GUIDE 98-3

Uncertainty of measurement —
Part 3:
Guide to the expression of uncertainty in measurement (GUM:1995)

Incertitude de mesure —
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreword</td>
<td>vi</td>
</tr>
<tr>
<td>0 Introduction</td>
<td>viii</td>
</tr>
<tr>
<td>1 Scope</td>
<td>1</td>
</tr>
<tr>
<td>2 Definitions</td>
<td>2</td>
</tr>
<tr>
<td>2.1 General metrological terms</td>
<td>2</td>
</tr>
<tr>
<td>2.2 The term “uncertainty”</td>
<td>2</td>
</tr>
<tr>
<td>2.3 Terms specific to this Guide</td>
<td>3</td>
</tr>
<tr>
<td>3 Basic concepts</td>
<td>4</td>
</tr>
<tr>
<td>3.1 Measurement</td>
<td>4</td>
</tr>
<tr>
<td>3.2 Errors, effects, and corrections</td>
<td>5</td>
</tr>
<tr>
<td>3.3 Uncertainty</td>
<td>5</td>
</tr>
<tr>
<td>3.4 Practical considerations</td>
<td>7</td>
</tr>
<tr>
<td>4 Evaluating standard uncertainty</td>
<td>8</td>
</tr>
<tr>
<td>4.1 Modelling the measurement</td>
<td>8</td>
</tr>
<tr>
<td>4.2 Type A evaluation of standard uncertainty</td>
<td>10</td>
</tr>
<tr>
<td>4.3 Type B evaluation of standard uncertainty</td>
<td>11</td>
</tr>
<tr>
<td>4.4 Graphical illustration of evaluating standard uncertainty</td>
<td>15</td>
</tr>
<tr>
<td>5 Determining combined standard uncertainty</td>
<td>18</td>
</tr>
<tr>
<td>5.1 Uncorrelated input quantities</td>
<td>18</td>
</tr>
<tr>
<td>5.2 Correlated input quantities</td>
<td>21</td>
</tr>
<tr>
<td>6 Determining expanded uncertainty</td>
<td>23</td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>23</td>
</tr>
<tr>
<td>6.2 Expanded uncertainty</td>
<td>23</td>
</tr>
<tr>
<td>6.3 Choosing a coverage factor</td>
<td>24</td>
</tr>
<tr>
<td>7 Reporting uncertainty</td>
<td>24</td>
</tr>
<tr>
<td>7.1 General guidance</td>
<td>24</td>
</tr>
<tr>
<td>7.2 Specific guidance</td>
<td>25</td>
</tr>
<tr>
<td>8 Summary of procedure for evaluating and expressing uncertainty</td>
<td>27</td>
</tr>
<tr>
<td>Annex A Recommendations of Working Group and CIPM</td>
<td>28</td>
</tr>
<tr>
<td>A.1 Recommendation INC-1 (1980)</td>
<td>28</td>
</tr>
<tr>
<td>A.2 Recommendation 1 (CI-1981)</td>
<td>29</td>
</tr>
<tr>
<td>A.3 Recommendation 1 (CI-1986)</td>
<td>29</td>
</tr>
<tr>
<td>Annex B General metrological terms</td>
<td>31</td>
</tr>
<tr>
<td>B.1 Source of definitions</td>
<td>31</td>
</tr>
<tr>
<td>B.2 Definitions</td>
<td>31</td>
</tr>
<tr>
<td>Annex C Basic statistical terms and concepts</td>
<td>39</td>
</tr>
<tr>
<td>C.1 Source of definitions</td>
<td>39</td>
</tr>
<tr>
<td>C.2 Definitions</td>
<td>39</td>
</tr>
<tr>
<td>C.3 Elaboration of terms and concepts</td>
<td>45</td>
</tr>
<tr>
<td>Annex D “True” value, error, and uncertainty</td>
<td>49</td>
</tr>
<tr>
<td>D.1 The measurand</td>
<td>49</td>
</tr>
<tr>
<td>D.2 The realized quantity</td>
<td>49</td>
</tr>
<tr>
<td>D.3 The “true” value and the corrected value</td>
<td>49</td>
</tr>
<tr>
<td>D.4 Error</td>
<td>50</td>
</tr>
</tbody>
</table>
D.5 Uncertainty ..........................................................51
D.6 Graphical representation ...........................................51

E.1 “Safe”, “random”, and “systematic” .................................54
E.2 Justification for realistic uncertainty evaluations .............54
E.3 Justification for treating all uncertainty components identically ...............................................................55
E.4 Standard deviations as measures of uncertainty ................58
E.5 A comparison of two views of uncertainty .......................59

Annex F Practical guidance on evaluating uncertainty components ..........................................................61
F.1 Components evaluated from repeated observations: Type A evaluation of standard uncertainty ..........61
F.2 Components evaluated by other means: Type B evaluation of standard uncertainty ...............................64

Annex G Degrees of freedom and levels of confidence .................................................................70
G.1 Introduction ..................................................................70
G.2 Central Limit Theorem ..................................................71
G.3 The t-distribution and degrees of freedom .........................71
G.4 Effective degrees of freedom ..........................................73
G.5 Other considerations ........................................................75
G.6 Summary and conclusions .............................................76

Annex H Examples .............................................................79
H.1 End-gauge calibration ...................................................79
H.2 Simultaneous resistance and reactance measurement ..........85
H.3 Calibration of a thermometer ........................................89
H.4 Measurement of activity ................................................93
H.5 Analysis of variance ......................................................98
H.6 Measurements on a reference scale: hardness ................104

Annex J Glossary of principal symbols ..................................109

Bibliography ...........................................................................114

Alphabetical index ................................................................116
This Guide establishes general rules for evaluating and expressing uncertainty in measurement that are intended to be applicable to a broad spectrum of measurements. The basis of the Guide is Recommendation 1 (CI-1981) of the Comité International des Poids et Mesures (CIPM) and Recommendation INC-1 (1980) of the Working Group on the Statement of Uncertainties. The Working Group was convened by the Bureau International des Poids et Mesures (BIPM) in response to a request of the CIPM. The CIPM Recommendation is the only recommendation concerning the expression of uncertainty in measurement adopted by an intergovernmental organization.

This Guide was prepared by a joint working group consisting of experts nominated by the BIPM, the International Electrotechnical Commission (IEC), the International Organization for Standardization (ISO), and the International Organization of Legal Metrology (OIML).

The following seven organizations* supported the development of this Guide, which is published in their name:

- BIPM: Bureau International des Poids et Mesures
- IEC: International Electrotechnical Commission
- IFCC: International Federation of Clinical Chemistry**
- ISO: International Organization for Standardization
- IUPAC: International Union of Pure and Applied Chemistry
- OIML: International Organization of Legal Metrology

Users of this Guide are invited to send their comments and requests for clarification to any of the seven supporting organizations, the mailing addresses of which are given on the inside front cover***.

* Footnote to the 2008 version:
In 2005, the International Laboratory Accreditation Cooperation (ILAC) officially joined the seven founding international organizations.

** Footnote to the 2008 version:
The name of this organization has changed since 1995. It is now:
IFCC: International Federation of Clinical Chemistry and Laboratory Medicine

*** Footnote to the 2008 version:
Links to the addresses of the eight organizations presently involved in the JCGM (Joint Committee for Guides in Metrology) are given on http://www.bipm.org/en/committees/jc/jcgm.
Foreword

In 1977, recognizing the lack of international consensus on the expression of uncertainty in measurement, the world's highest authority in metrology, the Comité International des Poids et Mesures (CIPM), requested the Bureau International des Poids et Mesures (BIPM) to address the problem in conjunction with the national standards laboratories and to make a recommendation.

The BIPM prepared a detailed questionnaire covering the issues involved and distributed it to 32 national metrology laboratories known to have an interest in the subject (and, for information, to five international organizations). By early 1979 responses were received from 21 laboratories. Almost all believed that it was important to arrive at an internationally accepted procedure for expressing measurement uncertainty and for combining individual uncertainty components into a single total uncertainty. However, a consensus was not apparent on the method to be used. The BIPM then convened a meeting for the purpose of arriving at a uniform and generally acceptable procedure for the specification of uncertainty; it was attended by experts from 11 national standards laboratories. This Working Group on the Statement of Uncertainties developed Recommendation INC-1 (1980), Expression of Experimental Uncertainties. The CIPM approved the Recommendation in 1981 and reaffirmed it in 1986.

The task of developing a detailed guide based on the Working Group Recommendation (which is a brief outline rather than a detailed prescription) was referred by the CIPM to the International Organization for Standardization (ISO), since ISO could better reflect the needs arising from the broad interests of industry and commerce.

Responsibility was assigned to the ISO Technical Advisory Group on Metrology (TAG 4) because one of its tasks is to coordinate the development of guidelines on measurement topics that are of common interest to ISO and the six organizations that participate with ISO in the work of TAG 4: the International Electrotechnical Commission (IEC), the partner of ISO in worldwide standardization; the CIPM and the International Organization of Legal Metrology (OIML), the two worldwide metrology organizations; the International Union of Pure and Applied Chemistry (IUPAC) and the International Union of Pure and Applied Physics (IUPAP), the two international unions that represent chemistry and physics; and the International Federation of Clinical Chemistry (IFCC).

TAG 4 in turn established Working Group 3 (ISO/TAG 4/WG 3) composed of experts nominated by the BIPM, IEC, ISO, and OIML and appointed by the Chairman of TAG 4. It was assigned the following terms of reference:

To develop a guidance document based upon the recommendation of the BIPM Working Group on the Statement of Uncertainties which provides rules on the expression of measurement uncertainty for use within standardization, calibration, laboratory accreditation, and metrology services;

The purpose of such guidance is

— to promote full information on how uncertainty statements are arrived at;
— to provide a basis for the international comparison of measurement results.


1) See the Bibliography.

* Footnote to the 2008 version:

In producing this 2008 version of the GUM, necessary corrections only to the printed 1995 version have been introduced by JCGM/WG 1. These corrections occur in Subclauses 4.2.2, 4.2.4, 5.1.2, B.2.17, C.3.2, C.3.4, E.4.3, H.4.3, H.5.2.5 and H.6.2.
This corrected version of ISO/IEC Guide 98-3:2008 incorporates the following corrections:

- on page v, Footnote ** has been corrected;
- in 4.1.1, the note has been indented;
- in the first line of the example in 5.1.5, \( \Delta V \) has been replaced with \( \Delta \bar{F} \);
- in the first lines of B.2 and C.2, Clause 0 has been corrected to Clause 2;
- in G.3.2, (G.1c) has been changed to (G.1c);
- in H.1.3.4, the formatting of the first equation has been improved.
0 Introduction

0.1 When reporting the result of a measurement of a physical quantity, it is obligatory that some quantitative
indication of the quality of the result be given so that those who use it can assess its reliability. Without such
an indication, measurement results cannot be compared, either among themselves or with reference values
given in a specification or standard. It is therefore necessary that there be a readily implemented, easily
understood, and generally accepted procedure for characterizing the quality of a result of a measurement, that
is, for evaluating and expressing its uncertainty.

0.2 The concept of uncertainty as a quantifiable attribute is relatively new in the history of measurement,
although error and error analysis have long been a part of the practice of measurement science or metrology.
It is now widely recognized that, when all of the known or suspected components of error have been
evaluated and the appropriate corrections have been applied, there still remains an uncertainty about the
correctness of the stated result, that is, a doubt about how well the result of the measurement represents the
value of the quantity being measured.

0.3 Just as the nearly universal use of the International System of Units (SI) has brought coherence to all
scientific and technological measurements, a worldwide consensus on the evaluation and expression of
uncertainty in measurement would permit the significance of a vast spectrum of measurement results in
science, engineering, commerce, industry, and regulation to be readily understood and properly interpreted. In
this era of the global marketplace, it is imperative that the method for evaluating and expressing uncertainty
be uniform throughout the world so that measurements performed in different countries can be easily
compared.

0.4 The ideal method for evaluating and expressing the uncertainty of the result of a measurement should
be:

— universal: the method should be applicable to all kinds of measurements and to all types of input data
used in measurements.

The actual quantity used to express uncertainty should be:

— internally consistent: it should be directly derivable from the components that contribute to it, as well as
independent of how these components are grouped and of the decomposition of the components into
subcomponents;

— transferable: it should be possible to use directly the uncertainty evaluated for one result as a component
in evaluating the uncertainty of another measurement in which the first result is used.

Further, in many industrial and commercial applications, as well as in the areas of health and safety, it is often
necessary to provide an interval about the measurement result that may be expected to encompass a large
fraction of the distribution of values that could reasonably be attributed to the quantity subject to measurement.
Thus the ideal method for evaluating and expressing uncertainty in measurement should be capable of readily
providing such an interval, in particular, one with a coverage probability or level of confidence that
corresponds in a realistic way with that required.

0.5 The approach upon which this guidance document is based is that outlined in Recommendation
INC-1 (1980) [2] of the Working Group on the Statement of Uncertainties, which was convened by the BIPM in
response to a request of the CIPM (see Foreword). This approach, the justification of which is discussed in
Annex E, meets all of the requirements outlined above. This is not the case for most other methods
in current use. Recommendation INC-1 (1980) was approved and reaffirmed by the CIPM in its own
Recommendations 1 (CI-1981) [3] and 1 (CI-1986) [4]; the English translations of these CIPM Recommendations
are reproduced in Annex A (see A.2 and A.3, respectively). Because Recommendation INC-1 (1980) is the
foundation upon which this document rests, the English translation is reproduced in 0.7 and the French text,
which is authoritative, is reproduced in A.1.
0.6 A succinct summary of the procedure specified in this guidance document for evaluating and expressing uncertainty in measurement is given in Clause 8 and a number of examples are presented in detail in Annex H. Other annexes deal with general terms in metrology (Annex B); basic statistical terms and concepts (Annex C); “true” value, error, and uncertainty (Annex D); practical suggestions for evaluating uncertainty components (Annex F); degrees of freedom and levels of confidence (Annex G); the principal mathematical symbols used throughout the document (Annex J); and bibliographical references (Bibliography). An alphabetical index concludes the document.

0.7 Recommendation INC-1 (1980) Expression of experimental uncertainties

1) The uncertainty in the result of a measurement generally consists of several components which may be grouped into two categories according to the way in which their numerical value is estimated:

   A. those which are evaluated by statistical methods,

   B. those which are evaluated by other means.

   There is not always a simple correspondence between the classification into categories A or B and the previously used classification into “random” and “systematic” uncertainties. The term “systematic uncertainty” can be misleading and should be avoided.

   Any detailed report of the uncertainty should consist of a complete list of the components, specifying for each the method used to obtain its numerical value.

2) The components in category A are characterized by the estimated variances \( s_i^2 \), (or the estimated “standard deviations” \( s_i \)) and the number of degrees of freedom \( v_i \). Where appropriate, the covariances should be given.

3) The components in category B should be characterized by quantities \( u_j^2 \), which may be considered as approximations to the corresponding variances, the existence of which is assumed. The quantities \( u_j^2 \) may be treated like variances and the quantities \( u_j \) like standard deviations. Where appropriate, the covariances should be treated in a similar way.

4) The combined uncertainty should be characterized by the numerical value obtained by applying the usual method for the combination of variances. The combined uncertainty and its components should be expressed in the form of “standard deviations”.

5) If, for particular applications, it is necessary to multiply the combined uncertainty by a factor to obtain an overall uncertainty, the multiplying factor used must always be stated.
Uncertainty of measurement —

Part 3:
Guide to the expression of uncertainty in measurement (GUM:1995)

1 Scope

1.1 This Guide establishes general rules for evaluating and expressing uncertainty in measurement that can be followed at various levels of accuracy and in many fields — from the shop floor to fundamental research. Therefore, the principles of this Guide are intended to be applicable to a broad spectrum of measurements, including those required for:

— maintaining quality control and quality assurance in production;
— complying with and enforcing laws and regulations;
— conducting basic research, and applied research and development, in science and engineering;
— calibrating standards and instruments and performing tests throughout a national measurement system in order to achieve traceability to national standards;
— developing, maintaining, and comparing international and national physical reference standards, including reference materials.

1.2 This Guide is primarily concerned with the expression of uncertainty in the measurement of a well-defined physical quantity — the measurand — that can be characterized by an essentially unique value. If the phenomenon of interest can be represented only as a distribution of values or is dependent on one or more parameters, such as time, then the measurands required for its description are the set of quantities describing that distribution or that dependence.

1.3 This Guide is also applicable to evaluating and expressing the uncertainty associated with the conceptual design and theoretical analysis of experiments, methods of measurement, and complex components and systems. Because a measurement result and its uncertainty may be conceptual and based entirely on hypothetical data, the term "result of a measurement" as used in this Guide should be interpreted in this broader context.

1.4 This Guide provides general rules for evaluating and expressing uncertainty in measurement rather than detailed, technology-specific instructions. Further, it does not discuss how the uncertainty of a particular measurement result, once evaluated, may be used for different purposes, for example, to draw conclusions about the compatibility of that result with other similar results, to establish tolerance limits in a manufacturing process, or to decide if a certain course of action may be safely undertaken. It may therefore be necessary to develop particular standards based on this Guide that deal with the problems peculiar to specific fields of measurement or with the various uses of quantitative expressions of uncertainty. These standards may be simplified versions of this Guide but should include the detail that is appropriate to the level of accuracy and complexity of the measurements and uses addressed.

NOTE There may be situations in which the concept of uncertainty of measurement is believed not to be fully applicable, such as when the precision of a test method is determined (see Reference [5], for example).

* Footnote to the 2008 version:
Several derivative general and specific applications documents have been published. Non-exhaustive compilations listing these documents can be found on http://www.bipm.org/en/committees/jc/jcgm/wg1_bibliography.html. In addition, up-to-date listings of documents that cite the Guide to the expression of uncertainty in measurement can be found by using the full-text search options on http://www.iso.org/ and http://www.iec.ch/.
2 Definitions

2.1 General metrological terms

The definition of a number of general metrological terms relevant to this Guide, such as “measurable quantity”, “measurand”, and “error of measurement”, are given in Annex B. These definitions are taken from the International vocabulary of basic and general terms in metrology (abbreviated VIM)* [6]. In addition, Annex C gives the definitions of a number of basic statistical terms taken mainly from International Standard ISO 3534-1 [7]. When one of these metrological or statistical terms (or a closely related term) is first used in the text, starting with Clause 3, it is printed in boldface and the number of the subclause in which it is defined is given in parentheses.

Because of its importance to this Guide, the definition of the general metrological term “uncertainty of measurement” is given both in Annex B and 2.2.3. The definitions of the most important terms specific to this Guide are given in 2.3.1 to 2.3.6. In all of these subclauses and in Annexes B and C, the use of parentheses around certain words of some terms means that these words may be omitted if this is unlikely to cause confusion.

2.2 The term “uncertainty”

The concept of uncertainty is discussed further in Clause 3 and Annex D.

2.2.1 The word “uncertainty” means doubt, and thus in its broadest sense “uncertainty of measurement” means doubt about the validity of the result of a measurement. Because of the lack of different words for this general concept of uncertainty and the specific quantities that provide quantitative measures of the concept, for example, the standard deviation, it is necessary to use the word “uncertainty” in these two different senses.

2.2.2 In this Guide, the word “uncertainty” without adjectives refers both to the general concept of uncertainty and to any or all quantitative measures of that concept. When a specific measure is intended, appropriate adjectives are used.

2.2.3 The formal definition of the term “uncertainty of measurement” developed for use in this Guide and in the VIM [6] (VIM:1993, definition 3.9) is as follows:

uncertainty (of measurement)
parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand

NOTE 1 The parameter may be, for example, a standard deviation (or a given multiple of it), or the half-width of an interval having a stated level of confidence.

NOTE 2 Uncertainty of measurement comprises, in general, many components. Some of these components may be evaluated from the statistical distribution of the results of series of measurements and can be characterized by experimental standard deviations. The other components, which also can be characterized by standard deviations, are evaluated from assumed probability distributions based on experience or other information.

NOTE 3 It is understood that the result of the measurement is the best estimate of the value of the measurand, and that all components of uncertainty, including those arising from systematic effects, such as components associated with corrections and reference standards, contribute to the dispersion.

2.2.4 The definition of uncertainty of measurement given in 2.2.3 is an operational one that focuses on the measurement result and its evaluated uncertainty. However, it is not inconsistent with other concepts of uncertainty of measurement, such as

* Footnote to the 2008 version:
— a measure of the possible error in the estimated value of the measurand as provided by the result of a measurement;

— an estimate characterizing the range of values within which the true value of a measurand lies (VIM:1984, definition 3.09).

Although these two traditional concepts are valid as ideals, they focus on unknowable quantities: the “error” of the result of a measurement and the “true value” of the measurand (in contrast to its estimated value), respectively. Nevertheless, whichever concept of uncertainty is adopted, an uncertainty component is always evaluated using the same data and related information. (See also E.5.)

2.3 Terms specific to this Guide

In general, terms that are specific to this Guide are defined in the text when first introduced. However, the definitions of the most important of these terms are given here for easy reference.

NOTE Further discussion related to these terms may be found as follows: for 2.3.2, see 3.3.3 and 4.2; for 2.3.3, see 3.3.3 and 4.3; for 2.3.4, see Clause 5 and Equations (10) and (13); and for 2.3.5 and 2.3.6, see Clause 6.

2.3.1 standard uncertainty
uncertainty of the result of a measurement expressed as a standard deviation

2.3.2 Type A evaluation (of uncertainty)
method of evaluation of uncertainty by the statistical analysis of series of observations

2.3.3 Type B evaluation (of uncertainty)
method of evaluation of uncertainty by means other than the statistical analysis of series of observations

2.3.4 combined standard uncertainty
standard uncertainty of the result of a measurement when that result is obtained from the values of a number of other quantities, equal to the positive square root of a sum of terms, the terms being the variances or covariances of these other quantities weighted according to how the measurement result varies with changes in these quantities

2.3.5 expanded uncertainty
quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand

NOTE 1 The fraction may be viewed as the coverage probability or level of confidence of the interval.

NOTE 2 To associate a specific level of confidence with the interval defined by the expanded uncertainty requires explicit or implicit assumptions regarding the probability distribution characterized by the measurement result and its combined standard uncertainty. The level of confidence that may be attributed to this interval can be known only to the extent to which such assumptions may be justified.

NOTE 3 Expanded uncertainty is termed overall uncertainty in paragraph 5 of Recommendation INC-1 (1980).

2.3.6 coverage factor
numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty

NOTE A coverage factor, \( k \), is typically in the range 2 to 3.
3 Basic concepts

Additional discussion of basic concepts may be found in Annex D, which focuses on the ideas of “true” value, error and uncertainty and includes graphical illustrations of these concepts; and in Annex E, which explores the motivation and statistical basis for Recommendation INC-1 (1980) upon which this Guide rests. Annex J is a glossary of the principal mathematical symbols used throughout the Guide.

3.1 Measurement

3.1.1 The objective of a measurement (B.2.5) is to determine the value (B.2.2) of the measurand (B.2.9), that is, the value of the particular quantity (B.2.1, Note 1) to be measured. A measurement therefore begins with an appropriate specification of the measurand, the method of measurement (B.2.7), and the measurement procedure (B.2.8).

NOTE The term “true value” (see Annex D) is not used in this Guide for the reasons given in D.3.5; the terms “value of a measurand” (or of a quantity) and “true value of a measurand” (or of a quantity) are viewed as equivalent.

3.1.2 In general, the result of a measurement (B.2.11) is only an approximation or estimate (C.2.26) of the value of the measurand and thus is complete only when accompanied by a statement of the uncertainty (B.2.18) of that estimate.

3.1.3 In practice, the required specification or definition of the measurand is dictated by the required accuracy of measurement (B.2.14). The measurand should be defined with sufficient completeness with respect to the required accuracy so that for all practical purposes associated with the measurement its value is unique. It is in this sense that the expression “value of the measurand” is used in this Guide.

EXAMPLE If the length of a nominally one-metre long steel bar is to be determined to micrometre accuracy, its specification should include the temperature and pressure at which the length is defined. Thus the measurand should be specified as, for example, the length of the bar at 25,00 °C * and 101 325 Pa (plus any other defining parameters deemed necessary, such as the way the bar is to be supported). However, if the length is to be determined to only millimetre accuracy, its specification would not require a defining temperature or pressure or a value for any other defining parameter.

NOTE Incomplete definition of the measurand can give rise to a component of uncertainty sufficiently large that it must be included in the evaluation of the uncertainty of the measurement result (see D.1.1, D.3.4, and D.6.2).

3.1.4 In many cases, the result of a measurement is determined on the basis of series of observations obtained under repeatability conditions (B.2.15, Note 1).

3.1.5 Variations in repeated observations are assumed to arise because influence quantities (B.2.10) that can affect the measurement result are not held completely constant.

3.1.6 The mathematical model of the measurement that transforms the set of repeated observations into the measurement result is of critical importance because, in addition to the observations, it generally includes various influence quantities that are inexactly known. This lack of knowledge contributes to the uncertainty of the measurement result, as do the variations of the repeated observations and any uncertainty associated with the mathematical model itself.

3.1.7 This Guide treats the measurand as a scalar (a single quantity). Extension to a set of related measurands determined simultaneously in the same measurement requires replacing the scalar measurand and its variance (C.2.11, C.2.20, C.3.2) by a vector measurand and covariance matrix (C.3.5). Such a replacement is considered in this Guide only in the examples (see H.2, H.3, and H.4).

* Footnote to the 2008 version:

According to Resolution 10 of the 22nd CGPM (2003) “... the symbol for the decimal marker shall be either the point on the line or the comma on the line...” The JCGM has decided to adopt, in its documents in English, the point on the line. However, in this document, the decimal comma has been retained for consistency with the 1995 printed version.
3.2 Errors, effects, and corrections

3.2.1 In general, a measurement has imperfections that give rise to an error (B.2.19) in the measurement result. Traditionally, an error is viewed as having two components, namely, a random (B.2.21) component and a systematic (B.2.22) component.

NOTE Error is an idealized concept and errors cannot be known exactly.

3.2.2 Random error presumably arises from unpredictable or stochastic temporal and spatial variations of influence quantities. The effects of such variations, hereafter termed random effects, give rise to variations in repeated observations of the measurand. Although it is not possible to compensate for the random error of a measurement result, it can usually be reduced by increasing the number of observations; its expectation or expected value (C.2.9, C.3.1) is zero.

NOTE 1 The experimental standard deviation of the arithmetic mean or average of a series of observations (see 4.2.3) is not the random error of the mean, although it is so designated in some publications. It is instead a measure of the uncertainty of the mean due to random effects. The exact value of the error in the mean arising from these effects cannot be known.

NOTE 2 In this Guide, great care is taken to distinguish between the terms "error" and "uncertainty". They are not synonyms, but represent completely different concepts; they should not be confused with one another or misused.

3.2.3 Systematic error, like random error, cannot be eliminated but it too can often be reduced. If a systematic error arises from a recognized effect of an influence quantity on a measurement result, hereafter termed a systematic effect, the effect can be quantified and, if it is significant in size relative to the required accuracy of the measurement, a correction (B.2.23) or correction factor (B.2.24) can be applied to compensate for the effect. It is assumed that, after correction, the expectation or expected value of the error arising from a systematic effect is zero.

NOTE The uncertainty of a correction applied to a measurement result to compensate for a systematic effect is not the systematic error, often termed bias, in the measurement result due to the effect as it is sometimes called. It is instead a measure of the uncertainty of the result due to incomplete knowledge of the required value of the correction. The error arising from imperfect compensation of a systematic effect cannot be exactly known. The terms "error" and "uncertainty" should be used properly and care taken to distinguish between them.

3.2.4 It is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects.

EXAMPLE A correction due to the finite impedance of a voltmeter used to determine the potential difference (the measurand) across a high-impedance resistor is applied to reduce the systematic effect on the result of the measurement arising from the loading effect of the voltmeter. However, the values of the impedances of the voltmeter and resistor, which are used to estimate the value of the correction and which are obtained from other measurements, are themselves uncertain. These uncertainties are used to evaluate the component of the uncertainty of the potential difference determination arising from the correction and thus from the systematic effect due to the finite impedance of the voltmeter.

NOTE 1 Often, measuring instruments and systems are adjusted or calibrated using measurement standards and reference materials to eliminate systematic effects; however, the uncertainties associated with these standards and materials must still be taken into account.

NOTE 2 The case where a correction for a known significant systematic effect is not applied is discussed in the Note to 6.3.1 and in F.2.4.5.

3.3 Uncertainty

3.3.1 The uncertainty of the result of a measurement reflects the lack of exact knowledge of the value of the measurand (see 2.2). 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.

3.3.2 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) nonrepresentative sampling — the sample measured may not represent the defined measurand;

d) inadequate knowledge of the effects of environmental conditions on the measurement or imperfect measurement of environmental conditions;

e) personal bias in reading analogue instruments;

f) finite instrument resolution or discrimination threshold;

g) inexact values of measurement standards and reference materials;

h) inexact values of constants and other parameters obtained from external sources and used in the data-reduction algorithm;

i) approximations and assumptions incorporated in the measurement method and procedure;

j) 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.

3.3.3 Recommendation INC-1 (1980) of the Working Group on the Statement of Uncertainties groups uncertainty components into two categories based on their method of evaluation, “A” and “B” (see 0.7, 2.3.2, and 2.3.3). These categories apply to uncertainty and are not substitutes for the words “random” and “systematic”. The uncertainty of a correction for a known systematic effect may in some cases be obtained by a Type A evaluation while in other cases by a Type B evaluation, as may the uncertainty characterizing a random effect.

NOTE In some publications, uncertainty components are categorized as “random” and “systematic” and are associated with errors arising from random effects and known systematic effects, respectively. Such categorization of components of uncertainty can be ambiguous when generally applied. For example, a “random” component of uncertainty in one measurement may become a “systematic” component of uncertainty in another measurement in which the result of the first measurement is used as an input datum. Categorizing the methods of evaluating uncertainty components rather than the components themselves avoids such ambiguity. At the same time, it does not preclude collecting individual components that have been evaluated by the two different methods into designated groups to be used for a particular purpose (see 3.4.3).

3.3.4 The purpose of the Type A and Type B classification is to indicate the two different ways of evaluating uncertainty components and is for convenience of discussion only; the classification is not meant to indicate that there is any difference in the nature of the components resulting from the two types of evaluation. Both types of evaluation are based on probability distributions (C.2.3), and the uncertainty components resulting from either type are quantified by variances or standard deviations.

3.3.5 The estimated variance $u^2$ characterizing an uncertainty component obtained from a Type A evaluation is calculated from series of repeated observations and is the familiar statistically estimated variance $s^2$ (see 4.2). The estimated standard deviation (C.2.12, C.2.21, C.3.3) $u$, the positive square root of $u^2$, is thus $u = s$ and for convenience is sometimes called a Type A standard uncertainty. For an uncertainty component obtained from a Type B evaluation, the estimated variance $u^2$ is evaluated using available
knowledge (see 4.3), and the estimated standard deviation \( \sigma \) is sometimes called a *Type B standard uncertainty*.

Thus a Type A standard uncertainty is obtained from a *probability density function* (C.2.5) derived from an *observed frequency distribution* (C.2.18), while a Type B standard uncertainty is obtained from an assumed probability density function based on the degree of belief that an event will occur [often called *subjective probability* (C.2.1)]. Both approaches employ recognized interpretations of probability.

**NOTE** A Type B evaluation of an uncertainty component is usually based on a pool of comparatively reliable information (see 4.3.1).

### 3.3.6 The standard uncertainty of the result of a measurement, when that result is obtained from the values of a number of other quantities, is termed *combined standard uncertainty* and denoted by \( \sigma_c \). It is the estimated standard deviation associated with the result and is equal to the positive square root of the combined variance obtained from all variance and *covariance* (C.3.4) components, however evaluated, using what is termed in this *Guide* the *law of propagation of uncertainty* (see Clause 5).

### 3.3.7 To meet the needs of some industrial and commercial applications, as well as requirements in the areas of health and safety, an *expanded uncertainty* \( U \) is obtained by multiplying the combined standard uncertainty \( \sigma_c \) by a *coverage factor* \( k \). The intended purpose of \( U \) is to provide an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand. The choice of the factor \( k \), which is usually in the range 2 to 3, is based on the coverage probability or level of confidence required of the interval (see Clause 6).

**NOTE** The coverage factor \( k \) is always to be stated, so that the standard uncertainty of the measured quantity can be recovered for use in calculating the combined standard uncertainty of other measurement results that may depend on that quantity.

### 3.4 Practical considerations

#### 3.4.1 If all of the quantities on which the result of a measurement depends are varied, its uncertainty can be evaluated by statistical means. However, because this is rarely possible in practice due to limited time and resources, the uncertainty of a measurement result is usually evaluated using a mathematical model of the measurement and the law of propagation of uncertainty. Thus implicit in this *Guide* is the assumption that a measurement can be modelled mathematically to the degree imposed by the required accuracy of the measurement.

#### 3.4.2 Because the mathematical model may be incomplete, all relevant quantities should be varied to the fullest practicable extent so that the evaluation of uncertainty can be based as much as possible on observed data. Whenever feasible, the use of empirical models of the measurement founded on long-term quantitative data, and the use of check standards and control charts that can indicate if a measurement is under statistical control, should be part of the effort to obtain reliable evaluations of uncertainty. The mathematical model should always be revised when the observed data, including the result of independent determinations of the same measurand, demonstrate that the model is incomplete. A well-designed experiment can greatly facilitate reliable evaluations of uncertainty and is an important part of the art of measurement.

#### 3.4.3 In order to decide if a measurement system is functioning properly, the experimentally observed variability of its output values, as measured by their observed standard deviation, is often compared with the predicted standard deviation obtained by combining the various uncertainty components that characterize the measurement. In such cases, only those components (whether obtained from Type A or Type B evaluations) that could contribute to the experimentally observed variability of these output values should be considered.

**NOTE** Such an analysis may be facilitated by gathering those components that contribute to the variability and those that do not into two separate and appropriately labelled groups.

#### 3.4.4 In some cases, the uncertainty of a correction for a systematic effect need not be included in the evaluation of the uncertainty of a measurement result. Although the uncertainty has been evaluated, it may be ignored if its contribution to the combined standard uncertainty of the measurement result is insignificant. If the value of the correction itself is insignificant relative to the combined standard uncertainty, it too may be ignored.
3.4.5 It often occurs in practice, especially in the domain of legal metrology, that a device is tested through a comparison with a measurement standard and the uncertainties associated with the standard and the comparison procedure are negligible relative to the required accuracy of the test. An example is the use of a set of well-calibrated standards of mass to test the accuracy of a commercial scale. In such cases, because the components of uncertainty are small enough to be ignored, the measurement may be viewed as determining the error of the device under test. (See also F.2.4.2.)

3.4.6 The estimate of the value of a measurand provided by the result of a measurement is sometimes expressed in terms of the adopted value of a measurement standard rather than in terms of the relevant unit of the International System of Units (SI). In such cases, the magnitude of the uncertainty ascribable to the measurement result may be significantly smaller than when that result is expressed in the relevant SI unit. (In effect, the measurand has been redefined to be the ratio of the value of the quantity to be measured to the adopted value of the standard.)

EXAMPLE A high-quality Zener voltage standard is calibrated by comparison with a Josephson effect voltage reference based on the conventional value of the Josephson constant recommended for international use by the CIPM. The relative combined standard uncertainty $u_c(V_S)/V_S$ (see 5.1.6) of the calibrated potential difference $V_S$ of the Zener standard is $2 \times 10^{-8}$ when $V_S$ is reported in terms of the conventional value, but $u_c(V_S)/V_S$ is $4 \times 10^{-7}$ when $V_S$ is reported in terms of the SI unit of potential difference, the volt (V), because of the additional uncertainty associated with the SI value of the Josephson constant.

3.4.7 Blunders in recording or analysing data can introduce a significant unknown error in the result of a measurement. Large blunders can usually be identified by a proper review of the data; small ones could be masked by, or even appear as, random variations. Measures of uncertainty are not intended to account for such mistakes.

3.4.8 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 integrity of those who contribute to the assignment of its value.

4 Evaluating standard uncertainty

Additional guidance on evaluating uncertainty components, mainly of a practical nature, may be found in Annex F.

4.1 Modelling the measurement

4.1.1 In most cases, a measurand $Y$ is not measured directly, but is determined from $N$ other quantities $X_1, X_2, \ldots, X_N$ through a functional relationship $f$: 

$$Y = f(X_1, X_2, \ldots, X_N)$$  \hspace*{1cm} (1)

NOTE 1 For economy of notation, in this Guide the same symbol is used for the physical quantity (the measurand) and for the random variable (see 4.2.1) that represents the possible outcome of an observation of that quantity. When it is stated that $X_i$ has a particular probability distribution, the symbol is used in the latter sense; it is assumed that the physical quantity itself can be characterized by an essentially unique value (see 1.2 and 3.1.3).

NOTE 2 In a series of observations, the $k$th observed value of $X_i$ is denoted by $X_{ik}$; hence if $R$ denotes the resistance of a resistor, the $k$th observed value of the resistance is denoted by $R_k$.

NOTE 3 The estimate of $X_i$ (strictly speaking, of its expectation) is denoted by $x_i$. 
EXAMPLE If a potential difference \( V \) is applied to the terminals of a temperature-dependent resistor that has a resistance \( R_0 \) at the defined temperature \( t_0 \) and a linear temperature coefficient of resistance \( \alpha \), the power \( P \) (the measurand) dissipated by the resistor at the temperature \( t \) depends on \( V, R_0, \alpha, \) and \( t \) according to

\[
P = f(V, R_0, \alpha, t) = V^2 [R_0 \left(1 + \alpha(t - t_0)\right)]
\]

NOTE Other methods of measuring \( P \) would be modelled by different mathematical expressions.

4.1.2 The input quantities \( X_1, X_2, \ldots, X_N \) upon which the output quantity \( Y \) depends may themselves be viewed as measurands and may themselves depend on other quantities, including corrections and correction factors for systematic effects, thereby leading to a complicated functional relationship \( f \) that may never be written down explicitly. Further, \( f \) may be determined experimentally (see 5.1.4) or exist only as an algorithm that must be evaluated numerically. The function \( f \) as it appears in this Guide is to be interpreted in this broader context, in particular as that function which contains every quantity, including all corrections and correction factors, that can contribute a significant component of uncertainty to the measurement result.

Thus, if data indicate that \( f \) does not model the measurement to the degree imposed by the required accuracy of the measurement result, additional input quantities must be included in \( f \) to eliminate the inadequacy (see 3.4.2). This may require introducing an input quantity to reflect incomplete knowledge of a phenomenon that affects the measurand. In the example of 4.1.1, additional input quantities might be needed to account for a known nonuniform temperature distribution across the resistor, a possible nonlinear temperature coefficient of resistance, or a possible dependence of resistance on barometric pressure.

NOTE Nonetheless, Equation (1) may be as elementary as \( Y = X_1 - X_2 \). This expression models, for example, the comparison of two determinations of the same quantity \( X \).

4.1.3 The set of input quantities \( X_1, X_2, \ldots, X_N \) may be categorized as:

- quantities whose values and uncertainties are directly determined in the current measurement. These values and uncertainties may be obtained from, for example, a single observation, repeated observations, or judgement based on experience, and may involve the determination of corrections to instrument readings and corrections for influence quantities, such as ambient temperature, barometric pressure, and humidity;

- quantities whose values and uncertainties are brought into the measurement from external sources, such as quantities associated with calibrated measurement standards, certified reference materials, and reference data obtained from handbooks.

4.1.4 An estimate of the measurand \( Y \), denoted by \( y \), is obtained from Equation (1) using input estimates \( x_1, x_2, \ldots, x_N \) for the values of the \( N \) quantities \( X_1, X_2, \ldots, X_N \). Thus the output estimate \( y \), which is the result of the measurement, is given by

\[
y = f(x_1, x_2, \ldots, x_N)
\]

NOTE In some cases, the estimate \( y \) may be obtained from

\[
y = \overline{Y} = \frac{1}{n} \sum_{k=1}^{n} y_k = \frac{1}{n} \sum_{k=1}^{n} f(X_{1k}, X_{2k}, \ldots, X_{Nk})
\]

That is, \( y \) is taken as the arithmetic mean or average (see 4.2.1) of \( n \) independent determinations \( Y_k \) of \( Y \), each determination having the same uncertainty and each being based on a complete set of observed values of the \( N \) input quantities \( X_i \) obtained at the same time. This way of averaging, rather than \( y = f(\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_N) \), where

\[
\overline{X}_j = \frac{1}{n} \sum_{k=1}^{n} X_{jk}
\]

is the arithmetic mean of the individual observations \( X_{ij} \), may be preferable when \( f \) is a nonlinear function of the input quantities \( X_1, X_2, \ldots, X_N \), but the two approaches are identical if \( f \) is a linear function of the \( X_i \) (see H.2 and H.4).

4.1.5 The estimated standard deviation associated with the output estimate or measurement result \( y \), termed combined standard uncertainty and denoted by \( u_c(y) \), is determined from the estimated standard deviation associated with each input estimate \( x_i \), termed standard uncertainty and denoted by \( u(x_i) \) (see 3.3.5 and 3.3.6).