufficient information for clients to work out (if not provided in the report): the combined standard uncertainty, the degrees of freedom, the expanded uncertainty and the associated level of confidence.

4.2 Recommendations

4.2.1 Only the measured value is an estimate

The primary outcome of a measurement is an estimate of the measurand. So, it is unhelpful to use the term 'estimate' in connection with statements about values of uncertainty too. This should be avoided: it is better to say that uncertainty has been 'calculated', 'evaluated', etc, but not 'estimated'.¹

Instead of:

These uncertainties are estimated by combining the uncertainties of the calibration process . . .

¹Many authoritative texts use 'estimated', but this is a bad habit that we should avoid. The MSL Guidelines on Reporting and Publishing discusses this further in the 'Uncertainty' section.

Prefer:

These uncertainties were calculated by combining the components of uncertainty associated with the calibration process . . .

4.2.2 Use the term ‘expanded uncertainty’

Our calibration reports refer to the current edition of the GUM for information about the terms we use. So it is unhelpful to abbreviate ‘expanded uncertainty’ unless the full term has been used elsewhere in context.

Instead of:

The uncertainty is based on a coverage factor of . . .

Prefer:

The expanded uncertainty is calculated with a coverage factor . . .

4.2.3 The level of confidence is assumed known

Given the level of confidence for a particular coverage factor, it is possible to work out the corresponding standard uncertainty and degrees of freedom (assuming a Gaussian error distribution).

The primary reason for reporting the level of confidence (or coverage probability) is to allow a client to work out degrees of freedom, so avoid language that would suggest otherwise.

Instead of:

. . . expanded uncertainties were calculated using a coverage factor $k = 2.1$ and define an interval estimated to have a 95 % level of confidence.

Prefer:

. . . expanded uncertainties were calculated using a coverage factor $k = 2.1$ for a 95 % level of confidence.

Nonetheless, if we consider the intended meaning of level of confidence, we can only expect the actual level of confidence to approximate the nominal value in practice. Situations can occur where we are confident that the actual coverage of a measurement procedure will be higher than the nominal value. For example, a value of the expanded uncertainty may be increased so that it falls within our current IANZ scope.

In such cases, do not write statements like ‘. . . at least 95 % level of confidence . . .’. Rather, simply write the report as if the expanded uncertainty were correct or include additional information to clarify how the client can use the information reported.

A References

- [1] *Guide to the Expression of Uncertainty in Measurement, JCGM 100:2008 (GUM 1995 with minor revisions)*; [online](#).
- [2] *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; [online](#).
- [3] *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; [online](#).
- [4] B D Hall, *Evaluating the measurement uncertainty of complex quantities: a selective review*, Metrologia 53 (2015) **S25**; [DOI](#).
- [5] B D Hall, *Rounding the expanded uncertainty*, Accred Qual Assur **24**, 369–373 (2019); [DOI](#).
- [6] *Good Laboratory Practice for Rounding Expanded Uncertainties and Calibration Values* (NIST, Sept 2014); [online](#).
- [7] R Willink, *On the uncertainty of the mean of digitized measurements*, Metrologia, 44 (2007) 73-81; [DOI](#).
- [8] D R White, *In pursuit of a fit-for-purpose uncertainty guide*, Metrologia 53 (2016) S107 – S124; [DOI](#).
- [9] G J Feldman and R D Cousins, *A Unified Approach to the Classical Statistical Analysis of Small Signals*, 1999 ; [online](#).