

GUIDE

**OIML G 19**

Edition 2017 (E)

---

The role of measurement uncertainty in conformity  
assessment decisions in legal metrology

Le rôle de l'incertitude de mesure dans les décisions pour l'évaluation  
de la conformité en métrologie légale

---





## Contents

Foreword .....	5
1 Scope and objectives .....	6
2 Terms and definitions .....	8
3 Introduction .....	16
3.1 Characterization of measuring instruments and systems .....	16
3.2 Conformity decisions using measurement uncertainty .....	16
3.3 Error versus uncertainty .....	17
3.4 Verification incorporating measurement uncertainty .....	17
3.5 MPEs and measurement uncertainty .....	17
4 Basic considerations pertaining to conformity testing decisions and measurement uncertainty .....	18
5 Conformity testing decisions that explicitly incorporate measurement uncertainty .....	20
5.1 Probability density function (PDF) .....	22
5.2 Probability of conformity .....	24
5.3 “Risks” and “decision rules” associated with conformity decisions .....	26
6 Conformity testing decisions that do not explicitly incorporate measurement uncertainty .....	32
7 Taking measurement uncertainty into account when establishing maximum permissible errors (MPEs) and accuracy classes .....	35
8 Options and suggested text pertaining to “measurement uncertainty” that should be considered for inclusion in OIML Recommendations and other OIML publications .....	36
8.1 Incorporating measurement uncertainty for laboratory testing .....	36
8.2 Calculating measurement uncertainty .....	37
8.3 Specifying MPEs and MPUs .....	38
8.4 Specifying acceptable levels of risk .....	38
8.5 Specifying uncertainty of error of indication when not using shared risk .....	39
8.6 Complexity of assessing uncertainty of error of indication .....	39
8.7 Recording measurement uncertainty in OIML test reports for type evaluation .....	40
8.8 Providing guidance on using measurement uncertainty for periodic verification .....	40
9 References .....	41
Annex A Coexistence of “measurement error” and “measurement uncertainty” in legal metrology (relationship between calibration and verification) .....	44
Annex B Use of the standard normal distribution table (Z-table) .....	52
Annex C Example of assessing measurement uncertainty of error of indication .....	57
Annex D Example of risk assessment incorporating measurement uncertainty .....	64
Annex E Measurement Capability Index ( $C_m$ ) .....	70
Annex F Establishing measurement uncertainty to use with conformity tested measuring instruments and systems .....	72



## Foreword

The International Organization of Legal Metrology (OIML) is a worldwide, intergovernmental organization whose primary aim is to harmonize the regulations and metrological controls applied by the national metrological services, or related organizations, of its Member States.

The main categories of OIML publications are:

- **International Recommendations (OIML R)**, which are model regulations that establish the metrological characteristics required of certain measuring instruments and which specify methods and equipment for checking their conformity. OIML Member States shall implement these Recommendations to the greatest possible extent;
- **International Documents (OIML D)**, which are informative in nature and which are intended to harmonize and improve work in the field of legal metrology;
- **International Guides (OIML G)**, which are also informative in nature and which are intended to give guidelines for the application of certain requirements to legal metrology; and
- **International Basic Publications (OIML B)**, which define the operating rules of the various OIML structures and systems.

OIML Draft Recommendations, Documents and Guides are developed by Project Groups linked to Technical Committees or Subcommittees which comprise representatives from the Member States. Certain international and regional institutions also participate on a consultation basis. Cooperative agreements have been established between the OIML and certain institutions, such as ISO and the IEC, with the objective of avoiding contradictory requirements. Consequently, manufacturers and users of measuring instruments, test laboratories, etc. may simultaneously apply OIML publications and those of other institutions.

International Recommendations, Documents, Guides and Basic Publications are published in English (E) and translated into French (F) and are subject to periodic revision.

Additionally, the OIML publishes or participates in the publication of **Vocabularies (OIML V)** and periodically commissions legal metrology experts to write **Expert Reports (OIML E)**. Expert Reports are intended to provide information and advice, and are written solely from the viewpoint of their author, without the involvement of a Technical Committee or Subcommittee, nor that of the CIML. Thus, they do not necessarily represent the views of the OIML.

This publication – reference OIML G 19, edition 2017 (E) – was developed by Project Group 2 of OIML Technical Subcommittee TC 3/SC 5 *New Guide: Expression of uncertainty in measurement in legal metrology applications*. It was approved for final publication by the President of the International Committee of Legal Metrology (CIML) in February 2017.

OIML Publications may be downloaded from the OIML website in the form of PDF files. Additional information on OIML Publications may be obtained from the Organization's headquarters:

Bureau International de Métrologie Légale  
11, rue Turgot - 75009 Paris – France  
Telephone: 33 (0)1 48 78 12 82  
Fax: 33 (0)1 42 82 17 27  
E-mail: [biml@oiml.org](mailto:biml@oiml.org)  
Internet: [www.oiml.org](http://www.oiml.org)

## 1 Scope and objectives

The scope of this OIML Guide is to provide guidance to OIML Secretariats and Conveners and to members of OIML Technical Committees, Subcommittees and Project Groups, on incorporating the concept of “measurement uncertainty” into OIML Recommendations and other OIML publications used for legal metrology purposes. It is assumed that the reader has at least a general familiarity with the concepts presented in the *Guide to the Expression of Uncertainty in Measurement* [1] (hereafter denoted by GUM), and possibly also with the concepts in its Supplements [2][3][4][5]. It is anticipated that this Guide will eventually become an OIML Document; as such, its contents will need to be incorporated, as appropriate, in OIML Recommendations and Documents. However, this initial version is presented as an OIML Guide in order to give OIML Technical Committees, Subcommittees and Project Groups additional time to consider the contents and how they can be incorporated into the Recommendations and Documents for which they are responsible.

The main objective of this OIML Guide is to provide guidance on incorporating text into OIML publications that describes when and how to take measurement uncertainty into account in conformity assessment by inspection, that is, when determining whether an entity (product, process, system, person or body) meets relevant standards or fulfils specified requirements. A particular focus is on conformity assessment of measuring instruments (or systems), especially when using measured values, obtained during the testing or verification of the measuring instruments or systems, as the basis for making pass-fail decisions in legal metrology. The role of measurement uncertainty for other important entities subject to conformity assessment in legal metrology, such as prepackages (as a prototype for more general product conformity assessment), and more generally for entities of all kinds, is dealt with elsewhere.

Practical procedures to incorporate this OIML Guide into other OIML publications are proposed in clause 8.

The proposals include providing and referencing information on how to assess the possible “risks” of erroneous conformity decisions. Such risks arise unavoidably from the measurement uncertainty associated with the measured values obtained during testing or verification of a measuring instrument or system. That is, measurement uncertainty in a test result – an apparent product dispersion arising from limited measurement quality – can be a concern in conformity assessment by inspection since if it is not accounted for it can lead to incorrect estimates of the consequences of entity error and increase the risk of making incorrect decisions, such as failing a conforming entity or passing a non-conforming entity when the test result is close to a tolerance limit.

This Guide also elaborates on the difference between “error” and “uncertainty” in a way that demonstrates how both concepts (and terms) are important in legal metrology. This Guide also provides guidelines and examples for the determination and expression of measurement uncertainty in legal metrology applications, consistent with the GUM and its Supplements.

The guidance provided in this Guide is intended to be applicable for both the type evaluation and verification of measuring instruments used in legal metrology. However, recognizing that in many cases the determination of measurement uncertainty can be a difficult, time-consuming and therefore expensive activity, guidance is also given on how the explicit determination of measurement uncertainty can be acceptably simplified or even avoided in certain measurement scenarios, such as verification.

Another important objective of this Guide is to demonstrate how measurement uncertainty can be taken into account, at least implicitly, for measuring instruments and systems that have been verified. This is important since uncertainty assessment is critical so that measurement results (values and uncertainties) that are obtained when subsequently using the verified instrument/system can have metrological traceability.

Harmonized methods for evaluating measurement uncertainties and implementing them into decision criteria used for the metrological evaluation of measuring instruments and systems are necessary so that test evaluations and metrological judgments may yield comparable results from one national responsible body in legal metrology to another. Such comparability is an important element for achieving trust between bodies in recognizing each other's type approvals, leading to the intended operation and function of the OIML Certification System (OIML-CS) [6], which will replace the OIML Basic Certificate System [7] and the OIML Mutual Acceptance Arrangement (MAA) [8]. Such trust is generally also necessary for providing confidence in verification processes and certificates.

The guidance provided in this Guide is intended to be consistent with ISO/IEC 17025 [9] *General requirements for the competence of testing and calibration laboratories* with regard to requirements involving the use of measurement uncertainty. Note, however, that the scope of this Guide is not intended to cover when a country must require accreditation of its own calibration and testing laboratories to ISO/IEC 17025, or even when and how a country shall specify the required use of measurement uncertainty in its national legislation. Testing laboratories used by Issuing Authorities in the OIML-CS must follow the requirements in OIML D 30 [10] *Guide for the application of ISO/IEC 17025 to the assessment of Testing Laboratories involved in legal metrology*.

As mentioned above, other topics related to risk assessment not covered in the scope of this Guide are

- sampling by attributes (e.g. broken seals, labeling, etc.),
- populations of instruments in a 'statistical analysis' sense, and
- net content and labeling of prepackages (see OIML R 87 [11] and OIML R 79 [12]).

## 2 Terms and definitions

Terms and definitions in clause 2 are taken from the following references: VIM3 [14], VIML [15] and JCGM 106 [5]. In general, Examples and Notes have not been included here, and the original reference should be consulted if necessary. In some cases, Notes have been included here when it is felt that they are important to the understanding of the definition.

### 2.1

#### **quantity** (VIM3 1.1)

property of a phenomenon, body or substance, where the property has a magnitude that can be expressed as a number and a reference

### 2.2

#### **quantity value** (VIM3 1.19)

number and reference together expressing magnitude of a quantity

### 2.3

#### **true quantity value** (VIM3 2.11)

quantity value consistent with the definition of a quantity

### 2.4

#### **measurand** (VIM3 2.3)

quantity intended to be measured

### 2.5

#### **measurement model** (VIM3 2.48)

mathematical relation among all quantities known to be involved in a measurement

### 2.6

#### **measurement function** (VIM3 2.49)

function of quantities, the value of which, when calculated using known quantity values for the input quantities in a measurement model, is a measured quantity value of the output quantity in the measurement model

### 2.7

#### **measured quantity value** (VIM3 2.10)

quantity value representing a measurement result

### 2.8

#### **measurement uncertainty** (VIM3 2.26)

non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used

*Note* (not in VIM3): In GUM Supplement JCGM 104 [4], measurement uncertainty is described as a measure of how well the essentially unique true value of a measurand is believed to be known.

## 2.9

**measurement result** (VIM3 2.9)

set of quantity values being attributed to a measurand together with any other available relevant information

## 2.10

**measurement error** (VIM3 2.16)

measured quantity value minus a reference quantity value

*Note 1:* The concept of ‘measurement error’ can be used both

- a) when there is a single reference quantity value to refer to, which occurs if a calibration is made by means of a measurement standard with a measured quantity value having a negligible measurement uncertainty or if a conventional quantity value is given, in which case the measurement error is known, and
- b) if a measurand is supposed to be represented by a unique true quantity value or a set of true quantity values of negligible range, in which case the measurement error is not known.

*Note 2:* Measurement error should not be confused with production error or mistake.

*Note 3* (not in VIM3): There are two positions on the consideration of ‘error’, whether it should be defined as a ‘value’, as in the above definition, or as a ‘quantity’ that has a value. Both uses of the term ‘error’ can be found in the metrology literature. In this Guide the definition given above will be used. Note that in reference [5] this is not the case.

## 2.11

**measurement bias** (VIM3 2.18)

estimate of a systematic error

## 2.12

**indication** (VIM3 4.1)

quantity value provided by a measuring instrument or a measuring system

*Note 1:* An indication may be presented in visual or acoustic form or may be transferred to another device. An indication is often given by the position of a pointer on the display for analog outputs, a displayed or printed number for digital outputs, a code pattern for code outputs, or an assigned quantity value for material measures.

*Note 2:* An indication and a corresponding value of the quantity being measured are not necessarily values of quantities of the same kind.

*Note 3* (not in VIM3): There are two positions on the consideration of ‘indication’, whether it should be defined as a ‘value’, as in the above definition, or as a ‘quantity’ that has a value. Both uses of the term ‘indication’ can be found in the metrology literature. In this Guide the definition given above will be used. Note that in reference [5] this is not the case.

2.13

**error of indication** (VIML 0.04)

indication minus a reference quantity value

*Note:* This reference value is sometimes referred to as a (conventional) true quantity value. See, however, also OIML V 2-200:2012, 2.12, Note 1.

2.14

**maximum permissible measurement error (MPE)** (VIM3 4.26)

extreme value of measurement error, with respect to a known reference quantity value, permitted by specifications or regulations for a given measurement, measuring instrument, or measuring system

*Note 1:* Usually, the term “maximum permissible errors” or “limits of error” is used where there are two extreme values.

*Note 2:* The term “tolerance” should not be used to designate ‘maximum permissible error’.

*Note 3* (not in VIM3): There are two positions on the consideration of ‘maximum permissible error’, whether it should be defined as a ‘value’, as in the above definition, or as a ‘quantity’ that has a value. In this Guide the definition given above will be used. Note that in reference [5] this is not the case.

2.15

**maximum permissible uncertainty (MPU<sub>E1</sub>)**

largest value that the measurement uncertainty of the error of indication for a test result can have for which the shared risk approach can be used

2.16

**metrological traceability** (VIM3 2.41)

property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty

2.17

**measurement capability index (C<sub>m</sub>)** (JCGM 106 3.3.17)

tolerance divided by a multiple of the standard measurement uncertainty associated with the measured value of a property of an item

*Note 1:*  $C_m = MPE/U_{E1}$

2.18

**risk of false acceptance** (JCGM 106 3.3.13 called **specific consumer’s risk**)

probability that a particular accepted item is non-conforming

2.19

**risk of false rejection** (JCGM 106 3.3.16 called **specific producer’s risk**)

probability that a particular rejected item is conforming

2.20

**shared risk**

risk which is based on an agreement between parties concerned with the outcome of a test that neither party will be given an advantage or suffer a disadvantage concerning consideration of measurement uncertainty

2.21

**guard band** (JCGM 106 3.3.11)

interval between a tolerance limit and a corresponding acceptance limit

2.22

**measuring system** (VIM3 3.2)

set of one or more measuring instruments and often other devices, including any reagent and supply, assembled and adapted to give information used to generate measured quantity values within specified intervals for quantities of specified kinds

*Note:* A measuring system may consist of only one measuring instrument.

2.23

**reference operating condition** (VIM3 4.11)

operating condition prescribed for evaluating the performance of a measuring instrument or measuring system or for comparison of measurement results

2.24

**rated operating condition** (VIM3 4.9)

operating condition that must be fulfilled during measurement in order that a measuring instrument or measuring system perform as designed

2.25

**conformity assessment** (VIML A.1)

demonstration that specified requirements relating to a product, process, system, person or body are fulfilled

2.26

**type (pattern) evaluation** (VIML 2.04)

conformity assessment procedure on one or more specimens of an identified type (pattern) of measuring instruments which results in an evaluation report and/or an evaluation certificate

2.27

**verification** (VIM3 2.44)

provision of objective evidence that a given item fulfils specified requirements

2.28

**verification of a measuring instrument** (VIML 2.09)

conformity assessment procedure (other than type evaluation) which results in the affixing of a verification mark and/or issuing of a verification certificate

2.29

**calibration** (VIM3 2.39)

operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication

2.30

**inspection** (VIML A.11)

examination of a product design, product, process or installation and determination of its conformity with specific requirements or, on the basis of professional judgment, with general requirements

*Note:* Inspection of a process may include inspection of persons, facilities, technology and methodology.

2.31

**metrology** (VIM3 2.2)

science of **measurement** and its application

Metrology includes all theoretical and practical aspects of measurement, whatever the **measurement uncertainty** and field of application.

2.32

**legal metrology** (VIML 0.01)

practice and process of applying statutory and regulatory structure and enforcement to metrology.

*Note 1:* The scope of legal metrology may be different from country to country.

*Note 2:* Legal metrology includes

- setting up legal requirements,
- control/conformity assessment of regulated products and regulated activities,
- supervision of regulated products and of regulated activities, and
- providing the necessary infrastructure for the traceability of regulated measurements and measuring instruments to SI or national standards.

*Note 3:* There are also regulations outside the area of legal metrology pertaining to the accuracy and correctness of measurement methods.

2.33

**measurement standard** (VIM3 5.1)

etalon

realization of the definition of a given **quantity**, with stated **quantity value** and associated **measurement uncertainty**, used as a reference

Example 1 1 kg mass measurement standard with an associated **standard measurement uncertainty** of 3  $\mu\text{g}$ .

Example 2 100  $\Omega$  measurement standard resistor with an associated standard measurement uncertainty of 1  $\mu\Omega$ .

- 
- Example 3 Caesium frequency standard with a relative standard measurement uncertainty of  $2 \times 10^{-15}$ .
- Example 4 Standard buffer solution with a pH of 7.072 with an associated standard measurement uncertainty of 0.006.
- Example 5 Set of reference solutions of cortisol in human serum having a certified quantity value with measurement uncertainty for each solution.
- Example 6 **Reference material** providing quantity values with measurement uncertainties for the mass concentration of each of ten different proteins.
- Note 1 A “realization of the definition of a given quantity” can be provided by a measuring system, a material measure, or a reference material.
- Note 2 A measurement standard is frequently used as a reference in establishing **measured quantity values** and associated measurement uncertainties for other quantities of the same **kind**, thereby establishing **metrological traceability** through **calibration** of other measurement standards, **measuring instruments**, or measuring systems.
- Note 3 The term “realization” is used here in the most general meaning. It denotes three procedures of “realization”:
- the first one consists in the physical realization of the **measurement unit** from its definition and is realization *sensu stricto*;
  - the second, termed “reproduction”, consists not in realizing the measurement unit from its definition but in setting up a highly reproducible measurement standard based on a physical phenomenon, as it happens, e.g. in case of use of frequency-stabilized lasers to establish a measurement standard for the metre, of the Josephson effect for the volt or of the quantum Hall effect for the ohm;
  - the third procedure consists in adopting a material measure as a measurement standard. It occurs in the case of the measurement standard of 1 kg.
- Note 4 A standard measurement uncertainty associated with a measurement standard is always a component of the **combined standard measurement uncertainty** (see GUM:1995, 2.3.4) in a **measurement result** obtained using the measurement standard. Frequently, this component is small compared with other components of the combined standard measurement uncertainty.
- Note 5 Quantity value and measurement uncertainty must be determined at the time when the measurement standard is used.
- Note 6 Several quantities of the same kind or of different kinds may be realized in one device which is commonly also called a measurement standard.
- Note 7 The word “embodiment” is sometimes used in the English language instead of “realization”.
-

Note 8 In science and technology, the English word “standard” is used with at least two different meanings: as a specification, technical recommendation, or similar normative document (in French “norme”) and as a measurement standard (in French “étalon”). This Guide is concerned solely with the second meaning.

Note 9 The term “measurement standard” is sometimes used to denote other metrological tools, e.g. ‘software measurement standard’ (see ISO 5436-2).

#### 2.34

##### **measurement accuracy** (VIM3 2.13)

accuracy of measurement

accuracy

closeness of agreement between a measured quantity value and a true quantity value of a measurand

*Note 1:* The concept ‘measurement accuracy’ is not a quantity and is not given a numerical quantity value. A measurement is said to be more accurate when it offers a smaller measurement error.

*Note 2:* The term “measurement accuracy” should not be used for measurement trueness and the term “measurement precision” should not be used for ‘measurement accuracy’, which, however, is related to both these concepts.

*Note 3:* The term “measurement accuracy” is sometimes understood as meaning closeness of agreement between measured quantity values that are being attributed to the measurand.

#### 2.35

##### **tolerance limit** (JCGM 106 3.3.4)

specified upper or lower bound of permissible values of a property

#### 2.36

##### **tolerance** (JCGM 106 3.3.5)

difference between upper and lower tolerance limits

## 2.a Abbreviations and symbols

BIPM	International Bureau of Weights and Measures
$E_I$	Error of Indication
$f_{E_I}$	= 1/TUR
$f_S$	= 1/TAR
GUM	Guide to the Expression of Uncertainty in Measurement [1]
IEC	International Electrotechnical Commission
IFCC	International Federation of Clinical Chemistry
ILAC	International Laboratory Accreditation Cooperation
ISO	International Organization for Standardization
IUPAC	International Union of Pure and Applied Chemistry
IUPAP	International Union of Pure and Applied Physics
IUT	Instrument Under Test
JCGM	Joint Committee for Guides in Metrology
MAA	OIML Mutual Acceptance Arrangement
MPE	Maximum Permissible Error
MPU	Maximum Permissible Uncertainty
OIML	International Organization of Legal Metrology
$p_n$	Probability of Non-Conformance
$p_{fa}$	Probability (Risk) of False Acceptance
$p_{fr}$	Probability (Risk) of False Rejection
PDF	Probability Density Function
SC	OIML Technical Subcommittee
TAR	Test Accuracy Ratio
TC	OIML Technical Committee
TUR	Test Uncertainty Ratio
$u_{E_I}$	Standard Measurement Uncertainty of Error of Indication
$u_S$	Standard Measurement Uncertainty of Measurement Standard (or System)
$u_I$	Standard Measurement Uncertainty of Indication
$u_{rep}$	Standard Measurement Uncertainty associated with Repeatability
$u_{roc}$	Standard Measurement Uncertainty associated with Rated Operating Conditions
VIM	International Vocabulary of Metrology
VIML	International Vocabulary of Legal Metrology
Z-table	Standard Normal Distribution Table

### **3 Introduction**

The concept of “measurement uncertainty”, as presented in the GUM, has revolutionized modern metrology. Consideration of measurement uncertainty is widely recognized, in both the metrology and laboratory accreditation communities, as being essential in order to have metrological traceability of measurement results.

There is a growing literature that provides methods for calculating and using measurement uncertainty for a variety of applications, including decision-making in legal metrology testing and verification. Some of these methods are more complex and time-consuming than others. While explicit, detailed determination and use of measurement uncertainty is usually appropriate for a calibration or testing laboratory environment, many measurements performed during the course of legal metrology activities are not performed in a laboratory. Rather, they are performed outside of a laboratory environment intended to allow for ‘quick and easy’ pass-or-fail decisions, and so the methods for determining and using measurement uncertainty (sometimes only implicitly) can be important for the efficiency and practicality of these activities (see 6).

#### **3.1 Characterization of measuring instruments and systems**

Legal metrology is the practice and process of applying statutory and regulatory structure and enforcement to metrology [15]. An important activity in legal metrology is the characterization of measuring instruments and systems that are used for the public good in areas such as commodity exchange, public health and safety, and environmental protection. This includes the classification and evaluation of instrument and system designs (or types), as well as the calibration or verification of the performance of individual instruments and systems, both immediately after their manufacture, as well as after their installation and use outside of a laboratory environment.

Conformity assessment in the context of legal metrology means an assessment of the design and performance of measuring instruments and systems according to technical and metrological requirements and specifications that are given in legal metrology documents. While individual countries, and jurisdictions within countries, may have their own requirements, one of the key objectives of the OIML is to provide authoritative requirements in OIML publications, especially OIML Recommendations, that can be adopted and used worldwide, either in their entirety or at least as a basis for harmonized legal metrology requirements.

#### **3.2 Conformity decisions using measurement uncertainty**

Making conformity decisions in legal metrology becomes more complex when including measurement uncertainty since the concepts of probability and risk enter the consideration (see 5.2 and 5.3). In particular, it becomes necessary to think in terms of the degree of belief (or level of confidence, expressed as a probability) that the essentially unique true value (denoted hereafter as “true value”) of a presumably constant systematic error of indication (denoted hereafter as just “error of indication”) actually lies outside the specified limits of maximum permissible error (MPE), even if the measured value lies within the MPE limits, and vice versa (see Annex A). Various “decision rules” can be established for deciding whether or not a particular test is considered to “pass” based on the expressed probability, and associated “risks” for making incorrect decisions can be calculated (see 5.2, Annex D). However, using techniques that will be described in this Guide (see 6) the need to

explicitly calculate probabilities and risks for many measurement-based decision scenarios can be minimized or even eliminated, while still taking measurement uncertainty into account.

### **3.3 Error versus uncertainty**

Not only does the incorporation of measurement uncertainty complicate conformity assessment decision-making, the language used to make such decisions can sometimes be confusing, and even appear to be contradictory. Most notably, while the concepts of “error” and “uncertainty” share a certain similarity, in that they are both related to the quality of a measurement, they are actually significantly different concepts. Perhaps ironically, an “error of indication” is something that can itself be measured, and thus have a value with an associated measurement uncertainty. This difference between “error” and “uncertainty,” and how they coexist in legal metrology (and other areas of metrology), is elaborated in [16] and in Annex A.

### **3.4 Verification incorporating measurement uncertainty**

While the concept of measurement uncertainty as elaborated in the GUM is relatively new (about 20 years), verification in legal metrology has always incorporated some notion of measurement uncertainty in the sense that Maximum Permissible Errors (MPEs) have usually been established so as to account for plausible measurement uncertainty, at least implicitly. One example is the practice of establishing conservative (in-service) MPEs in order to draw “safe” conclusions concerning whether measured errors of indication are within acceptable limits. The practice of specifying a fraction, such as 1/3 or 1/5, for the maximum allowed ratio of the error (actually, uncertainty) of the standard (reference) measuring instrument to the MPE is another example of at least implicitly accounting for measurement uncertainty. One of the important topics discussed in this Guide is when and how to implicitly, rather than explicitly, incorporate measurement uncertainty into conformity decisions for testing and verification scenarios, so that measurement traceability can be established and maintained (see 6) when subsequently using the measuring instrument or system.

### **3.5 MPEs and measurement uncertainty**

Measurement uncertainty considerations also enter into establishing appropriate MPEs for given testing scenarios. The costs to the consumer, vendor or manufacturer associated with the specification of MPEs that are unnecessarily large or small can be reduced through taking likely measurement uncertainty into account when first establishing the MPEs. Setting MPEs that are very small can be costly to the instrument manufacturer that will have to design and build a more costly instrument to meet the tighter requirements for a given application and will most likely pass the additional cost on to the consumer. By considering likely levels of measurement uncertainty for different applications and uses of measuring instruments, MPEs can be set to yield acceptable levels of risk more cost-effectively. Clause 7 briefly discusses options for taking measurement uncertainty into account when prescribing MPEs in OIML Recommendations and other OIML publications (see also OIML R 34 *Accuracy classes of measuring instruments* [13]).

For convenience, options containing specific language pertaining to explicitly and implicitly incorporating measurement uncertainty that should be considered for inclusion in OIML Recommendations or other OIML publications are provided in clause 8.

#### 4 Basic considerations pertaining to conformity testing decisions and measurement uncertainty

One of the key roles of legal metrology is to evaluate the performance and conformance of designs (or types) of measuring instruments and systems (type evaluation), as well as the performance of individual measuring instruments and systems (initial or subsequent verification), for various regulated applications. The basic kind of test that is used to conduct such evaluations involves comparing the ‘error of indication’ with a ‘maximum permissible error’ (MPE) that is specified for the particular application. The error of indication (denoted  $E_I$ ) is typically defined as the difference between the indicated value of the measuring instrument or system obtained when measuring the measurand, and the ‘true’ value of that measurand. Since it is not possible to perform a ‘perfect’ measurement, and so the ‘true’ value of the measurand cannot be known, the error of indication is operationally taken to be the difference between the indicated value ( $Y_I$ ) of the measuring instrument or system obtained when measuring the measurand, and the value ( $Y_S$ ) of the same measurand as determined when using a measurement standard. Expressed mathematically:

$$E_I = Y_I - Y_S. \quad (4.1)$$

Note that, historically, in legal metrology the term “true value” is usually not used in the sense given here, but rather is used to mean the value associated with a measurement standard that is used in the process of testing a measuring instrument. This latter meaning is not the meaning of the term ‘true’ value in this Guide; see Annex A and JCGM 106 [5] for more detail.

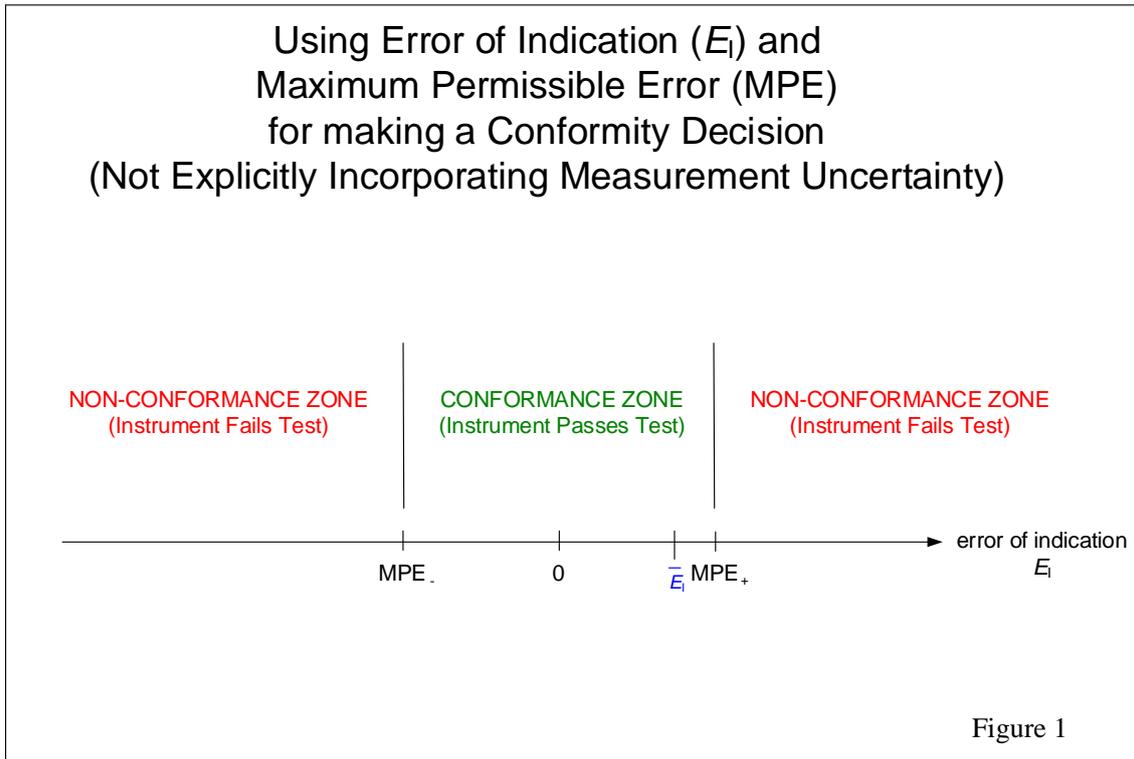
Most commonly,  $Y_S$  is an indicated value obtained directly from the indication of a measurement standard, or from a calibration certificate of the measurement standard.

For more complicated measurement standards (or systems),  $Y_S$  can be determined through use of a ‘measurement model’ [1][2][3][4] that relates the value of the measurand to values ( $X_i$ ) of ‘input quantities in a measurement model’ [4] (that is,  $Y_S$  depends on, or is a function ( $f$ ) of, the values  $X_i$ ):

$$Y_S = f(X_1, X_2, \dots, X_n). \quad (4.2)$$

Depending on the category of test being performed (type evaluation, initial verification, or subsequent verification), there can be wide variation in conducting the test. The specification of a particular test may include the number of individual errors of indication that should be obtained (through repeated measurements), and when and how the operating conditions of the instrument should be controlled (if at all). Common to all of the categories of tests, however, is that conformity decisions are ultimately made based on the results of one or more tests that compare measured errors of indication with MPEs.

The comparison of a measured error of indication with the MPEs, for the purpose of making a conformity decision, is shown schematically in Figure 1. The horizontal axis represents possible values of the error of indication  $E_I$ . The upper and lower MPEs, denoted  $MPE_+$  and  $MPE_-$ , respectively, are shown to be symmetric about 0, but this may not always be necessary. If only a single measured error of indication is to be used to make a conformity decision, then if that single measured error of indication lies within the interval defined by the MPEs (denoted as “Conformance Zone” in Figure 1), the instrument is considered to pass that particular test (as shown in Figure 1). Otherwise, the instrument is considered to fail that test. Note that measurement uncertainty is not being explicitly considered in this discussion or in Figure 1, however the MPEs are assumed to have been established on the basis of likely levels of measurement uncertainty for the particular type of measurement.



Note that in some OIML Recommendations, in order to account for random variation in the measured values of error of indication, tests are structured such that individual conformity decisions are not based on a single measured error of indication, but rather are based on obtaining two or more errors of indication and using the average value as the basis of the conformity decision. This is illustrated by the use of the symbol  $\bar{E}_I$  in Figure 1, where the test would be considered to pass since  $\bar{E}_I$  lies in the conformance zone. Yet another variation is to obtain two or more measured errors of indication, and then require that a certain fraction of them (say, two out of three) lie in the conformance zone. When measurement uncertainty is taken into account, as will be demonstrated in clause 5, the differences between these ways of making a conformity decision may not lead to different decision outcomes, since random variation in the measurement is incorporated into the measurement uncertainty.

## 5 Conformity testing decisions that explicitly incorporate measurement uncertainty

As indicated in the Introduction, incorporating the concept of measurement uncertainty into conformity testing decisions in legal metrology requires a different way of thinking and talking about such decisions than is described in clause 4 (see Annex A and JCGM 106 [5]). Rather than being able to definitively state that a measuring instrument meets specified MPE requirements and so passes a particular conformity test, only a degree of belief (or probability, expressed as a level of confidence) can be stated that the measuring instrument conforms for each MPE requirement. Inherent in such a probabilistic approach is that certain risks should be considered (e.g. a risk that a decision is incorrect) when ultimately making a pass/fail decision. Measurement uncertainty is used in the process of establishing quantitative values of such probabilities and risks.

It is assumed that the reader of this Guide has some familiarity with the concept of measurement uncertainty and with the GUM procedure for calculating it. However, for those who are not familiar, examples are provided in Annex C. A GUM Supplement [2] is also available that discusses another approach to calculating measurement uncertainty based on the Monte Carlo method.

ISO/IEC 17025 [9] has become a widely accepted standard used in the international laboratory accreditation community for assessing the competence of calibration and testing laboratories. This standard states that “[5.4.6.2] testing laboratories shall have and shall apply procedures for estimating uncertainty of measurement,” and, further “[5.4.6.3] When estimating the uncertainty of measurement, all uncertainty components which are of importance in the given situation shall be taken into account using appropriate methods of analysis”.

*Note:* See 5.4.6.2 Note 2 of ISO/IEC 17025: *In those cases where a well-recognized test method specifies limits to the values of the major sources of uncertainty of measurement and specifies the form of presentation of calculated results, the laboratory is considered to have satisfied this clause by following the test method and reporting instructions (see 5.10).*

Clause 5 focuses on the explicit use of measurement uncertainty for the purposes of making conformity decisions, such as when measurements are performed in a laboratory environment. Clause 6 focuses on the implicit use of measurement uncertainty for making conformity decisions, such as when measurements are performed outside of a laboratory environment or when using a measurement system for which it would be difficult to assess the uncertainty of the results explicitly. It is important to realize, however, that measurement uncertainty is being accounted for in both measurement environments, so that measurement results obtained when later using the tested measuring instruments can be traceable.

Independent of the measurement applied in the process of testing for type evaluation purposes specified in an OIML Recommendation (or other OIML publication), guidance should be provided in OIML publications on practical and efficient methods that can be used

---

to calculate measurement uncertainty for the measurement model(s) that is appropriate to the type of instrument(s) covered in the Recommendation.

In particular, guidance should be provided in OIML publications on how to describe the test apparatus (including the measurement standard and any additional measuring equipment), and on how to set up a measurement model (as in Equation 4.2) and identify the input quantities. In order to establish the acceptance or failure of a particular measuring instrument or system, it may be necessary to separately keep track of uncertainties arising from influences affecting the test apparatus and uncertainties due to influence quantities affecting the measuring instrument or system under test. If this is not done, an otherwise good measuring instrument could be improperly rejected on the basis of uncertainty considerations.

Guidance should then be provided in OIML publications on methods that can be used to identify or calculate the standard measurement uncertainty ( $u_S$ ) associated with the test apparatus (including the measurement standard and any additional measuring equipment). Ideally, uncertainties associated with influence quantities affecting the test apparatus can be kept small with respect to the maximum permissible errors (MPEs).

Similarly, guidance should be provided in OIML publications on methods that can be used to calculate a standard uncertainty ( $u_I$ ) associated with the indicated value of the measurand (including uncertainty components due to indicator resolution, jitter, etc.), and a standard uncertainty ( $u_{rep}$ ) associated with repeatability or reproducibility of both the instrument under test and the measuring system and/or procedure.

If the indication of the measuring instrument is found to vary over the range of rated operating conditions of the instrument (for a fixed input to the instrument), then a component of measurement uncertainty ( $u_{roc}$ ) should be included to cover this.

Finally, guidance should be provided in OIML publications on how to combine these components of measurement uncertainty in order to calculate a combined standard uncertainty ( $u_{E_I}$ ) associated with the error of indication (based on using Equation 4.1).

All of this guidance should be based on and consistent with the methods of the GUM and its Supplements. Examples of GUM-consistent procedures for establishing a measurement model, identifying and estimating individual components of measurement uncertainty, and finally calculating the combined standard and expanded uncertainties associated with the error of indication, are provided in Annex C.

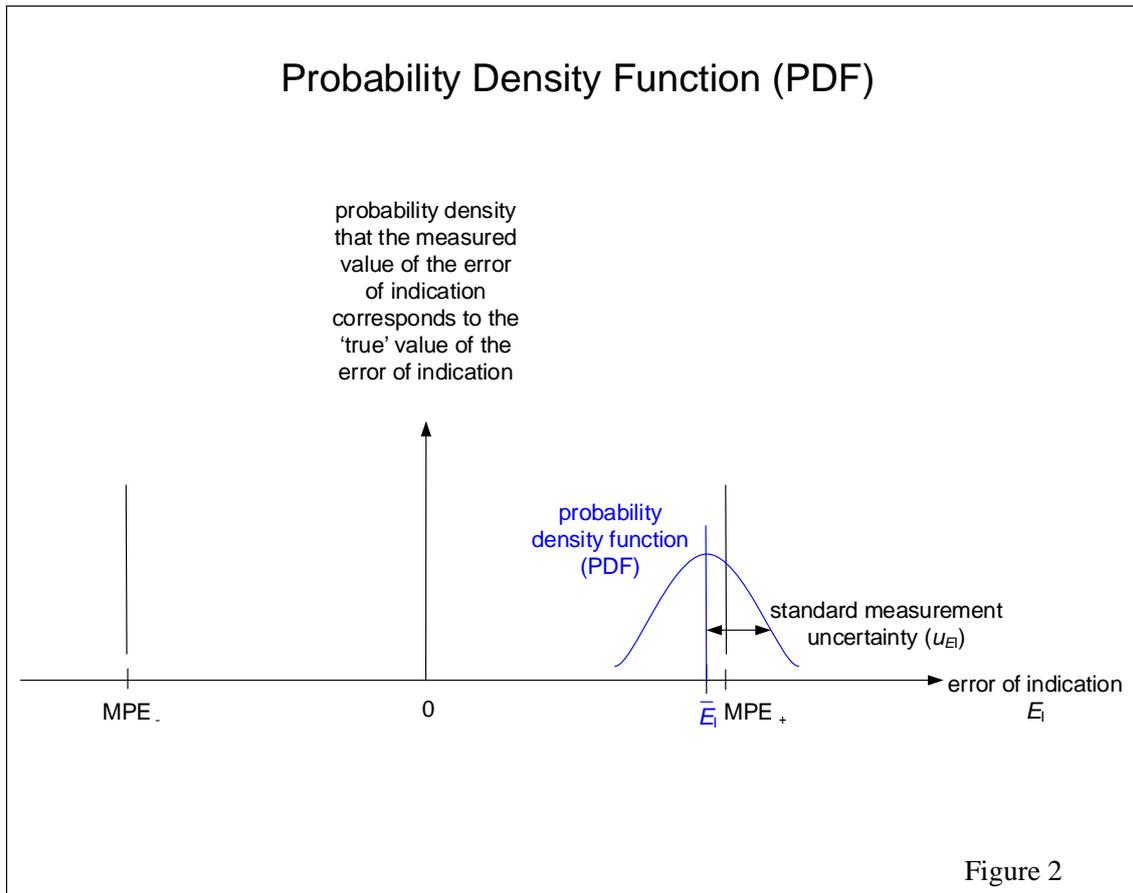
It is important to note that in cases where multiple measurements of a particular error of indication are made for the purpose of assessing the repeatability or reproducibility of the measurement process, it is not necessary to assess the measurement uncertainty associated with each of the individual measured values of error of indication. Rather, the mean value of error of indication ( $\bar{E}_I$ ) can be calculated from the set of individual measured values and used as the 'measured' error of indication, and the standard deviation of the mean of the set of individual values (i.e.,  $S/n^{1/2}$  where  $S$  is the experimental standard deviation and  $n$  is the number of measured values in the set) can be used as a component of the measurement

uncertainty that should be associated with the mean value. OIML Recommendations (and other OIML publications) should emphasize, however, that a measurement uncertainty based on random effects alone is not the entire measurement uncertainty, and that other components of measurement uncertainty, such as from systematic effects, must also be included.

Subclauses 5.1 to 5.3.6 discuss the ways in which the calculated combined standard measurement uncertainty associated with the error of indication ( $u_{E_I}$ ) can and should be used in order to make conformity decisions for instruments/systems under test. See JCGM 106 [5] for further options and details.

### **5.1 Probability density function (PDF)**

Inherent in the concept of measurement uncertainty is that the ‘true’ value of the quantity that is intended to be measured cannot be known, since it is impossible to know whether a mistake was made when performing the measurement. And even if it were known that no mistakes had been made in performing a measurement, virtually all measurements have some associated unknown systematic aspects and random variations that are not fully controlled or understood. Accordingly, one should talk in terms of knowing the true value of the measurand on a probabilistic basis, where some values are thought to be more likely than others to correspond to the true value of the measurand. One way of viewing this is that a function, known as a probability density function, can be constructed that gives one’s degree of belief (level of confidence) about knowing the true value of the measurand.



The concept of the probability density function (PDF) is shown schematically in Figure 2. As in Figure 1, the horizontal axis represents possible values of error of indication  $E_I$ . In Figure 2, a vertical axis has been added that represents possible probability densities that the true value of the error of indication of an individual measuring instrument lies within an infinitesimal region around a particular value of error of indication. The probability (or degree of belief, expressed as a level of confidence, based on the assumption that no mistakes have been made) that the true value of the error of indication lies somewhere on the horizontal axis between two specified values of error of indication can be obtained by mathematically integrating the area under the probability density function curve bounded by the two specified values.

The PDF curve is shown as Gaussian in shape, which is commonly used (but not always; e.g. see [2]). The mean value ( $\bar{E}_I$ ) of the curve and the standard measurement uncertainty (denoted by  $u_{E_I}$ ) are indicated. The curve is normalized such that the total area under the curve is 1, meaning that there is a 100 % probability of finding the true value of the error of indication somewhere along the horizontal axis. While this must be the case, it is worth noting that the 'true' value of the error of indication might actually be very far from the mean of the PDF curve, such as if a mistake was made in performing the measurement. Note that for a Gaussian PDF, the probability (degree of belief or level of confidence) that the true value of the error of indication is within the interval  $\bar{E}_I \pm u_{E_I}$  is 68 %, and within the interval  $\bar{E}_I \pm 2 \cdot u_{E_I}$

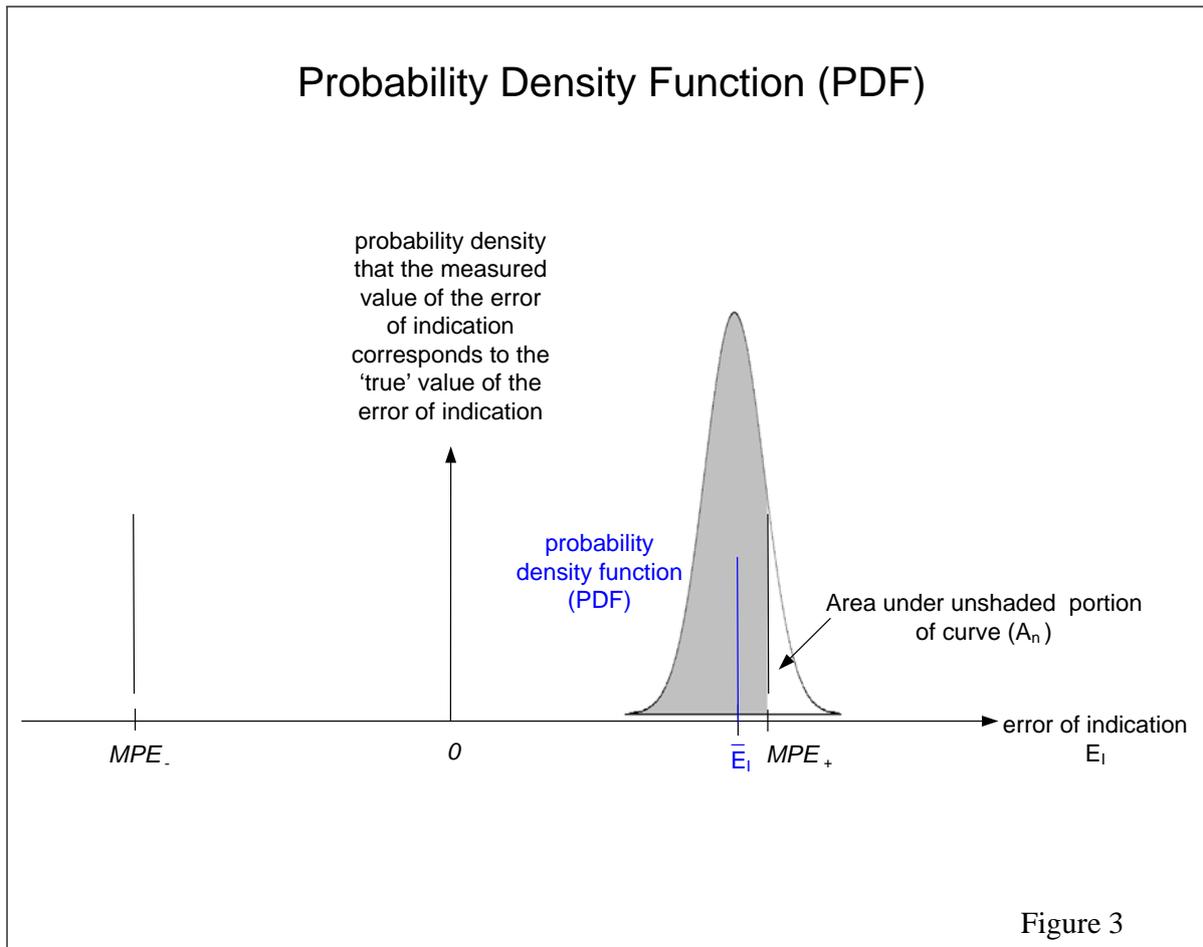
is 95 %. In general, the interval can be written  $\bar{E}_1 \pm U_{E_1}$ , where  $U_{E_1} (= k \cdot u_{E_1})$  is called the expanded measurement uncertainty, and  $k$  is called the coverage factor.

It is also worth reemphasizing that the PDF encodes all of the known information about the measurand, including both systematic and random effects. While a curve fit to a histogram of random fluctuations alone frequently has a Gaussian shape, the PDF is not such a fit to a histogram, but rather contains additional information coming from systematic effects in the measurement.

## 5.2 Probability of conformity

Figure 2 can be used to demonstrate the important differences in making conformity decisions using the classical approach, discussed in clause 4, and using the GUM uncertainty approach. Using the classical approach, since the mean value ( $\bar{E}_1$ ) of the error of indication is within the conformance zone as defined in Figure 1, the measuring instrument would be considered to pass the particular test shown in Figure 2.

Using the uncertainty approach and taking measurement uncertainty into account for the particular test, it can be seen in Figure 2 that there is a considerable area under the PDF curve that lies outside of the conformance zone (that is, to the right of  $MPE_+$ ), which means that there is a considerable probability (degree of belief or level of confidence) that the true value of the error of indication lies outside of the conformance zone, even though the mean value ( $\bar{E}_1$ ) of the error of indication is within the conformance zone.



If the area under the PDF curve that lies outside of the conformance zone (as indicated by the un-shaded area under the Gaussian curve in Figure 3) is denoted by  $A_n$ , (where “n” stands for ‘nonconformance’) then the probability  $p_n$  that the true value of the error of indication is outside of the conformance zone, and hence that the measuring instrument does not conform to the MPE requirement, is given by  $p_n = A_n$ , or  $p_n$  (in %) =  $100 \cdot A_n$ . A decision about whether or not the measuring instrument is considered to pass the particular test could then depend upon whether acceptable levels of probability (risk) were met for that kind of test. For example, the measuring instrument could be considered to pass the particular test if there was less than a 10 % probability that it was non-conforming, meaning  $p_n = A_n < 0.1 = 10\%$ .

Note that if the mean value of the error of indication ( $\bar{E}_I$ ) is just slightly outside of the conformance zone, there can still be a significant probability that the true value of the error of indication lies within the conformance zone. Although the measuring instrument would fail the particular test if measurement uncertainty is not taken into account, the test could still result in a “pass” in this case when taking into account the acceptable level of risk and the risk bearing party. If  $\bar{E}_I$  is exactly equal to  $MPE_+$ , then there is a 50 % probability that the error of indication lies within the conformance zone and a 50 % probability that it is outside the conformance zone. The issue of risk assessment, along with rules for deciding whether a particular test results in a pass or fail, will be addressed in the next subclause.

Constructing PDFs and calculating areas under a PDF curve is in general a nontrivial matter, and so any advice and assistance to be provided in this regard in the Recommendation(s) should be carefully considered during its development. When the PDF can be treated as Gaussian, there is a convenient method that incorporates what is known as the ‘standard normal distribution table’ (or Z-table) for calculating the area under the curve for a specified  $\bar{E}_L$ ,  $MPE_+$  and  $u_{E_L}$  [19]. Annex B provides information about the standard normal distribution table, along with an example of how to use it.

### **5.3 “Risks” and “decision rules” associated with conformity decisions**

As already discussed, because of the probabilistic nature of the GUM approach to measurement uncertainty, making a pass-fail decision based on whether or not the measured value of the error of indication lies within the region bounded by the MPEs carries with it the possibility (or risk) that an incorrect decision has been made. That is, the true value of the error of indication may actually lie in a region bounded by the MPEs that is different than the region where the measured value lies. This subclause discusses the types of risks associated with incorporating measurement uncertainty into the decision-making process, and the rules that can be applied to making conformity decisions for testing in legal metrology. These rules should be considered for possible incorporation into OIML Recommendations and other OIML publications during their review or development.

Various treatments and names have been given to the different types of risks associated with making conformity decisions for tests that are based on meeting tolerance interval requirements such as MPEs [5][19]. As a summary, there are three fundamental types of risks: 1) risk of false acceptance of a test, 2) risk of false rejection of a test result, and 3) shared risk.

#### **5.3.1 Risk and decision rule for false acceptance**

The risk of false acceptance means that the test is considered to have been passed, but in reality the MPE requirement might not have been met. In this case, the measured value of the error of indication lies within the region bounded by the MPEs, but the PDF extends into the region outside of the region bounded by the MPEs, as shown in Figure 3, meaning that the true value of the error of indication is believed to possibly lie outside of the region bounded by the MPEs. Note that the risk of false acceptance is taken by the evaluator or user of the measuring instrument or system. The risk is that the instrument or system is not performing ‘within specification’ even though the test result says it is. The value of the risk of false acceptance is calculated as the area  $A_n$  under the PDF curve that is outside of the region bounded by the MPEs, which is the un-shaded area under the curve in Figure 3.

A possible decision rule that can be associated with a legal metrology test is that the probability or risk of false acceptance ( $p_{fa}$ ) be less than some stated value (for example, 5 %). This risk would favor the evaluator or user of the instrument/system, to the detriment of the

manufacturer or seller of the instrument/system, since the measured value of the error of indication  $\bar{E}_1$  would lie within the region bounded by the MPEs, and, further, could usually not even lie very close to the relevant MPE boundary if the decision rule is to be met (see example in Annex B).

### 5.3.2 Risk and decision rule for false rejection

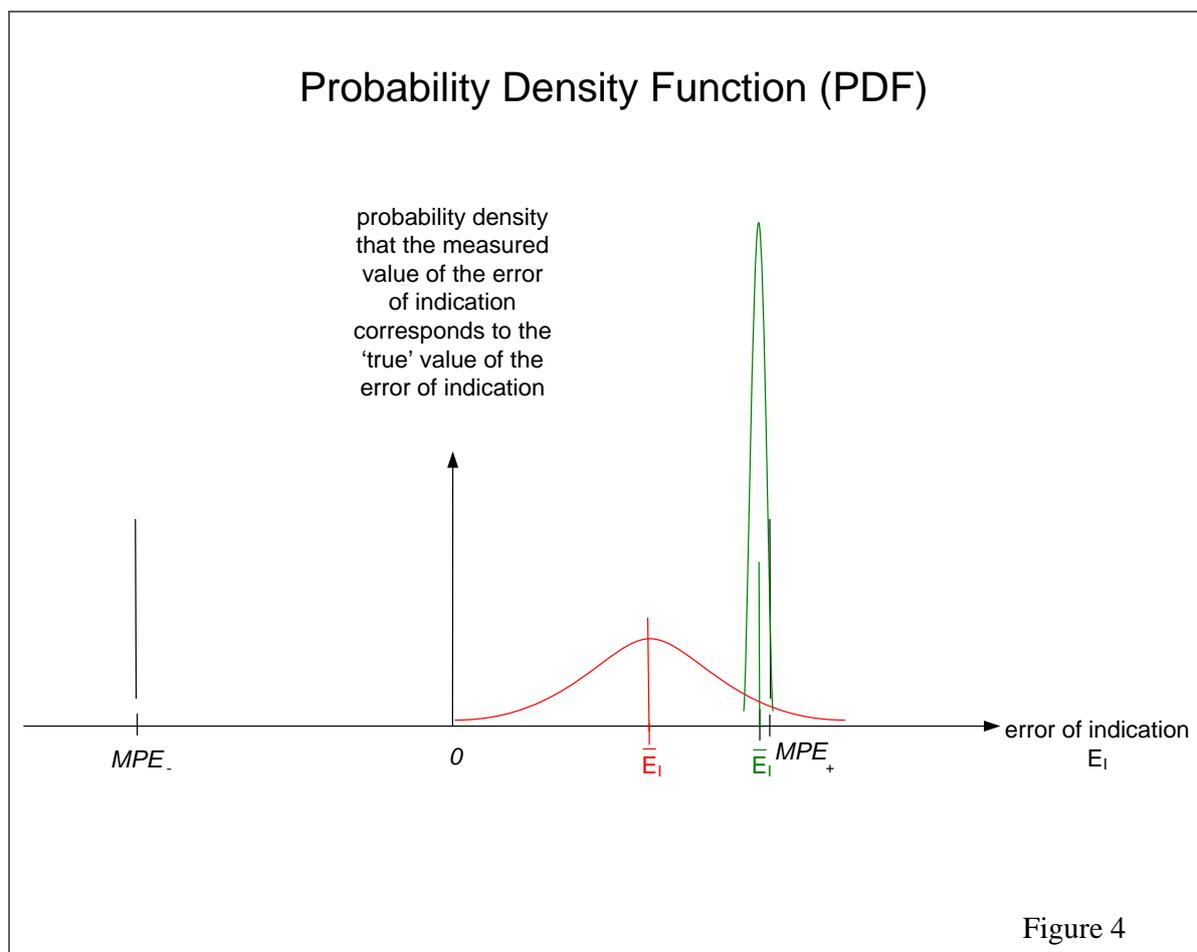
Conversely, risk of false rejection means that the test is considered to have failed, but in reality the MPE requirement might have been met. In this case, the measured value of the error of indication lies outside the region bounded by the MPEs, but the PDF extends into the region inside of the region bounded by the MPEs. Note that the risk of false rejection is taken by the manufacturer or seller of the measuring instrument or system. The risk is that the instrument/system is performing ‘within specification’, even though the test result says it is not. The value of the risk of false rejection is calculated as the area under the PDF that is inside of the region bounded by the MPEs when the measured value of the error of indication lies outside the region bounded by the MPEs.

A possible decision rule that can be associated with a legal metrology test is that the risk of false rejection ( $p_{fr}$ ) be less than some stated value (for example, 2 %). This risk would favor the manufacturer or seller of the instrument/system, to the detriment of the evaluator or user of the instrument/system, since the measured value of the error of indication  $\bar{E}_1$  would lie outside of the region bounded by the MPEs, and, further, could usually not even lie very close to the relevant MPE boundary if the decision rule is to be met.

It is important to note that it is not possible to have a decision rule for a given test that incorporates both risk of false acceptance and risk of false rejection. That is, the ‘advantage’ can go to either the evaluator/user or the manufacturer/seller, but not to both at the same time! It is also important to note that the PDF must be known in order to calculate the risk of false acceptance or false rejection.

### 5.3.3 Shared risk

Shared risk, on the other hand, is an agreement between the parties concerned with the outcome of the testing that neither will be given an advantage or disadvantage concerning consideration of measurement uncertainty. Implicit in such an agreement is that the expanded measurement uncertainty  $U_{E_1}$  is ‘small’ with respect to the MPE (i.e. the ratio  $(U_{E_1}/\text{MPE})$  is ‘small’) so that the significant risk of an erroneous decision exists for values of  $\bar{E}_1$  that are only very close to the MPE boundaries. This is illustrated in Figure 4 for two possible different PDFs for a given measurement. The uncertainty  $U_{E_1}$  associated with the leftmost (red) Gaussian curve is probably too large for a shared risk arrangement, whereas the uncertainty  $U_{E_1}$  associated with the rightmost (green) Gaussian curve would probably be acceptable for most applications.



An important advantage of the shared risk approach is that it is not necessary to know the PDF for the error of indication, since the risk is shared equally and so no risk calculations are necessary. This advantage makes use of the shared risk approach, which is highly desirable when considering what decision rule to propose in an OIML Recommendation or other OIML publication, since it at least partially simplifies the decision-making process.

In fact, many OIML Recommendations currently, at least implicitly, use the shared risk approach. In order to meet the requirements in ISO/IEC 17025 [9] that measurement uncertainty be taken into account, at least at some level of rigor, for all measurements it is highly recommended that, when this is the case, explicit text should be included in OIML publications specifying that the shared risk principle is being used.

Note that with the shared risk approach it is still necessary to calculate the measurement uncertainty  $U_{E_i}$  so that the ratio ( $U_{E_i}/MPE$ ) can be examined to see if it is 'small enough', as discussed in 5.3.4. Also note that if the maximum permissible errors are to be adjusted for some reason (for example, allowance for in-service conditions) using the guard band method (see 5.3.6), the shared risk approach can still be used with the new or guard banded MPEs.

### 5.3.4 Maximum permissible uncertainty of error of indication

It is becoming common (e.g. [20]) to refer to the maximum value that the ratio ( $U_{E_1}/\text{MPE}$ ) is allowed to have in terms of a “maximum permissible uncertainty” (denoted symbolically by  $\text{MPU}_{E_1}$ ) of the error of indication, defined by:

$$\text{MPU}_{E_1} \equiv f_{E_1} \cdot \text{MPE} \quad (5.1)$$

where  $f_{E_1}$  is a specified number less than one, usually of the order 1/3 or 1/5 (0.33 or 0.2) [17].

*Note:*  $f_{E_1} < 1$  is not always true in some OIML Recommendations (e.g. R 76 for non-automatic weighing instruments, and also sometimes for load cells and automatic weighing instruments), especially when the measured values of the errors of indication are all very close to the maximum permissible errors.

The maximum permissible uncertainty ( $\text{MPU}_{E_1}$ ) is typically thought of as the largest value that  $U_{E_1}$  can have for a given measurement of the error of indication  $\bar{E}_1$  for which the shared risk approach can be used. The decision rule to be applied concerning  $\text{MPU}_{E_1}$  is that if  $U_{E_1}$  is greater than  $\text{MPU}_{E_1}$  then the test is considered to fail, and means for reducing  $U_{E_1}$  (or for incorporating an increased MPE) will need to be developed.

Another way of thinking about the need for specifying an  $\text{MPU}_{E_1}$  is that if  $U_{E_1}$  is comparable to the MPE, then for values of  $\bar{E}_1$  that are, say, around halfway between 0 and  $\text{MPE}_+$ , as shown by the leftmost curve in Figure 4, there can be a relatively large probability that the true value of the error of indication lies far to the right of  $\text{MPE}_+$  (i.e., when  $E_1$  lies very close to  $\text{MPE}_+$ ), which is an unacceptable risk in many cases. By having an  $\text{MPU}_{E_1}$ , such a possibility is eliminated.

Note that  $1/f_{E_1}$  is sometimes called the test uncertainty ratio (TUR). Also note that if the uncertainty associated with the measurement standard ( $U_S$ ) is much larger than the uncertainty associated with the other components contributing to  $U_{E_1}$ , then  $\text{MPU}_{E_1}$  is about equal to the ‘maximum permissible uncertainty (of the measurement standard)’ (denoted symbolically by  $\text{MPU}_S$ ) (see 5.3.5).

It is worth reemphasizing here that  $U_{E_1}$  is not just the expanded uncertainty associated with the measuring instrument under evaluation, but encompasses the uncertainty associated with the entire test apparatus and any effects due to environmental conditions. That is, the measuring instrument under evaluation is assumed to be operating within its specified rated operating conditions when the measured errors of indication are obtained. If the actual operating conditions fluctuate outside of the rated operating conditions, then additional measurement uncertainty might need to be taken into account.

### 5.3.5 Maximum permissible uncertainty of the measurement standard

Besides the need for specification of a ‘maximum permissible uncertainty (of error of indication)’, for the reasons given above, another decision rule that is frequently used is to specify a ‘maximum permissible uncertainty (of the measurement standard)’ (denoted symbolically by  $MPU_S$ ), defined by:

$$MPU_S \equiv f_S \cdot MPE \quad (5.2)$$

where  $f_S$  is a specified number less than one, also usually of the order 1/3 or 1/5 (0.33 or 0.2). Then the maximum permissible uncertainty ( $MPU_S$ ) is the largest value that  $U_S$  is allowed have for a given measurement of the error of indication  $\bar{E}_I$ .

The rationale for this requirement is that if  $MPU_S$  is too large, then the pass-fail decision based on  $MPU_{E_I}$  above can become dominated by the quality of the measurement standard and/or testing laboratory, rather than on the quality of the instrument/system being tested (note that  $U_{E_I}$  contains  $U_S$  as well as other components of uncertainty). It could be considered unfair to test the instrument manufacturer’s instrument with a measurement standard that has an uncertainty that comprises most of  $U_{E_I}$ , since then the uncertainty associated with the indicated value ( $U_I$ ), as well as other possible components of uncertainty associated with the instrument/system, would need to be relatively small in order that the uncertainty associated with the error of indication remains acceptably small for the particular test (i.e. less than  $MPU_{E_I}$ ). By requiring that  $f_S$  be relatively small (say, less than 1/5), then any significant differences or discrepancies among testing laboratories can be avoided. Individual OIML Recommendations should therefore specify an acceptable  $f_S$  (or  $MPU_S$ ) that is appropriate to each particular kind of test.

Note that  $1/f_S$  is sometimes called the test accuracy ratio (TAR), although in such case  $MPU_S$  is treated as a maximum permissible error (of the standard), since TAR is customarily considered to be a ratio of errors. Also note that if the uncertainty of the measurement standard is the major component of the total uncertainty, then  $MPU_{E_I}$  is about equal to  $MPU_S$ , which is an undesirable situation unless the total uncertainty is much smaller than the MPE.

### 5.3.6 Summary of considerations for decision rules

When considering which decision rules should be incorporated into OIML Recommendations and other OIML publications, the consequences of an incorrect decision when proposing acceptable levels of risk should be taken into account. If the consequences of false acceptance are not considered to be too severe, incorporating the shared risk approach should be promoted, since it is a relatively efficient means of deciding conformity while still taking measurement uncertainty into account. It is usually the case in legal metrology that the shared risk approach can be used successfully for a test, as long as the corresponding MPE for that kind of test does not need to be too ‘small’ (see 7) and that the measurement uncertainty of the error of indication is not too large.

For many legal metrology situations, MPEs are used that anticipate the likely level of measurement uncertainty, so that risk has already been taken into consideration. This condition should be documented when this is the case so that double accounting of measurement uncertainty does not occur.

Whether to use  $f_{E_1}$  or  $f_S$  (TUR or TAR) for purposes of deciding whether the shared risk approach is appropriate depends on the level of information and resources available, and the consequences of making an incorrect decision. While  $f_S$  (TAR) alone is the easiest to determine, typically by using only a manufacturer's accuracy specification,  $f_{E_1}$  (TUR) is the safest to use, since it takes all significant components of uncertainty explicitly into account.

If the shared risk approach cannot be used, and it is instead necessary to use the risk of false acceptance for making a conformity decision, there is a convenient means of doing this, that can minimize the time and effort required by the test evaluator, utilizing the concept of the "measurement capability index" [5], defined for purposes of legal metrology as  $C_m = \text{MPE}/(2 \cdot u_{E_1})$ . Note that  $C_m$  is proportional to  $\text{MPU}_{E_1}$ , and inversely proportional to  $f_{E_1}$ . Annex E provides a discussion and example of how the measurement capability index can be used to make a relatively 'quick' decision on a test when the MPE, risk of false acceptance ( $p_{fa}$ ), measured  $E_1$  and calculated  $u_{E_1}$  are all known.

For those special cases of using risk of false acceptance (or false rejection) where the standard uncertainty associated with the error of indication ( $u_{E_1}$ ) can be considered to be constant (i.e. it is the same for each error of indication), then a particularly convenient method can be used for making conformity decisions, known as "guard banding". Under such conditions, the MPE boundaries are simply 'shifted' inward (for false acceptance) or outward (for false rejection) by an amount corresponding to the respective risks, and conformity decisions are then made on the basis of whether the measured error of indication ( $E_1$ ) lies within or outside of the shifted conformity boundaries. Reference [5] provides a very useful discussion of the guard band principle. For type approval in legal metrology, only guard bands that are shifted inward are used.

While decision rules and associated risks, along with their consequences, should be considered and discussed in OIML Recommendations, it should also be carefully considered whether specified levels of acceptable probability for various types of tests are required or even suggested. If so, this should be done only in the context of regulatory matters. Suggestions may be provided in Recommendations, although typically this should be left up to national or regional regulations. Different risks may have serious economic consequences for different parties, and the specification of such risks is typically outside the scope of a Recommendation.

## **6 Conformity testing decisions that do not explicitly incorporate measurement uncertainty**

As mentioned in the Introduction, consideration of measurement uncertainty is widely recognized, in both the metrology and laboratory accreditation communities, as being essential in order to have metrological traceability of measurement results. As also mentioned, many tests in legal metrology are performed outside of a laboratory environment intended to allow for ‘quick and easy’ pass-or-fail decisions, and so measurement uncertainty is sometimes only provided implicitly. It is therefore important to consider how to maintain metrological traceability of measurement results outside of a laboratory environment when measurement uncertainty is not explicitly provided.

As an example of the necessity of taking measurement uncertainty into account when performing a measurement, according to 3.1.2 of the GUM “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”. The VIM3 [14] defines “metrological traceability” as “2.41 property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty”, so that in order for a measurement result to have traceability the measurement uncertainty must at least be considered, if not explicitly provided.

Sometimes when measurements are performed outside of a laboratory environment the measurement uncertainty is believed to be insignificant. The GUM states in 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”.

In this case, “error” means “error of indication”. While the GUM recognizes here that there are situations where “the components of uncertainty are small enough to be ignored”, it is important to recognize that this must be demonstrated somehow and documented somewhere, and not just assumed.

The difficulties sometimes associated with assessing measurement uncertainty, even in a testing laboratory environment, are recognized in ISO/IEC 17025 [9], where in 5.4.6.2 it is stated that:

“Testing laboratories shall have and shall apply procedures for estimating uncertainty of measurement. In certain cases, the nature of the test method may preclude rigorous, metrologically and statistically valid, calculation of uncertainty of measurement. In these cases the laboratory shall at least attempt to identify all the components of

uncertainty and make a reasonable estimation, and shall ensure that the form of reporting of the result does not give a wrong impression of the uncertainty. Reasonable estimation shall be based on knowledge of the performance of the method and on the measurement scope and shall make use of, for example, previous experience and validation data. NOTE 1 The degree of rigor needed in an estimation of uncertainty depends on factors such as – the requirements of the test method; the requirements of the customer; - the existence of narrow limits on which decisions on conformity to a specification are based. NOTE 2 In those cases where a well-recognized test method specifies limits to the values of the major sources of uncertainty of measurement and specifies the form of presentation of calculated results, the laboratory is considered to have satisfied this clause by following the test method and reporting instructions”.

The same is also assumed to be true for measurements performed outside of a laboratory environment.

Finally, the GUM also states in 7.1.3:

“Numerous measurements are made every day in industry and commerce without any explicit report of uncertainty. However, many are performed with instruments subject to periodic calibration or legal inspection. If the instruments are known to be in conformance with their specifications or with the existing normative documents that apply, the uncertainties of their indications may be inferred from these specifications or from these normative documents”.

The use of normative documents or specifications is important for legal metrology, since it opens the way for not having to explicitly report measurement uncertainty for every measurement performed outside of a laboratory environment, where reporting uncertainties is frequently impractical.

Everything presented above in this clause is intended to reinforce the notion of not having to always explicitly compute and report measurement uncertainty in order to still be able to claim metrological traceability of measurement results. What is important to appreciate, however, is that there is always an underlying understanding that the level of uncertainty in the measurement results has been assured and that the method of assurance is well documented. This assurance is usually conducted at higher levels of the organization responsible for providing regulation, and not by the person who actually performs the verification measurements. It is further understood that, if challenged, a credible measurement uncertainty can be provided for such measurements performed outside of a laboratory environment (such as for purposes of verification of an individual measuring instrument).

For example, when specified test accuracy ratios (TARs) have been maintained throughout the chain of calibrations in a laboratory or measurement or testing system in an environment outside of a laboratory, an upper bound on the standard uncertainty of the resulting

measurements often can be obtained fairly easily. This can be done without explicit provision of measurement uncertainties at each step assuming that at each link in the calibration chain<sup>1</sup>

- the TAR in use is maintained at a high enough level, typically 4:1,
- systems are in place that mitigate the effects of any potential sources of uncertainty not accounted for in the TAR being used, and
- the test accuracy ratio (TAR) in use is defined so that its numerator and denominator can be related to multiples<sup>2</sup> (i.e. coverage factors) of the standard uncertainties for future measurements made using the device under test and of the measurement standard used to establish the traceability of the device under test.

When the conditions above have been met, a standard uncertainty for the measurements made using the device under test can be evaluated fairly simply using methods described in the GUM. The value of this standard uncertainty will be based on the performance specification for the device under test, the TAR level maintained throughout the calibration chain, and the ratio of the relevant multiples of the standard uncertainties for the device under test and the calibration correction discussed above and inherent in the definition of the TAR in use.

To illustrate this further, consider the definition of the TAR where the numerator of the TAR is the half-range of the device specification, which would correspond to a range of about three standard uncertainties assuming the measurement errors of the device follow a Gaussian distribution. Assume the denominator is the half-range of the potential measurement standard values at the 95 % confidence level, which would generally correspond to a range of about two standard uncertainties. Therefore the ratio of the relevant multiples of the standard uncertainties for this TAR would be 6/4. When combined with the value of the TAR maintained at each step in the calibration chain (e.g.  $\text{TAR} \geq 4$ ), this ratio can then be used to determine an upper bound on the uncertainty used in the establishment of the traceability of measurements made using the device at the end of the calibration chain [25].

---

<sup>1</sup> Note that this chain may only be one or two links.

<sup>2</sup> Although these multiples are frequently integers (e.g.  $k = 2$  or  $k = 3$ ), this is not necessary.

## 7 Taking measurement uncertainty into account when establishing maximum permissible errors (MPEs) and accuracy classes

Many OIML Recommendations, and some other OIML publications, specify MPEs that are to be used for particular tests. Establishing what values the MPEs should have usually involves a balance of considerations, including adequately protecting the consumer or user of the measuring instrument/system for reasons of cost and sometimes safety, but also protecting the manufacturer or distributor, again for reasons of cost. What is sometimes overlooked is consideration of the lowest level of measurement uncertainty that can be physically attained for the particular test, which sets a lower limit on the MPE that can be used. This should be taken into account when specifying an MPE for a particular test, or when establishing accuracy classes for a type of instrument, especially in cases where MPUs are specified.

For example, in testing cases where the expanded uncertainty  $U_{E_1}$  is known to typically be of a certain amount (and cannot easily be reduced), then the MPE corresponding to that test should be appropriately specified such that the ratio ( $f_{E_1} = U_{E_1}/\text{MPE}$ ) discussed in 5.3.4 can be kept acceptably low. In this case, since  $u_{E_1}$  cannot be reduced, it may become necessary to increase the MPE such that the condition illustrated by the rightmost curve in Figure 4 can be obtained.

Similarly for the measurement standard, if  $f_S (= U_S / \text{MPE})$  is typically too large for a given type of test, then the MPE might not be appropriate and so, if possible, specifying a larger MPE in the Recommendation might be necessary. If the MPE cannot be increased for other reasons, then it might be necessary to specify a type of measurement standard/system that has a lower measurement uncertainty ( $U_S$ ).

Related literature exists that may be consulted (e.g. [21][22]) when considering advice to include in Recommendations concerning the specification of appropriate MPEs and accuracy classes.

## **8 Options and suggested text pertaining to “measurement uncertainty” that should be considered for inclusion in OIML Recommendations and other OIML publications**

When deciding how to incorporate measurement uncertainty considerations into OIML Recommendations and other OIML publications the following possibilities should be considered and the appropriate ones determined. Suggested text is provided (in *italics*) that could be included in the OIML Recommendation or other OIML publication.

### **8.1 Incorporating measurement uncertainty for laboratory testing**

When the OIML Recommendation or publication involves the type evaluation or other testing in a laboratory of a measuring instrument or system, a clause shall be provided that emphasizes how measurement uncertainty can and should be incorporated into conformity decisions that are associated with the Recommendation (see clause 5). Suggested text (in *italics*):

“XX *Measurement uncertainty*

*The evaluation and use of measurement uncertainty have become important and essential elements in all aspects of metrology, including legal metrology. The OIML Guide G 19 on “The role of measurement uncertainty in conformity assessment decisions in legal metrology” should be consulted for a general understanding of the terminology and concepts related to measurement uncertainty, and for guidance on how to assess and use measurement uncertainty.*

*Measurement uncertainty shall be considered in all aspects of measurement and conformity assessment decisions associated with the type evaluation or other laboratory testing of this OIML Recommendation. Guidance is provided in xxx. Each test comprises measurements applying harmonized test setups for the verification of compliance with requirements. Measurement uncertainty is an attribute of each measurement. The uncertainty associated with a test method shall be taken into account in the decision on the applicability of the test method.*

*Measurement results that are reported during laboratory testing of a measuring instrument/system shall include a measured value along with its associated measurement uncertainty. Exceptions include those cases where individual measured values are obtained for the purpose of assessing a component of measurement uncertainty associated with the repeatability or reproducibility of the measuring instrument/system and/or testing procedure, in which case a measurement uncertainty is instead associated with the mean value of the individual measured values, or where it is determined that a component of measurement uncertainty is not significant in a particular measurement application (this should be so noted)”.*

## 8.2 Calculating measurement uncertainty

Individual OIML Recommendations should provide guidance, as appropriate, on calculating measurement uncertainty for the measurement model(s) appropriate to the type of instrument(s), testing systems and processes covered in the Recommendation (see examples in Annex C). Examples of such guidance are given in the seven steps below. In general, guidance should be provided on the following:

- (Step 1) Describe the instrument under test (IUT), along with the measuring system that will be used for performing the test(s). Include in the description all quantities that can affect the measuring instrument, all influence quantities that can affect the measuring instrument/system, and specify the conditions (if any) at which the (influence) quantities will be maintained during the testing, or the range(s) that the (influence) quantities shall remain within during the testing (e.g. rated operating conditions and/or reference operating conditions of both the measuring instrument/system and IUT).
- (Step 2) Identify all of the different kinds of tests that will need to be performed for the type evaluation and/or verification in the laboratory. Based on the description in Step 1, develop a mathematical model of the measurement (as in Equation 4.2) to be used for performing each kind of test. Each model shall ultimately provide an expression for the ‘error of indication’, and also include an expression for the standard measurement uncertainty to be associated with each measured error of indication (unless repeated measurements of error of indication are to be obtained, in which case the mean value of the error of indication is to be presented, along with an associated standard measurement uncertainty that incorporates a component obtained from the repeated measurements; see Step 5 below);
- (Step 3) Calculate the associated standard measurement uncertainty ( $u_s$ ) of the measurement standard or system;
- (Step 4) Calculate a standard measurement uncertainty ( $u_I$ ) associated with the indicated value of the measurand (including components due to indicator resolution and/or random fluctuation);
- (Step 5) Calculate a standard measurement uncertainty ( $u_{rep}$ ) associated with the repeatability or reproducibility of the measuring instrument/system and/or testing procedure;
- (Step 6) Calculate a standard measurement uncertainty ( $u_{roc}$ ) if the indication of the measuring instrument is found to vary when the instrument is operated over its range of rated operating conditions for a fixed input to the instrument. (Note that this component of measurement uncertainty is sometimes explicitly considered to be part of the measurement uncertainty of the error of indication, but is sometimes considered to be incorporated by convention into the MPE, so care should be taken in specifying what convention is being used [36]);

- (Step 7) Combine all of these components of measurement uncertainty (using the methodology of [1][2][3][4][5]) in order to calculate a combined standard measurement uncertainty ( $u_{E_1}$ ) associated with the error of indication. (See the example in Annex C).

Steps 2–7 can be summarized as:

- for each test, develop a model for that test which takes into account all the components of uncertainty;
- calculate the standard uncertainty for each component; and
- calculate the combined standard uncertainty for that test.

OIML Recommendations (and other OIML publications) should emphasize that the component of measurement uncertainty coming from the standard deviation of the individual measured values (Type A component) is not the entire measurement uncertainty, and that Type B components coming from steps 3–6 above shall also be included in the combined standard measurement uncertainty.

If they exist, include discussion of special or unusual aspects of assessing the components of measurement uncertainty (see also 8.8).

### **8.3 Specifying MPEs and MPUs**

For each kind of test identified above in 8.2, Step 2, the OIML Recommendation should discuss and specify what the appropriate MPE is for that kind of test. For example, for a type evaluation test, the MPE that is specified could correspond to one of several possible accuracy classes that the instrument is being tested for. For a verification test, the specified MPE could be based on a variety of considerations, as discussed in 7.

There should also be discussion of what the likely values of  $u_{E_1}$  and  $u_S$  will be during the test, in order to decide whether values of  $MPU_{E_1}$  and  $MPU_S$  should be specified and, if so, what those values should be (or, rather, what  $f_{E_1}$  and  $f_S$  should be; see 0, 5.3.5 and 7.)

### **8.4 Specifying acceptable levels of risk**

During the development of an OIML Recommendation or other OIML publication, it should be considered whether ‘acceptable’ levels of risk for various types of tests should be suggested. Decision rules and associated risks, along with their consequences, should be considered and discussed in OIML Recommendations or other OIML publications. However, this should be done only in the context of regulatory matters. Risks to a manufacturer may have serious economic consequences that are typically outside the scope of a Recommendation (see 7).

Depending on the values of  $MPU_{E_1}$  and  $MPU_S$  specified above in 8.3 (if any), discussion should be provided on whether the ‘shared risk’ principle is to be used (see 5.3.3), or whether there is a specified risk (probability) that is to be used and, if so, whether it is a risk of false acceptance (see 5.3.1) or a risk of false rejection (see 5.3.2). Note that if the ‘shared risk’

approach is used in an OIML Recommendation (or in other OIML publications), it should not be used in an implicit manner but, rather, an explicit statement of its use should be provided in the Recommendation.

### **8.5 Specifying uncertainty of error of indication when not using shared risk**

If risk of false acceptance or risk of false rejection is to be used, it is further necessary to specify whether  $u_{E_1}$  is to be considered as fixed for each measurement, in which case a guard band can be used for deciding conformity (see 5.3.6), or whether  $u_{E_1}$  is to be calculated separately for each error of indication, in which case the Standard Normal Distribution Table or Measurement Capability Index can be used each time. Reference to Annex B and Annex E of the present OIML Guide should be provided, along with possible additional discussion of how to use the Standard Normal Distribution Table and/or Measurement Capability Index for the particular Recommendation.

Constructing PDFs and calculating areas under a PDF curve is in general a nontrivial matter, and so any advice and assistance to be provided in this regard in the Recommendation(s) should be carefully considered during its development.

### **8.6 Complexity of assessing uncertainty of error of indication**

Assessing the measurement uncertainty of the error of indication for an individual measurement for a specified type of measuring instrument may be somewhat complex. It is important to note, however, that once all of the derivation has been performed, and values and associated measurement uncertainties are obtained for typical measurement conditions, the process of obtaining a value of  $u_{E_1}$  for each subsequent individual measurement performed during a given type evaluation or verification test in a laboratory should become relatively straightforward. Most components of measurement uncertainty will not change from one individual measurement to another. This aspect of the treatment of measurement uncertainty should be included in the discussion in each OIML Recommendation where measurement uncertainty is relevant. Suggested text (in *italics*):

*“Assessing the measurement uncertainty of the error of indication for an individual measurement for a specified type of measuring instrument may be somewhat complex. It is important to note, however, that once all of the derivation has been performed, and values and associated measurement uncertainties are obtained for typical measurement conditions, the process of obtaining a value of  $u_{E_1}$  for each subsequent individual measurement performed during a given type evaluation test should become relatively straightforward. Most components of measurement uncertainty will not change from one individual measurement to another. This can simplify the process of incorporating measurement uncertainty in situations outside of a laboratory environment, since guard bands or straightforward Measurement Capability Index tables can be used (e.g., see Annex E in the OIML Guide G 19 on “The role of measurement uncertainty in conformity assessment decisions in legal metrology”).”*

Alternatively, a reference to the guidance given in the present OIML Guide should be provided, with a Note to refer to Annex E.

### **8.7 Recording measurement uncertainty in OIML test reports for type evaluation**

For type evaluation, OIML Recommendations should provide for explicit entries in the Format of the Test Report document for recording measurement uncertainty (see 5), to accompany every measured value that is recorded (except when measurements for repeatability and/or reproducibility are being obtained). In those cases where measurement uncertainty can be assumed to be negligible, this should be documented with an appropriate notation, rather than leaving a blank entry. Also, if the 'Measurement Capability Index' ( $C_m$ ) method or the 'Guard band' method is to be used, this should also be noted in the Format of the Test Report, along with spaces for recording values of the appropriate parameters (e.g. the size of the guard band), along with the outcome of the test. A space for reference to where to find the  $C_m$  chart that was used should also be provided.

### **8.8 Providing guidance on using measurement uncertainty for periodic verification**

OIML Recommendations should provide guidance on how to treat measurement uncertainty at the stage of periodic verification testing, emphasizing any differences, precautions and/or special considerations from the guidance provided for type evaluation testing. For example, for a given type of measuring instrument, it might be recommended to include some sources of measurement uncertainty during type evaluation, whereas those sources of uncertainty may not be as significant in cases where the MPEs are increased, e.g. during periodic verification testing of certain instruments. Also, the use of guard bands might be recommended for type evaluation, whereas shared risk might be acceptable for a periodic verification test.

## 9 References

- [1] JCGM 100:2008 (GUM 1995 with minor corrections) - BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, *Evaluation of measurement data - Guide to the expression of uncertainty in measurement* (Published by the OIML as OIML G 1-100:2008 at [www.oiml.org/en/publications/guides/](http://www.oiml.org/en/publications/guides/))
- [2] JCGM 101:2008 - BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, *Evaluation of measurement data - Supplement 1 to the "Guide to the expression of uncertainty in measurement" - Propagation of distributions using a Monte Carlo method* (Published by the OIML as OIML G 1-101:2008 at [www.oiml.org/en/publications/guides/](http://www.oiml.org/en/publications/guides/))
- [3] JCGM 102:2011 - BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, *Evaluation of measurement data - Supplement 2 to the "Guide to the expression of uncertainty in measurement" - Extension to any number of output quantities* - (Published by the OIML as OIML G 1-102:2008 at [www.oiml.org/en/publications/guides/](http://www.oiml.org/en/publications/guides/))
- [4] JCGM 104:2009 - BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, *Evaluation of measurement data - An introduction to the "Guide to the expression of uncertainty in measurement" and related documents*, (Published by the OIML as OIML G 1-104:2011 at [www.oiml.org/en/publications/guides/](http://www.oiml.org/en/publications/guides/))
- [5] JCGM 106:2012 - BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, *Evaluation of measurement data - The role of measurement uncertainty in conformity assessment* (Published by the OIML as OIML G 1-106:2012 at [www.oiml.org/en/publications/guides/](http://www.oiml.org/en/publications/guides/))
- [6] OIML B 18:2016, *Framework for the OIML Certification System (OIML-CS)*, ([www.oiml.org/en/publications/basic](http://www.oiml.org/en/publications/basic))
- [7] OIML B 3:2013, *OIML Basic Certificate System for OIML Type Evaluation of Measuring Instruments* ([www.oiml.org/en/publications/basic](http://www.oiml.org/en/publications/basic))
- [8] OIML B 10:2011, amended 2012, *Framework for a Mutual Acceptance Arrangement on OIML Type Evaluations (MAA)*, ([www.oiml.org/en/publications/basic](http://www.oiml.org/en/publications/basic))
- [9] ISO/IEC 17025:2005, *General requirements for the competence of testing and calibration laboratories*
- [10] OIML D 30:2008, *Guide for the application of ISO/IEC 17025 to the assessment of Testing Laboratories involved in legal metrology* ([www.oiml.org/en/publications/documents](http://www.oiml.org/en/publications/documents))
- [11] OIML R 87:2016, *Quantity of product in prepackages*, ([www.oiml.org/en/publications/recommendations](http://www.oiml.org/en/publications/recommendations))

- [12] OIML R 79:2015, *Labeling requirements for prepackages*, ([www.oiml.org/en/publications/recommendations](http://www.oiml.org/en/publications/recommendations))
- [13] OIML R 34:1979, *Accuracy classes of measuring instruments*, ([www.oiml.org/en/publications/recommendations](http://www.oiml.org/en/publications/recommendations))
- [14] JCGM 200:2012 - BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM)* - (JCGM 200:2008 with minor corrections) (Published by the OIML as OIML V 2-200:2012 at [www.oiml.org/en/publications/guides/](http://www.oiml.org/en/publications/guides/)) Also referred to in the present Guide as “VIM3”.
- [15] OIML V 1:2013 *International vocabulary of terms in legal metrology (VIML) (bilingual French-English) / Vocabulaire international des termes de métrologie légale (VIML) (bilingue français-anglais)*, ([www.oiml.org/en/publications/vocabularies/](http://www.oiml.org/en/publications/vocabularies/))
- [16] Ehrlich, C., Dybkaer, R., and Woger, W., *Evolution of philosophy and description of measurement (preliminary rationale for VIM3)*, OIML Bulletin, April 2007, pp. 23–35 ([www.oiml.org/en/publications/bulletin](http://www.oiml.org/en/publications/bulletin))
- [17] Sommer, K.-D. and Kochsiek, M., *Role of measurement uncertainty in deciding conformance in legal metrology*, OIML Bulletin, Volume XLIII, Number 2, April 2002, pp. 19–24 ([www.oiml.org/en/publications/bulletin](http://www.oiml.org/en/publications/bulletin))
- [18] Ehrlich, C. and Rasberry, S., *Metrological Timelines in Traceability*, Metrologia, 1997, 34, pp. 503–514
- [19] [www.itl.nist.gov/div898/handbook/eda/section3/eda3671.htm](http://www.itl.nist.gov/div898/handbook/eda/section3/eda3671.htm)
- [20] Kallgren, H. and Pendrill, L., *Uncertainty in conformity assessment in legal metrology (related to the MID)*, OIML Bulletin, Volume XLVII, Number 3, July 2006, pp. 15–21 ([www.oiml.org/en/publications/bulletin](http://www.oiml.org/en/publications/bulletin))
- [21] WELMEC Guide 4.2, Elements for deciding the appropriate level of confidence in regulated measurements, Issue 1, June 2006
- [22] The Expression of Uncertainty in EMC Testing, United Kingdom Accreditation Service (LAB34), Edition 1, August 2002
- [23] Deaver, D., An Application of the Guide to Measurement Uncertainty, 2000, Proceedings of the Measurement Science Conference (US).
- [24] Philips, Steven, Eberhardt, K.R., Estler, W.T., Measurement Uncertainty and Uncorrected Bias, Proceedings of the NCSL Workshop & Symposium, 1999, pp. 831–849

- 
- [25] Guthrie, William, Easy Translation of TAR or TUR into Uncertainty. NCSLI 2013 - Session 6B:  
([www.nist.gov/itl/sed/pubarchives/loader.cfm?csModule=security/getfile&pageid=3204506](http://www.nist.gov/itl/sed/pubarchives/loader.cfm?csModule=security/getfile&pageid=3204506))

Additional references/reading :

- [26] OIML D 8 *Measurement standards. Choice, recognition, use, conservation and documentation* ([www.oiml.org/en/publications/documents/](http://www.oiml.org/en/publications/documents/))
- [27] OIML D 18 *The use of certified reference materials in fields covered by metrological control exercised by national services of legal metrology. Basic principles* ([www.oiml.org/en/publications/documents/](http://www.oiml.org/en/publications/documents/))
- [28] OIML D 23 *Principles for metrological control of equipment used for verification* ([www.oiml.org/en/publications/documents/](http://www.oiml.org/en/publications/documents/))
- [29] EA-4/02 M:2013 *Evaluation of the Uncertainty of Measurement In Calibration*
- [30] ILAC P-14 *ILAC Policy for Uncertainty in Calibration*
- [31] ILAC G-8 *Guidelines on the Reporting of Compliance with Specification*
- [32] L R Pendrill 2014, “Using measurement uncertainty in decision-making & conformity assessment”, *Metrologia* **51** S206, doi:10.1088/0026-1394/51/4/S206
- [33] H Karlsson, Å Olsen, L R Pendrill 2016, “Conformance assessment of electrical energy meters investigated by risk analysis – a case study”, *OIML Bulletin* **LVI I** (no. 2 April), 5 – 20 ([www.oiml.org/en/publications/bulletin](http://www.oiml.org/en/publications/bulletin))
- [34] I Hendrikx, B Jovanoski, N Tuneski 2016, “Dynamic simulations of market surveillance actions”, IEEE Product Safety conference, May 2016, California (US)
- [35] Ehrlich, Charles, Terminological aspects of the Guide to the expression of Uncertainty in Measurement (GUM), *Metrologia*, 51 (2014) S145-S154
- [36] Ehrlich, Charles D., Traceability Considerations for the Characterization and Use of Measuring Systems, NCSLI Measure J. Meas. Sci., Vol. 10, No. 4, December 2015

## **Annex A**

### **Coexistence of “measurement error” and “measurement uncertainty” in legal metrology (relationship between calibration and verification)**

The introduction of the 1993 edition of the *Guide to the Expression of Uncertainty in Measurement* (also referred to as the GUM) opened a new way of thinking about both measurement and about expressing the perceived quality of the result of a measurement. Rather than express the result of a measurement by providing a best-estimate of the true value of the quantity being measured, along with information about known systematic and random errors, the GUM provided an alternative approach whereby the result of a measurement is expressed as a best-estimate of the *essentially-unique* true value (denoted hereafter as ‘true’ value) of the quantity *intended to be* measured (the ‘measurand’), along with an associated ‘measurement uncertainty’. (Note that, historically, in legal metrology the term “true value” is sometimes used to mean the value associated with a measurement standard that is used in the process of verifying a measuring instrument. This is not the meaning of the term in this OIML Guide).

The concept of measurement uncertainty can be described as a measure of how well the ‘true’ value of the measurand is believed to be known. (Note that according to the GUM approach it is not possible to know how well the ‘true’ value of the measurand is known, but only how well it is believed to be known). The notion of ‘belief’ is an important one, since it moves metrology (and legal metrology) into a realm where results of measurements must be considered and expressed (sometimes only implicitly) in terms of probabilities or degrees of belief. When making decisions in legal metrology about whether measuring systems are performing according to specified requirements, if the GUM approach is to be followed it becomes necessary to make such decisions on a probabilistic basis. This OIML Guide provides guidance on how to incorporate the GUM approach and take into account the concepts of measurement uncertainty and probability when making such conformity assessment decisions.

Legal metrology is the process and the practice of applying regulatory structure and enforcement to metrology, which is the science and application of measurement. Much of legal metrology involves testing or verifying measuring instrument/system design and use, in both laboratory and environments outside of a laboratory, to ensure that credible measurements can be, and are being, made when using the instrument/system in regulated situations. Testing or verifying in this context means that a decision is being made about whether the measuring system under test is providing indicated values of a quantity being measured that are believed to be ‘close enough’ to the ‘true’ value, as determined by using measurement standards, for the regulatory purpose at hand. The close enough conditions are specified in regulations, usually in terms of ‘maximum permissible errors’ (MPEs) or ‘accuracy classes’. Using the GUM approach, the objective of verification then becomes to determine the degree of belief (level of confidence) that the ‘true’ value of the ‘error of

---

indication' lies within the maximum permissible errors when taking measurement uncertainty (of the 'error of indication'!) into account.

Using the concepts of 'measurement error' and 'measurement uncertainty' at the same time like this may at first glance seem inconsistent or confusing. The GUM seemingly discourages use of the concept of measurement error in favor of measurement uncertainty. However, it must be kept in mind that the focus of the GUM is on using calibrated measuring instruments to perform measurements, and not on testing or verifying measuring instruments themselves. From the GUM perspective, known measurement errors that arise when using a measuring instrument are to be 'corrected for', so that no known (systematic) measurement error remains. (Methods for treating known systematic error or bias in this context exist, however. See, for example, references [23] and [24]). By contrast, in the context of verification in legal metrology (as well as in some other areas of metrology), error is used to assess the performance of a measuring instrument (and is not corrected for), and error (or, actually, error of indication) can in fact be considered to be a perfectly reasonable value to assess, along with its associated uncertainty. This approach to use of the term 'error' is the approach that is taken in this OIML Guide.

As already indicated, conformity testing in legal metrology typically involves comparing the measured error of indication of a measuring instrument or system to an MPE that is specified in a legal regulation. The error of indication is typically calculated in legal metrology as the difference between the indicated value and a value as given by a measurement standard. It is known that the value as given by the measurement standard is very likely not the 'true' value of the quantity being measured, but it is typically thought to be very close for a given situation. However, since the 'error of indication' is actually meant to be the difference between the indicated value and the 'true' value of the measurement standard, the uncertainty associated with the value given by the measurement standard (such as is stated in its calibration certificate) must be taken into consideration when making a conformity assessment decision. This will be elaborated on below.

By utilizing a first-principles approach that incorporates a simple example involving a mass standard and a weighing instrument to be verified, this Annex will now elaborate on how measurement error and measurement uncertainty can coexist when considering measurement in the context of verification.

As in clause 3 of the GUM, the initial focus of this Annex will be to consider measurement error and measurement uncertainty from the perspective of describing the objective of measurement. The terminology used to do this will be that of the VIM3 [10], which in some cases is somewhat different to that of the GUM, for reasons that will be explained when necessary. Several relevant definitions from the VIM3 are provided in clause 2 of this OIML Guide.

The objective of a measurement can be thought of as developing, through some type of 'experiment', a quantitative expression about the 'measurand'. The expression usually involves the concept and term 'value' ('quantity value' in VIM3), which is a number and

reference that together express the magnitude of a ‘quantity’. The reference is typically a measurement ‘unit’, which is adopted by convention such that other quantities of the same kind can be compared to it.

Prior to the concept of measurement uncertainty, the objective of measurement was to obtain a measurement result that was typically expressed as a best estimate of the ‘true’ value of the measurand and was sometimes accompanied by an ‘error analysis’ that contained any systematic errors (that were to be ‘corrected’ for when calculating the best estimate) and a description of the ‘spread’ of the random errors (if more than one observation was made) that occurred during the measurement. The concept of metrological traceability was used for expressing the measurement result in terms of an appropriate measurement unit by establishing a chain of comparisons or calibrations to a realization of the measurement unit. Besides stating possible systematic errors associated with the traceability chain, nothing further was typically stated about other possible sources of systematic error.

As discussed earlier, the concept of measurement uncertainty fundamentally changed the way that metrologists think about the objective of measurement. Most notably, one of the basic premises of the GUM approach is that it is possible to characterize the quality of a measurement by accounting for both random and systematic ‘effects’ on an equal footing, thus refining the information previously provided in an error analysis, and putting it on a probabilistic basis. Rather than express a measurement result as a best estimate of the ‘true’ value of the measurand, along with an error analysis, a measurement result is instead to be expressed as a best estimate of the ‘true’ value of the measurand along with a measurement uncertainty, which is a measure of how well the stated best estimate is believed to be known (based on the measured data and other knowledge, typically relating to systematic effects, and on the assumption that no mistakes were made when performing the measurement).

The probabilistic basis of the GUM approach derives primarily from another basic premise of the GUM (3.3.1), which is that it is not possible to know the true value of a 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”. This is a very fundamental and important point to keep in mind. Another related consideration, discussed in D.3.4 of the GUM, is that there is no such thing as a unique true value of a measurand, since at some level there is always an ‘intrinsic’ uncertainty due to the necessarily incomplete definition of the measurand (VIM3 refers to this as “definitional uncertainty”). Clause 1.2 of the GUM elaborates that, therefore, it is not possible to have a unique, true value of a measurand, but rather that it is only possible to have an “essentially unique” true value, which, as mentioned earlier, for shorthand has been referred to in this Guide as a ‘true’ value.

Note that the Note in 3.1.1 of the GUM explains why the GUM views the terms “value of a measurand” and “true value of a measurand” to be “equivalent”, and so uses only the term “value” when what is meant is the concept of ‘true’ value (as it is defined in B.2.3 of the GUM), namely, a value consistent with the definition of the measurand. The VIM3 [10] and this Guide do not adopt this GUM convention, and utilize the term “true value” when that

concept is what is intended, since the term “value” is already used in the more general sense given above. It is otherwise confusing to use the single term “value” for two different concepts [12].

Besides the concept of “error”, another concept (and term) that is discouraged in the GUM, at least in a quantitative sense, is “accuracy”. This is because “accuracy” is typically thought of in the inverse sense as “error”, in that the larger the error, the lower the accuracy. Since “error” cannot be known in the GUM sense, neither can “accuracy”. Therefore, care should be taken in OIML Recommendations to be sensitive to how the term “accuracy” is used, both in connection with “accuracy classes” as well as in the general sense. Accuracy classes are intended to convey information about what level of MPE a measuring instrument that meets a specified accuracy class is capable of achieving.

Metrological traceability continues to be a very important concept in the uncertainty (GUM) approach to measurement, and in fact takes on an additional aspect that links it very closely to the concept of measurement uncertainty. Besides serving as the basis for establishing a chain of comparisons or calibrations back to the measurement unit so as to be able to express the ‘measured value’ in terms of a measurement unit, the concept of metrological traceability is also used to be able to track the progression of measurement uncertainty along the traceability chain. In this regard, metrological traceability and measurement uncertainty are inextricably linked [14], as explicitly evidenced in the VIM3 (and VIM2) definition of metrological traceability.

## A.1 Calibration

The concepts of ‘measurement unit’, ‘true value’, ‘measurement error’ and ‘standard measurement uncertainty’ are illustrated in Figure A1, in the context of measuring (calibrating) a standard weight, which is shown schematically at the top right. It is assumed that the weight is calibrated using a high quality measuring system that is not otherwise mentioned or shown. The calibration certificate of the standard weight contains the measured mass value ( $M_{\text{calibrated}}$ ) of the standard weight, along with the associated standard measurement uncertainty ( $u_{\text{calibrated}}$ ). The standard measurement uncertainty (or the expanded uncertainty,  $U_{\text{calibrated}}$ ) is obtained during the calibration of the standard weight, through the use of the traceability principle, back to the measurement unit shown on the horizontal axis of the figure. The ‘true’ value of the mass of the standard weight is also indicated in the figure, both at the top right and on the horizontal axis, where it is indicated that it exists, but is unknowable in principle. The small vertical bars around the ‘true’ value of the mass of the standard weight on the horizontal axis are intended to denote the definitional uncertainty associated with the ‘true’ value.

Also shown in Figure A1 is a probability density function (PDF) which, as described in 5.1, provides probability densities that the ‘true’ value of the mass of the standard weight lies within an infinitesimal region around a particular possible ‘true’ value of the mass of the standard weight. The standard measurement uncertainty ( $u_{\text{calibrated}}$ ) is obtained from the PDF, usually as the standard deviation, as indicated.

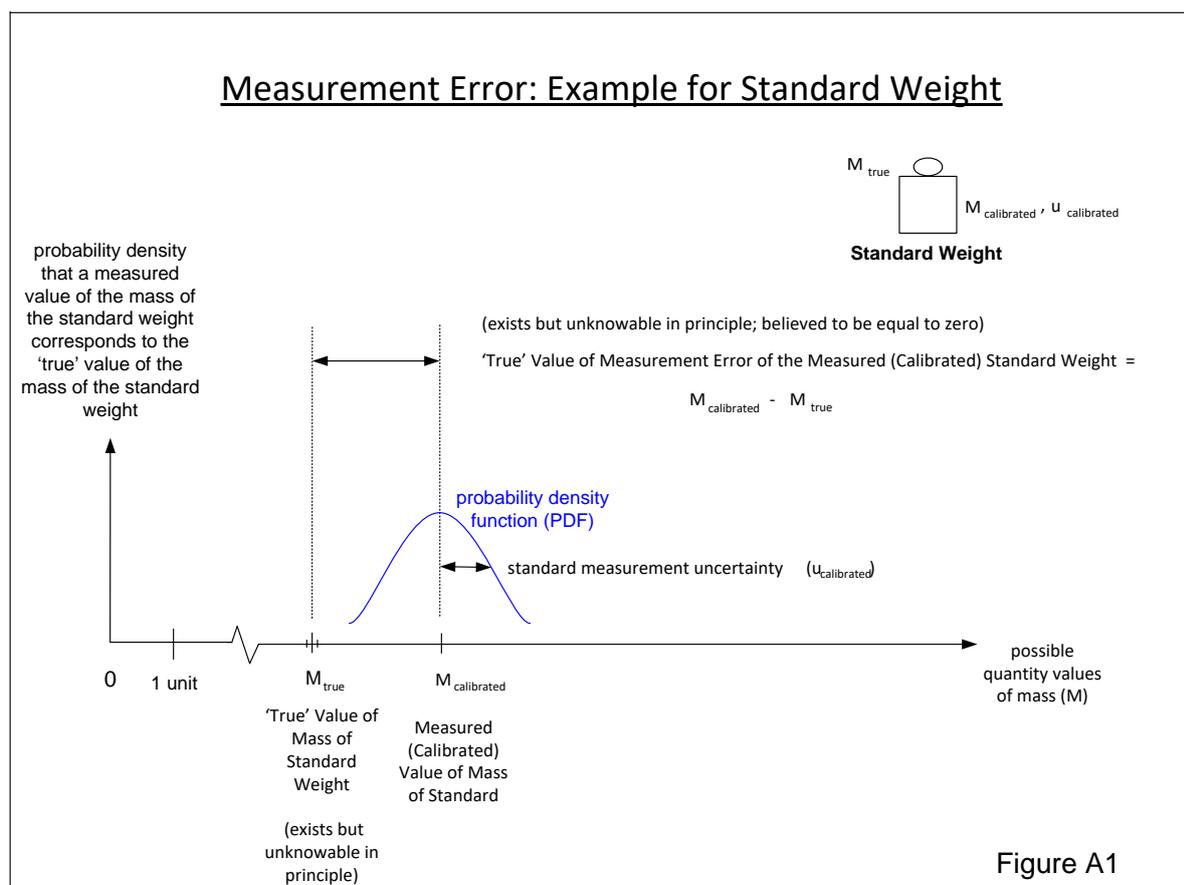


Figure A1 also illustrates the 'true' value of the 'measurement error' of the mass of the standard weight, defined as the difference between the measured (calibrated) value of the mass of the standard weight and the 'true' value of the mass of the standard weight. An important point to note in Figure A1 is that this error is considered as unknowable, since the 'true' value of the mass of the standard weight is unknowable. The GUM discourages use of the concept of error in some contexts since it is 'unknowable' in this measurement context, and instead favors use of measurement uncertainty, since measurement uncertainty can be calculated, and gives a measure of how well one believes one knows the 'true' value of the mass of the standard weight. It is very important to keep in mind that, in the context of measurement, despite the possible reality illustrated in Figure A1, the 'true' value of the measurement error of the measured (calibrated) mass of the measurement standard is believed to be zero, based on all of the available information from the measurement (calibration), since corrections are to be applied for all known systematic errors.

## A.2 Verification

Now consider the situation where the calibrated standard weight is used for the purpose of verifying, not calibrating, a weighing instrument, as illustrated in Figure A2. In a verification

test, indicated values of a quantity being measured when verifying a measuring instrument under test are compared with calibrated values (of the same quantity) as obtained when using a measurement standard.

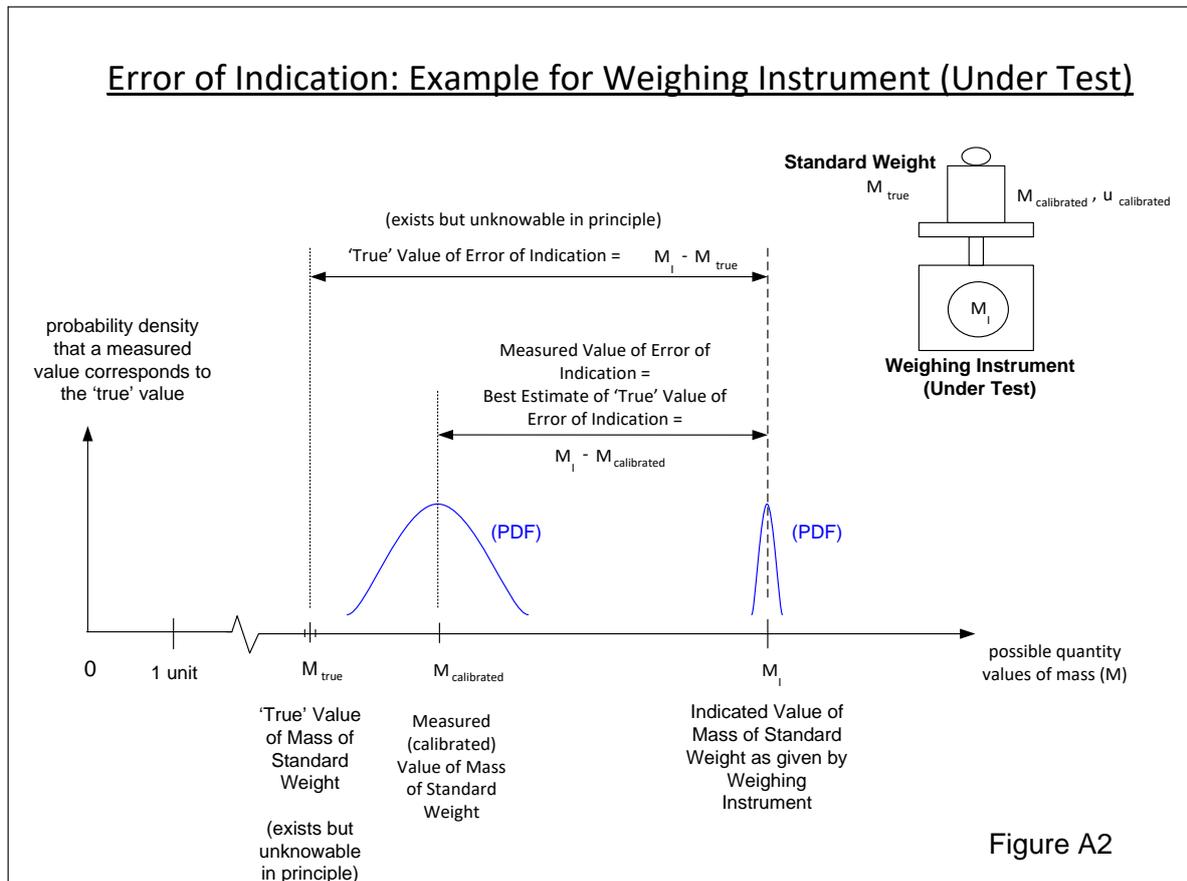


Figure A2 contains much of the same information as Figure A1, but in addition shows the value ( $M_I$ ) of the indication of the mass of the standard weight as obtained from the weighing instrument under test. Two ‘errors of indication’ are also shown, one with respect to the ‘true’ value of the mass of the standard weight (which is still unknowable), and another with respect to the measured (calibrated) value of the mass of the standard weight (which is known). As noted in Figure A2, the measured value of the error of indication is taken as the ‘best estimate’ of the ‘true’ value of the error of indication since, as discussed above, the ‘true’ value of the error of the measured (calibrated) mass of the measurement standard (standard weight) is believed to be zero.

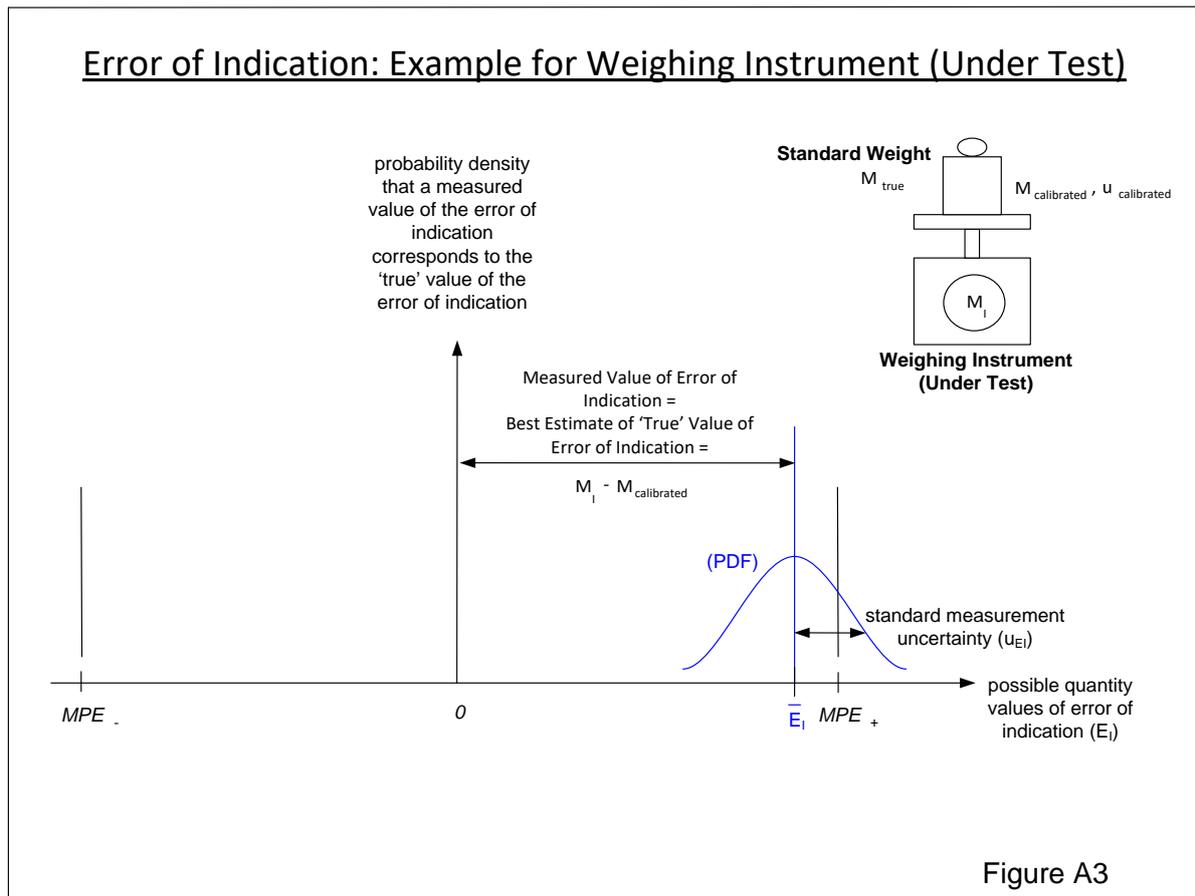
Verification testing is frequently performed in both ‘laboratory’ and environments outside of a laboratory. In a verification testing scenario, the objective is not to ‘correct’ or ‘adjust’ the indicated value to the measured (calibrated) value of the mass standard, but rather to assess whether the difference (error of indication) between the indicated value and the calibrated value of the mass standard is within acceptable limits of maximum permissible errors (MPEs, see clause 4), as expressed in a regulation (e.g. in an OIML Recommendation). While it is

highly desirable that the error of indication be small (and even zero), this is typically not the case in verification testing.

For type evaluation testing in a controlled laboratory environment, it is common to include tests where quantities that influence the indicated value of the measuring instrument (so-called influence quantities, such as ambient temperature and humidity) are varied in a controlled manner while everything else remains the same (including the quantity being measured, which in this example is the standard weight). The allowed variation in the error of indication under such conditions is either expressed in the OIML Recommendation or left to national regulation. When assessing whether such an influence quantity test passes, it is important to take into account the measurement uncertainty associated with the measurement of the influence quantity.

Also shown in Figure A2 are two PDFs. The one on the left is the same PDF as shown in Figure A1. The one on the right is for the indicated value of the mass of the standard weight (sources of this uncertainty could come from instability (jitter) of the indicated value, finite resolution of the indicator, and other random effects that generally contribute to lack of repeatability when obtaining multiple values of error of indication). Note that the width of the PDF for  $M_I$  is shown in the figure to be significantly smaller than the width of the PDF for  $M_{\text{calibrated}}$  (say, if the resolution and repeatability of the indicated mass value is quite small compared to the uncertainty of the calibrated mass), but this is not necessarily always true. (Note also that the two PDFs are not shown to scale with respect to each other). What is desired is to use the information in these two PDFs to be able to make a statement about how well the ‘true’ value of the error of indication is believed to be known. This is illustrated in Figure A3.

Note that the horizontal axis in Figure A3 is now changed from that in Figures A1 and A2, and is labeled ‘possible quantity values of error of indication’. The magnitude of the measured value of the error of indication is the same as is given in Figure A2 and, as discussed earlier, is the best estimate of the ‘true’ value of the error of indication. A PDF can be constructed giving the probability density that the ‘true’ value of the error of indication lies within an infinitesimal region around a particular possible ‘true’ value of the error of indication. Such a PDF is illustrated in Figure A3, along with the associated standard measurement uncertainty ( $u_{E_I}$ ). This PDF is obtained by combining (sometimes called convolving) the two PDFs in Figure A2 [2]. It is interesting to note that  $u_{E_I}$  is the ‘standard uncertainty of the error (of indication)’, which explicitly demonstrates the coexistence of the terms and concepts ‘uncertainty’ and ‘error’ in a testing scenario.



### A.3 Brief summary

In summary, while the concept of 'measurement uncertainty' was developed to replace the concept of 'measurement error' and 'error analysis' in the context of performing measurements, the term and concept of 'error' remains useful in the context of verifying measuring instruments and systems. In fact, it makes sense to talk about the uncertainty of a measured error of indication! The measurement uncertainty associated with the measurement standard(s) used when performing the verification test must be taken into account when making (probabilistic) conformity assessment decisions, since they contribute to the standard measurement uncertainty of the error of indication ( $u_{E_I}$ ).

## Annex B

### Use of the standard normal distribution table (Z-table)

This Annex is an adaptation from the NIST/SEMATECH e-Handbook of Statistical Methods (<http://www.itl.nist.gov/div898/handbook/eda/section3/eda3671.htm>).

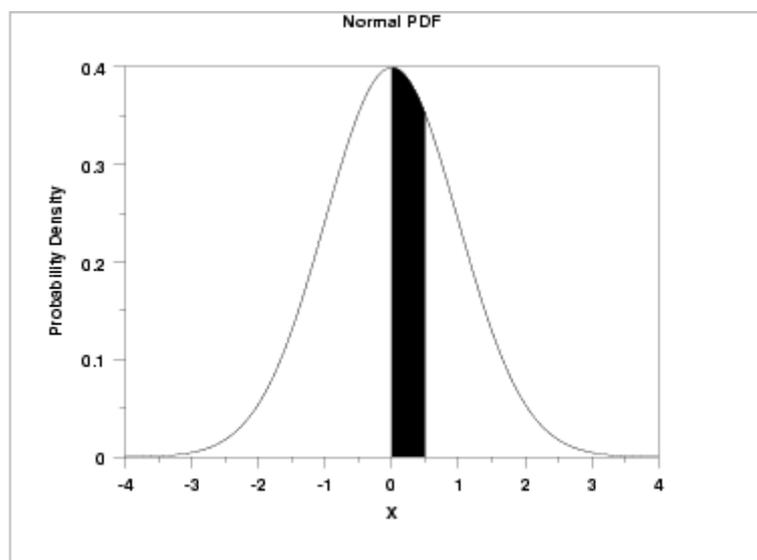
The general formula for the [probability density function](#) of the normal distribution is

$$f(x) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$$

Where  $\mu$  is the [location parameter](#) and  $\sigma$  is the [scale parameter](#). The case where  $\mu = 0$  and  $\sigma = 1$  is called the **standard normal distribution**. The equation for the standard normal distribution is

$$f(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

The figure below illustrates the standard normal distribution (sometimes also referred to as a normalized Gaussian distribution). The shaded area under the curve represents the probability that the parameter  $x$  is between 0 and  $\alpha$  ( $\alpha = 0.5$  in the figure).



Values of areas under the curve for discrete values of  $\alpha$  can be obtained from the standard normal distribution table below.

### Standard normal distribution table

The table below contains the area under the standard normal curve from  $x = 0$  to a specified value  $x = \alpha$ .

Area under the Normal Curve from  $X = 0$  to  $X = \alpha$

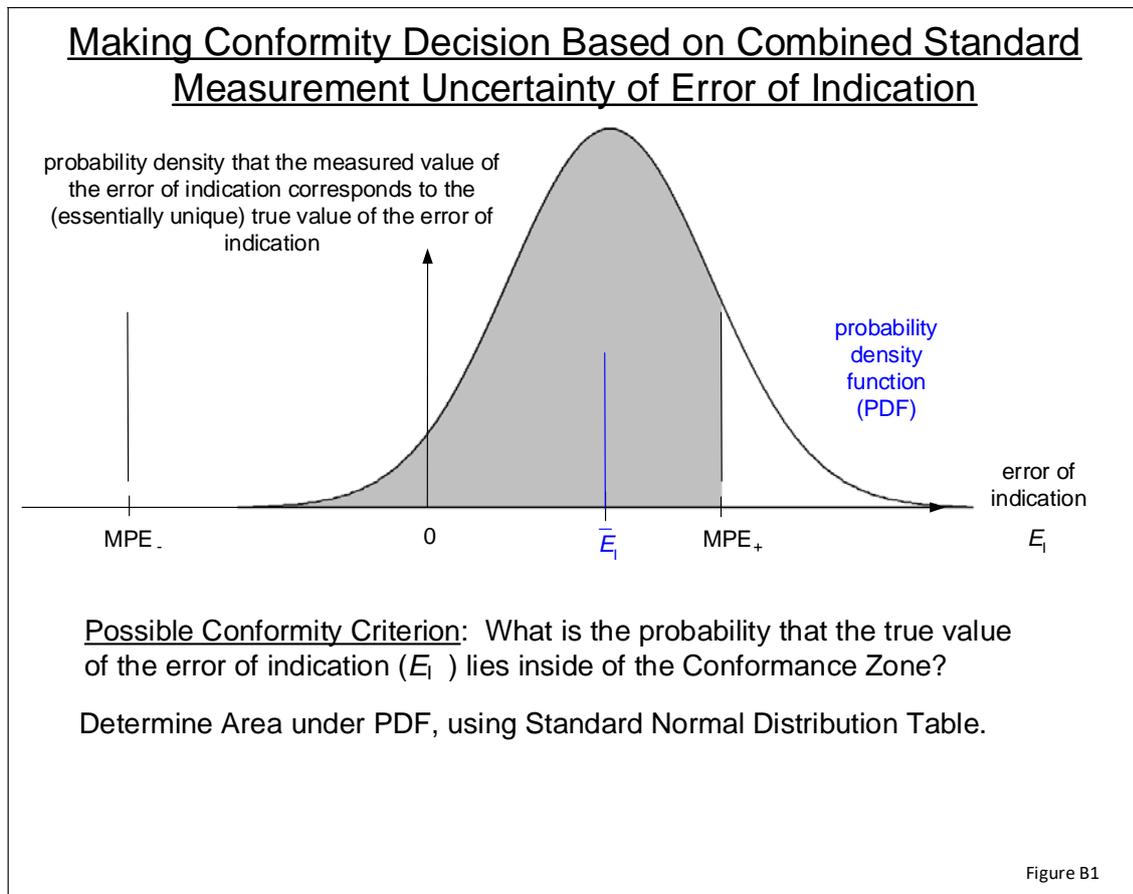
$\alpha$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.00000	0.00399	0.00798	0.01197	0.01595	0.01994	0.02392	0.02790	0.03188	0.03586
0.1	0.03983	0.04380	0.04776	0.05172	0.05567	0.05962	0.06356	0.06749	0.07142	0.07535
0.2	0.07926	0.08317	0.08706	0.09095	0.09483	0.09871	0.10257	0.10642	0.11026	0.11409
0.3	0.11791	0.12172	0.12552	0.12930	0.13307	0.13683	0.14058	0.14431	0.14803	0.15173
0.4	0.15542	0.15910	0.16276	0.16640	0.17003	0.17364	0.17724	0.18082	0.18439	0.18793
0.5	0.19146	0.19497	0.19847	0.20194	0.20540	0.20884	0.21226	0.21566	0.21904	0.22240
0.6	0.22575	0.22907	0.23237	0.23565	0.23891	0.24215	0.24537	0.24857	0.25175	0.25490
0.7	0.25804	0.26115	0.26424	0.26730	0.27035	0.27337	0.27637	0.27935	0.28230	0.28524
0.8	0.28814	0.29103	0.29389	0.29673	0.29955	0.30234	0.30511	0.30785	0.31057	0.31327
0.9	0.31594	0.31859	0.32121	0.32381	0.32639	0.32894	0.33147	0.33398	0.33646	0.33891
1.0	0.34134	0.34375	0.34614	0.34849	0.35083	0.35314	0.35543	0.35769	0.35993	0.36214
1.1	0.36433	0.36650	0.36864	0.37076	0.37286	0.37493	0.37698	0.37900	0.38100	0.38298
1.2	0.38493	0.38686	0.38877	0.39065	0.39251	0.39435	0.39617	0.39796	0.39973	0.40147
1.3	0.40320	0.40490	0.40658	0.40824	0.40988	0.41149	0.41308	0.41466	0.41621	0.41774
1.4	0.41924	0.42073	0.42220	0.42364	0.42507	0.42647	0.42785	0.42922	0.43056	0.43189
1.5	0.43319	0.43448	0.43574	0.43699	0.43822	0.43943	0.44062	0.44179	0.44295	0.44408
1.6	0.44520	0.44630	0.44738	0.44845	0.44950	0.45053	0.45154	0.45254	0.45352	0.45449
1.7	0.45543	0.45637	0.45728	0.45818	0.45907	0.45994	0.46080	0.46164	0.46246	0.46327
1.8	0.46407	0.46485	0.46562	0.46638	0.46712	0.46784	0.46856	0.46926	0.46995	0.47062
1.9	0.47128	0.47193	0.47257	0.47320	0.47381	0.47441	0.47500	0.47558	0.47615	0.47670
2.0	0.47725	0.47778	0.47831	0.47882	0.47932	0.47982	0.48030	0.48077	0.48124	0.48169
2.1	0.48214	0.48257	0.48300	0.48341	0.48382	0.48422	0.48461	0.48500	0.48537	0.48574
2.2	0.48610	0.48645	0.48679	0.48713	0.48745	0.48778	0.48809	0.48840	0.48870	0.48899
2.3	0.48928	0.48956	0.48983	0.49010	0.49036	0.49061	0.49086	0.49111	0.49134	0.49158
2.4	0.49180	0.49202	0.49224	0.49245	0.49266	0.49286	0.49305	0.49324	0.49343	0.49361
2.5	0.49379	0.49396	0.49413	0.49430	0.49446	0.49461	0.49477	0.49492	0.49506	0.49520
2.6	0.49534	0.49547	0.49560	0.49573	0.49585	0.49598	0.49609	0.49621	0.49632	0.49643
2.7	0.49653	0.49664	0.49674	0.49683	0.49693	0.49702	0.49711	0.49720	0.49728	0.49736
2.8	0.49744	0.49752	0.49760	0.49767	0.49774	0.49781	0.49788	0.49795	0.49801	0.49807
2.9	0.49813	0.49819	0.49825	0.49831	0.49836	0.49841	0.49846	0.49851	0.49856	0.49861
3.0	0.49865	0.49869	0.49874	0.49878	0.49882	0.49886	0.49889	0.49893	0.49896	0.49900
3.1	0.49903	0.49906	0.49910	0.49913	0.49916	0.49918	0.49921	0.49924	0.49926	0.49929
3.2	0.49931	0.49934	0.49936	0.49938	0.49940	0.49942	0.49944	0.49946	0.49948	0.49950
3.3	0.49952	0.49953	0.49955	0.49957	0.49958	0.49960	0.49961	0.49962	0.49964	0.49965
3.4	0.49966	0.49968	0.49969	0.49970	0.49971	0.49972	0.49973	0.49974	0.49975	0.49976
3.5	0.49977	0.49978	0.49978	0.49979	0.49980	0.49981	0.49981	0.49982	0.49983	0.49983
3.6	0.49984	0.49985	0.49985	0.49986	0.49986	0.49987	0.49987	0.49988	0.49988	0.49989
3.7	0.49989	0.49990	0.49990	0.49990	0.49991	0.49991	0.49992	0.49992	0.49992	0.49992
3.8	0.49993	0.49993	0.49993	0.49994	0.49994	0.49994	0.49994	0.49995	0.49995	0.49995
3.9	0.49995	0.49995	0.49996	0.49996	0.49996	0.49996	0.49996	0.49996	0.49997	0.49997

In this Guide,  $\alpha$  is defined as

$$\alpha = [(MPE_+ - \bar{E}_I) / u_{E_I}] \quad (\text{B.1})$$

for the case where  $\bar{E}_I > 0$ . The case where  $\bar{E}_I < 0$  is discussed later.

Figure B1 illustrates the relevant parameters.



### Example

Consider an individual test of a length measuring instrument, such as a line measure, where the indicated value of length ( $L_I$ ) is 1.0006 m when the value of the reference length of a high-precision line measure ( $L_R$ ), as obtained from its calibration certificate, is 1.0003 m. The measured value of the error of indication is then:

$$E_I = L_I - L_R = 0.0003 \text{ m} = 300 \mu\text{m} \quad (\text{B.2})$$

Say that an evaluation of the standard uncertainty of the error of indication gives

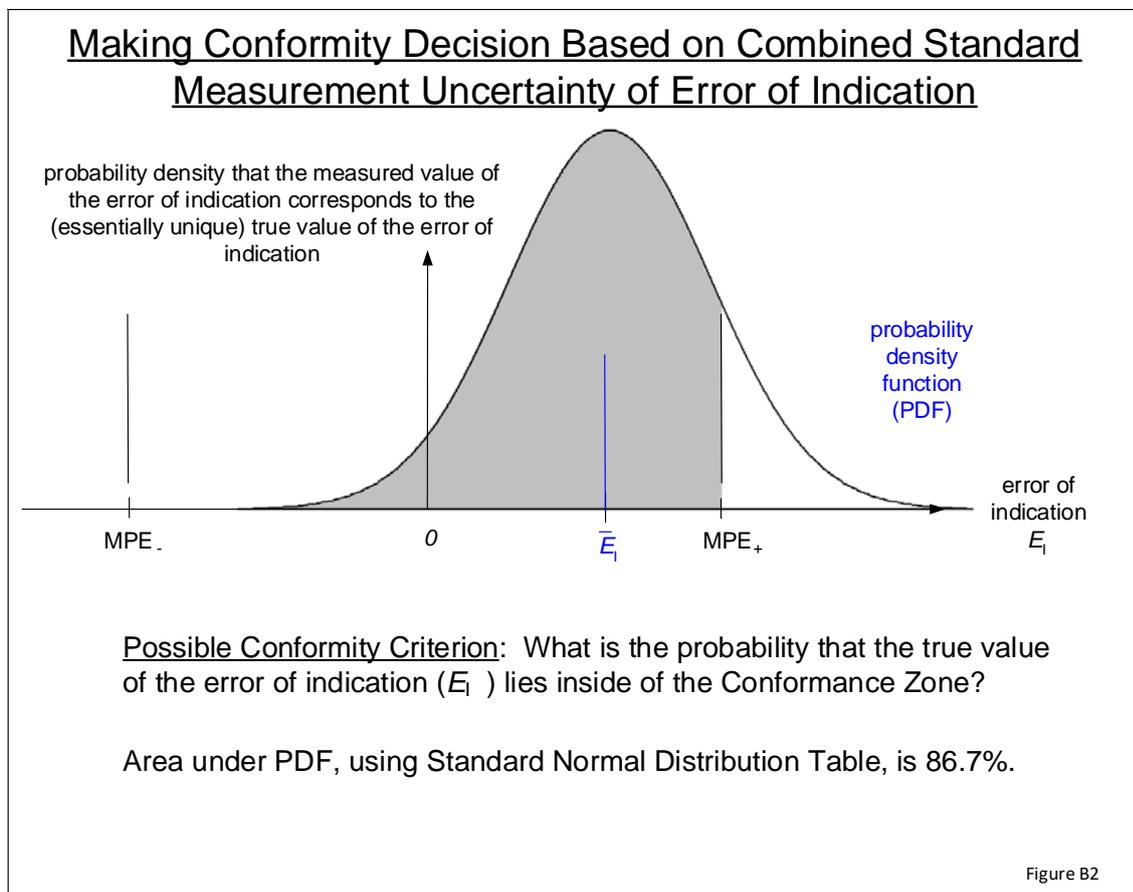
$$u_{E_I} = 180 \mu\text{m} \quad (\text{B.3})$$

If the MPE for this particular test is given as  $500 \mu\text{m}$ , then  $\alpha$  is calculated as

$$\alpha = [(MPE_+ - \bar{E}_I) / u_{E_I}] = [(500 - 300) / 180] = 1.11 \quad (\text{B.4})$$

From the standard normal distribution table above, find the entry for 1.11 by reading down the left column under  $\alpha$  to “1.1”, then read across the top row to the heading “0.01”, then read the entry in the table at which the column and row intersect, which is 0.3665.

This means that the area under the curve between  $\bar{E}_I$  and  $MPE_+$  in Figure B1 is 0.3665. Therefore, since the area under the curve to the left of  $\bar{E}_I$  is 0.5000, the probability (assuming that no mistakes were made in the measurement) that the true value of the error of indication is within the conformance zone is  $0.3665 + 0.5000$ , or 0.8665 (86.7 %). Thus, the risk of false acceptance is  $p_{fa} = 1 - 0.8665 = 0.133 = 13.3 \%$ . Note that  $f_{E_I} = u_{E_I} / MPE = 0.36$ , so that the maximum permissible uncertainty test would fail if the maximum value of  $f_{E_I}$  was specified as  $1/3$  for this test.



In the case where  $\bar{E}_1$  is less than 0 the standard normal distribution table can again be used, taking advantage of the symmetry of the Gaussian curve, but it is then necessary to define  $\alpha$  according to:

$$\alpha = [(\bar{E}_1 - \text{MPE.}) / u_{E_1}] \quad (\text{B.5})$$

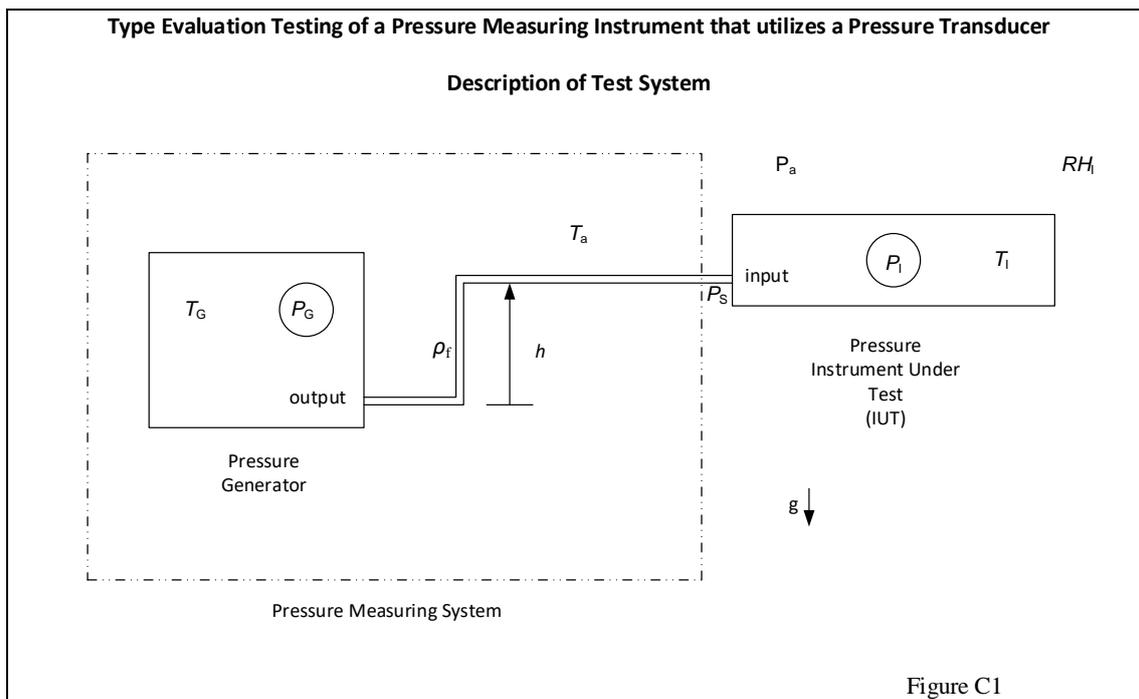
## Annex C

### Example of assessing measurement uncertainty of error of indication

Consider the case of incorporating measurement uncertainty into the decision process (pass/fail) for type evaluation testing of a pressure measuring instrument that utilizes a pressure transducer.

Following the steps given in 8.2 (the steps are presented in *italics*):

*(Step 1) Describe the instrument under test (IUT), along with the measuring system that will be used for performing the test(s). Include in the description all influence quantities that can affect the measuring instrument, all influence quantities that can affect the measuring system, and specify the conditions (if any) at which the influence quantities will be maintained during the testing, or the range(s) that the influence quantities shall remain within during the testing (e.g. rated operating conditions and/or reference operating conditions of both the measuring system and IUT).*



The instrument under test (IUT) is a pressure measuring instrument that utilizes a pressure transducer that, for the sake of illustration, will be considered to be configured in the so-called ‘gauge mode’, meaning that one side of the transducer is open to ambient (atmospheric) pressure (denoted in Figure C1 by  $P_a$ ).

The IUT is located such that it either sits on a bench that is open to the atmosphere (as in Figure C1), or is placed in a chamber where the temperature and relative humidity can be controlled. The temperature of the IUT is indicated as  $T_i$ , and the relative humidity is

indicated as  $RH_I$ . The input to the IUT is indicated in the figure, and this establishes the reference level of the IUT with respect to the indicated gage-mode pressure  $P_I$ .

The pressure measuring system is indicated by the dotted rectangle, and consists of a pressure generator and rigid tubing that connects the output of the pressure generator to the input of the IUT. The operating fluid (which must be specified) is known to have a mass density denoted by  $\rho_f$ , and the height of the reference level of the IUT above the height of the pressure generator is denoted by  $h$  (even if the pressure generator sits on the same bench, the two reference levels will likely be different). The gage mode pressure generated by the pressure generator at its reference level is denoted by  $P_G$ , and the temperature of the pressure generator is denoted by  $T_G$ , which might be different than the temperature of the ambient air, denoted by  $T_a$ . The local acceleration of gravity at the test location is given as  $g$ .

The influence quantities that can affect the outcome of the test are then  $P_a$ ,  $T_a$ ,  $T_G$ ,  $RH_I$  and  $T_I$ . The first three will not be controlled during any of the tests, but rather will only be measured ( $T_G$  will be monitored to make sure that the pressure generator is always operating within its rated operating conditions). On the other hand, some of the tests will involve changing (and measuring) the temperature of the IUT ( $T_I$ ) and the relative humidity of the air surrounding the IUT ( $RH_I$ ).

The other test parameters  $h$ ,  $g$  and  $\rho_f$  are not considered as influence quantities since they do not affect the IUT (or pressure generator).

*(Step 2) Identify all of the different kinds of tests that will need to be performed for the type evaluation. Based on the description in Step 1, develop a mathematical model of the measurement to be used for performing each kind of test. Each model shall ultimately provide an expression for the ‘error of indication,’ and also include an expression for the standard measurement uncertainty to be associated with each measured error of indication (unless repeated measurements of error of indication are to be obtained, in which case the mean value of the error of indication is to be presented, along with an associated standard measurement uncertainty that incorporates a component obtained from the repeated measurements). Account should also be taken in the uncertainty analysis of the range of values of error of indication that could be obtained when the IUT is operating anywhere within its rated operating conditions.*

The different kinds of tests that will need to be performed are given in OIML R 101 *Indicating and recording pressure gauges, vacuum gauges and pressure-vacuum gauges with elastic sensing elements (ordinary instruments)* and R 109 *Pressure gauges and vacuum gauges with elastic sensing elements (standard instruments)*. Included are temperature tests, humidity tests, and hysteresis tests.

The basic mathematical model (for error of indication) for all of these types of tests is based on first generating a mathematical expression for the best estimate of the ‘true’ value of the hydrostatic gage pressure delivered by the pressure measuring system to the input of the IUT (this pressure is denoted by  $P_S$  in Figure C1):

$$P_S = P_G + (\rho_f - \rho_a) \cdot g \cdot h \quad (\text{C.1})$$

where  $\rho_a$  is the density of the ambient air. The mathematical model for the error of indication ( $E_I$ ) of the measuring instrument is then taken as the difference between the indication that would be obtained if no random influences affected the measuring instrument ( $P_I$ ) and the best estimate of the ‘true’ value of the hydrostatic gage pressure ( $P_S$ ) delivered by the pressure measuring system to the input of the IUT:

$$E_I = P_I - P_S \quad (\text{C.2})$$

The combined standard uncertainty of an individually measured value of the error of indication is then obtained from the use of equation 10 of the GUM:

$$u_{E_I}^2 = u_{P_I}^2 + u_{P_S}^2 \quad (\text{C.3})$$

where  $u_{P_I}$  incorporates only resolution limitation and ‘jitter’ of the indication of the IUT, and

$$u_{P_S}^2 = \sum_i \left( \frac{\partial P_S}{\partial x_i} \right)^2 u_{x_i}^2 \quad (\text{C.4})$$

The summation over the index  $i$  covers all of the quantities upon which  $P_S$  depends. (Note that equations C.3 and C.4 are based on the assumption that there is no correlation among the quantities. If such correlation exists, equation 13 of the GUM must be used). From equations C.1 and C.4:

$$u_{P_S}^2 = u_{P_G}^2 + (g \cdot h)^2 \cdot u_{\rho_f}^2 + (-g \cdot h)^2 \cdot u_{\rho_a}^2 + [(\rho_f - \rho_a) \cdot h]^2 \cdot u_g^2 + [(\rho_f - \rho_a) \cdot g]^2 \cdot u_h^2 \quad (\text{C.5})$$

where the individual components of measurement uncertainty must be obtained from various sources, such as tables or calibration certificates. (Note that  $u_{\rho_a}$  itself depends on the temperature and relative humidity of the air). Equation C.5 can then be combined with equation 3 to obtain an expression for the combined standard uncertainty to associate with an individually measured value of the error of indication.

However, for each type of test for the type evaluation, it is necessary to also incorporate a component of measurement uncertainty for the repeatability of the test (denoted  $u_{\text{rep}}$ ). This can be obtained by performing a series of repeated ‘identical’ measurements and calculating the standard deviation of the measured values, or by obtaining such information from measurements that were performed earlier (the method used should be specified).

Also, the IUT should be evaluated to determine how the indication changes (for a fixed input) as the instrument is subjected to likely simultaneous changes in its operating conditions during use in environments outside of a laboratory. A component of uncertainty (denoted  $u_{\text{roc}}$ ), perhaps obtained as the standard deviation of a set of values obtained as the operating conditions of the IUT are randomly varied over the range of rated operating conditions, should also be considered for inclusion in the final expression for  $u_{E_1}$ :

$$u_{E_1}^2 = u_{P_1}^2 + u_{P_S}^2 + u_{\text{rep}}^2 + u_{\text{roc}}^2 \quad \text{C.6}$$

where  $u_{P_S}^2$  is obtained from equation C.5.

Consider a particular type evaluation test where the IUT is operated at its nominal maximum rated operating pressure of 1.01 MPa (10 atmospheres). Let the pressure generator be set to generate a pressure ( $P_G$ ) of 1.0000 MPa, with an uncertainty ( $u_{P_G}$ ), as given from its calibration certificate, of 0.0001 MPa (or 100 Pa).

The operating fluid is a liquid with a mass density (as given by the manufacturer) of 900 kg/m<sup>3</sup> and a corresponding stated measurement uncertainty ( $u_{\rho_f}$ ) of 10 %, or 90 kg/m<sup>3</sup>. The ambient air density ( $\rho_a$ ) depends on the air temperature ( $T_a$ ) [measured to be 23 °C, with an uncertainty of 0.01 °C], the atmospheric pressure ( $P_a$ ) [measured to be 0.10147 MPa, with an uncertainty of 0.00010 MPa], and the relative humidity ( $RH_1$ ) [measured to be 60 %, with an uncertainty of 5 %]. Using known equations for calculating air density,  $\rho_a$  is calculated to be 1.194 kg/m<sup>3</sup>, with an uncertainty of 0.005 kg/m<sup>3</sup>.

As the total variation in the local acceleration of gravity ( $g$ ) over the surface of the Earth can be as much as 0.5 %, the value of the local gravity needs to be established with an uncertainty appropriate for this use. Tables accounting for latitude and height above sea level are available. For this particular test,  $g$  is obtained from such a table to be 9.79560 m/s<sup>2</sup>, with an uncertainty ( $u_g$ ) of 0.00005 m/s<sup>2</sup>.

The height ( $h$ ) of the reference level of the IUT above the reference level of the pressure generator is measured to be 0.0213 m, with a measurement uncertainty ( $u_h$ ) of 0.0001 m.

*(Step 3) calculate the associated standard measurement uncertainty ( $u_{P_S}$ ) of the measurement standard or system.*

The standard measurement uncertainty ( $u_{P_S}$ ) of the pressure delivered by the measurement system to the input of the IUT can be calculated using equation C.5 as

$$\begin{aligned}
 u_{p_s}^2 &= (100)^2 + (9.79560 \cdot 0.0213)^2 \cdot (90)^2 + (-9.79560 \cdot 0.0213)^2 \cdot (0.005)^2 \\
 &\quad + [(900 - 1.194) \cdot 0.0213]^2 \cdot (0.00005)^2 + [(900 - 1.194) \cdot 0.0213]^2 \cdot (0.0001)^2 \\
 &= [10^4 + 352.62 + 1.09 \times 10^{-6} + 9.16 \times 10^{-7} + 0.775] \text{ Pa}^2 \\
 &\approx 10\,353 \text{ Pa}^2,
 \end{aligned}
 \tag{C.7}$$

or

$$u_{p_s} \approx 102 \text{ Pa}$$

It can be seen immediately from this analysis that the uncertainty in the value of the generated pressure dominates the total uncertainty of the pressure delivered to the input of the IUT, followed next by the uncertainty of the density of the operating fluid. Such an analysis helps to identify where efforts could be best spent, if necessary, trying to reduce the uncertainty of the pressure delivered to the input of the IUT.

*(Step 4) calculate a standard measurement uncertainty ( $u_{p_i}$ ) associated with the indicated value of the measurand (including components due to indicator resolution and/or random fluctuations).*

Observed random fluctuations (jitter) in the indicated pressure ( $P_i$ ) of the IUT for a fixed input pressure of 1.01 MPa, and for the operating conditions of the IUT maintained under specified reference conditions, are found to be  $\pm 15$  Pa, which translates into a component of uncertainty of  $u_{p_i}$  of  $15/\sqrt{3} = 8.7$  Pa. Note that the factor  $\sqrt{3}$  comes from treating the uncertainty as a rectangular distribution (see GUM, subclause 4.3.7). Other assumed distributions would result in different divisors.

The resolution of the indication is found to be 10 Pa, which yields a component of uncertainty of  $u_{p_i}$  of  $10/\sqrt{12} = 2.9$  Pa.

The combined standard uncertainty associated with the indication of the IUT is then

$$u_{p_i} = \sqrt{\left[ \left( \frac{15}{\sqrt{3}} \right)^2 + \left( \frac{10}{\sqrt{12}} \right)^2 \right]} = 9.13 \text{ Pa}$$

*(Step 5) calculate a standard measurement uncertainty ( $u_{\text{rep}}$ ) associated with the repeatability or reproducibility of the measuring instrument/system and/or testing procedure.*

A series of repeatability tests is performed on the IUT, where the repeatability condition is that the pressure from the pressure generator is alternately applied and then removed 50 times, everything else remaining constant. Sufficient time is left between pressurizations to allow for thermal equilibrium to be established. Effects due to possible hysteresis are also

analyzed. The calculated standard deviation of the 50 values ( $u_{SD}$ ) is then taken as a component of measurement uncertainty to attribute to repeatability/reproducibility for this particular type of test. For purposes of example, assume that a value of  $u_{SD} = u_{rep} = 20$  Pa is calculated.

*(Step 6) calculate a standard measurement uncertainty ( $u_{roc}$ ) if the indication of the measuring instrument is found to vary when the instrument is operated over its range of rated operating conditions for a fixed input to the instrument.*

Returning to the test conditions in Step 4, now systematically vary (if possible) the operating conditions of the IUT over its range of rated operating conditions, and observe the corresponding variation in the indicated pressure  $P_I$ . Again if possible, vary the operating conditions both individually and also all at once, to simulate possible conditions (temperature test, humidity test, hysteresis test, etc.) that the IUT could experience in environments outside of a laboratory. Say that for such a test the indicated pressure is found to vary by  $\pm 30$  Pa. The corresponding component of measurement uncertainty ( $u_{roc}$ ) due to (possible) variation in operating conditions over the range of rated operating conditions is then

$$u_{roc} = 30/\sqrt{3} = 17.3 \text{ Pa}$$

Note that this component of measurement uncertainty is sometimes explicitly considered to be part of the measurement uncertainty of the error of indication, but is sometimes considered to be incorporated by convention into the MPE, so care should be taken in specifying what convention is being used. In this example this component of measurement uncertainty is not considered to be incorporated into the MPE.

*(Step 7) combine these components of measurement uncertainty in order to calculate a combined standard measurement uncertainty ( $u_{E_I}$ ) associated with the error of indication.*

It is now possible to calculate the combined standard measurement uncertainty of the error of indication ( $u_{E_I}$ ) for the particular type evaluation test where the IUT is operated at its nominal maximum rated operating pressure of 1.01 MPa (10 atmospheres), as described above. Using equation C.6:

$$\begin{aligned} u_{E_I}^2 &= u_{P_I}^2 + u_{P_S}^2 + u_{rep}^2 + u_{roc}^2 \\ &= (9.13)^2 + (102)^2 + (20)^2 + (17.3)^2 = 11\,187 \text{ Pa}^2 \end{aligned}$$

or

$$u_{E_I} = 105.8 \text{ Pa} \tag{C.8}$$

While this example shows virtually everything that it is necessary to consider in order to assess the measurement uncertainty of the error of indication for an individual measurement for this type of measuring instrument (and so may appear somewhat complex), it is important to note that, once all of this derivation has been performed, and values and associated

measurement uncertainties are obtained for typical measurement conditions, the process of obtaining a value of  $u_{E_1}$  for each subsequent individual pressure measurement performed during a given type evaluation test should become relatively straightforward, since most components of measurement uncertainty will not change from one individual measurement to another.

It is in fact interesting to note that the uncertainty of the error of indication presented in equation C.8 was obtained without ever obtaining an explicit value for an individual error of indication, but rather only a nominal (maximum) pressure value was specified for the test. While some of the components of measurement uncertainty may decrease at lower pressures, it is sometimes convenient to just (conservatively) use what is believed to be the maximum uncertainty throughout the testing for a particular type of test.

It is also interesting to note that, in this case, almost all of the uncertainty in the error of indication comes from the measurement standard (i.e. the pressure generator). This is not always the case.

Now that how to assess the measurement uncertainty of the error of indication for the test arrangement in this example has been presented, it is now possible to extend the example to consider how to establish requirements on MPEs, maximum permissible uncertainties and risk options to be considered for making conformity decisions. This will be done in Annex D.

## Annex D

### Example of risk assessment incorporating measurement uncertainty

*For each kind of test identified, the OIML Recommendation should discuss and specify what the appropriate MPE is for that kind of test. For example, for a type evaluation test, the MPE that is specified could correspond to one of several possible accuracy classes that the instrument is being tested for. For a verification test, the specified MPE should be based on a variety of considerations, as discussed in clause 6.*

*There should also be discussion of what the likely values of  $u_{E_1}$  and  $u_S$  will be during the test, in order to decide whether values of  $MPU_{E_1}$  and  $MPU_S$  should be specified and, if so, what those values should be (or, rather, what  $f_{E_1}$  and  $f_S$  should be). See 5.3.4, 5.3.5 and 6.*

Continuing with the example from Annex C1, consider the case where the IUT is to be tested to determine if that type of instrument can be classified as belonging to a specified accuracy class (say, class 0.06 as specified in OIML R 109), that has a corresponding MPE that is designated as  $MPE_{.06} = 0.06\% (1 \text{ MPa}) = 0.0006 \text{ MPa} = 600 \text{ Pa}$ .

An analysis must be performed of whether, for the type of test covered in Annex C on this type of instrument, it is most appropriate to use consumer's risk, producer's risk or shared risk. Things to be considered in the analysis are what the consequences would be (safety, economic and otherwise) to the instrument user and instrument manufacturer of an incorrect pass-fail decision (for either the specified, or likeliest, use of the type of instrument), and what likely values are of  $u_{E_1}$  during the test.

For example, if the type of IUT is typically used to monitor atmospheric pressure for weather forecasting, it might be decided that the shared risk approach is adequate, as long as a specified value of  $f_{E_1}$  (such as 1/3) is adhered to. On the other hand, for a type of pressure measuring instrument being used to monitor critical vessel pressure in a nuclear power plant, or being used for aviation altimetry, the consumer's risk approach should probably be used, with a relatively conservative (smaller)  $f_{E_1}$ .

Before deciding on which risk approach to use, it might be necessary (or at least useful) to first perform some preliminary measurements to determine typical values of  $u_{E_1}$  (which has already been established in equation C.8 in Annex C as being around 105 Pa). These measurements could also be used to help establish an appropriate specified value of  $f_{E_1}$  such that there would be a very small probability of an incorrect pass-fail decision.

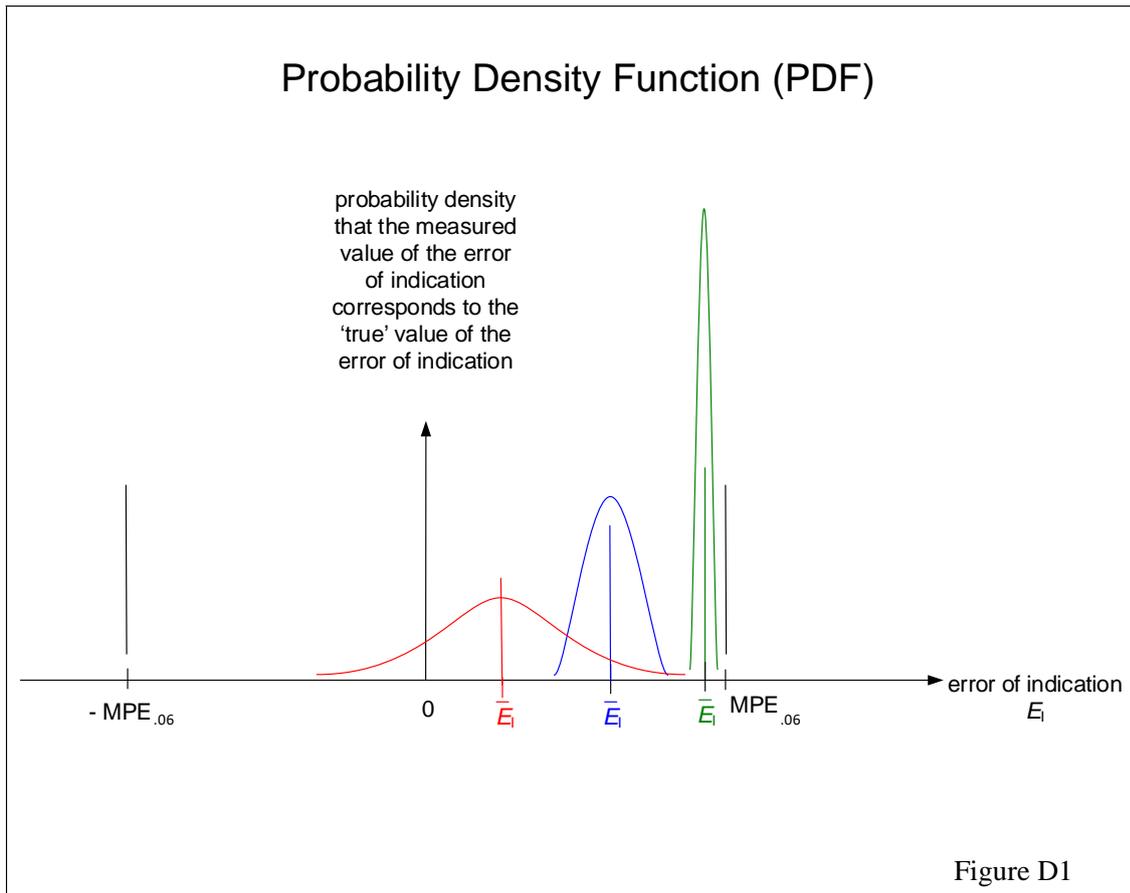


Figure D1 illustrates the situation for the example being discussed. The middle (blue) curve represents a Gaussian PDF where the uncertainty (standard deviation of the curve) is about 1/6 of the MPE ( $u_{E_i} / \text{MPE}_{.06} = 105/600$ ). The leftmost (red) curve represents a Gaussian PDF where the uncertainty is about 1/3 of the MPE. By examining these two curves it is then possible, on just a visual basis, to decide a level of comfort with which either the ratios (corresponding to  $f_{E_i}$ , as discussed in 5.3.4), or a different ratio, should be specified as a requirement in the Recommendation. For the particular example being discussed, assume that the type of IUT will be used in a non-critical application, and so an  $f_{E_i}$  of 1/3 is considered to be acceptable. For a critical application, an  $f_{E_i}$  of 1/20, as indicated schematically by the rightmost (green) curve in Figure D1, might be more appropriate. In this latter case, in order to achieve this smaller value of  $f_{E_i}$ , it would be necessary to either reduce  $u_{E_i}$ , or choose a larger MPE (accuracy class) for this type of instrument to belong to.

Turning to requirements on the measurement standard, values of  $u_{P_S}$  can be obtained from an analysis of the measuring system, including incorporating information contained in the calibration certificate of the measurement standard, in order to help decide whether the measurement standard and measuring system are appropriate to be used for the particular kind of test. This aspect of the test should be discussed and specified in the Recommendation (e.g., an appropriate value of  $f_S$  should possibly be specified, as discussed in 5.3.5). For the pressure example being discussed, the uncertainty associated with the pressure delivered and

measured by the ‘standard’ is given in equation C.7 as  $u_{PS} = 102$  Pa, which is only slightly less than  $u_{E1}$ , and so the middle curve in Figure D1 can again be used for deciding (on a visual basis) whether the uncertainty of the standard is acceptable. In this case the decision to be made is whether the uncertainty due to the standard unfairly affects the pass-fail test decision from the manufacturer’s point of view, in that most of the uncertainty is due to the measurement standard and not the IUT. For the particular example being discussed, a required value of  $f_S = 1/3$  would be acceptable (since the measured value is  $1/6$ ).

*During the development of an OIML Recommendation or other OIML publication, it should be considered whether ‘acceptable’ levels of risk for various types of tests should be suggested. Decision rules and associated risks, along with their consequences, should be considered and discussed in OIML Recommendations. However, this should be done only in the context of regulatory matters. Risks to a manufacturer may have serious economic consequences that are typically outside the scope of a Recommendation.*

*Depending on the values of  $MPU_{E1}$  and  $MPU_S$  (or  $f_{E1}$  and  $f_{PS}$ ) specified in the prior step (if any), discussion should be provided on whether the ‘shared risk’ principle is to be used, or whether there is a specified risk (probability) that is to be used and, if so, whether it is a Risk of False Acceptance or a Risk of False Rejection. Note that if the ‘shared risk’ approach is used in an OIML Recommendation (or in other OIML publications), it should not be used in an implicit manner but, rather, an explicit statement of its use should be provided in the Recommendation.*

Continuing further with the example from Annex C1, next consider the case where the IUT is to be tested for initial verification requirements. In this case, an MPE for initial verification ( $MPE_{iv}$ ) is to be specified in the Recommendation, and so the Recommendation should discuss the various considerations that go into choosing an appropriate  $MPE_{iv}$ , such as needs of the regulator and consumer, and achievable levels of operation of the instrument in environments outside of a laboratory.

As was the case for the type evaluation test, the question of what type of risk and decision rules to use for initial verification must be analyzed, only with now a (typically) larger MPE (it is frequently the case that the  $MPE_{iv}$  is chosen to be twice the MPE, however this is not always necessary), and so the answer to the question might be different. For example, for the type evaluation test it might be decided that using consumer’s risk is appropriate, along with a specified value of  $f_{E1}$ , whereas for the initial (or subsequent) verification test, the use of shared risk (which is easier to handle in environments outside of a laboratory) is adequate, since, with a larger MPE, the PDF might now look more like the rightmost curve in Figure D1, rather than like the middle curve. In such a case it makes sense to avoid computational complication and share the risk, since an ‘incorrect’ decision could be made only over the relative width (which is very small) of the rightmost curve.

*If Risk of False Acceptance or Risk of False Rejection is used, it is further necessary to specify whether  $u_{E1}$  is to be considered as fixed for each measurement, in which case a guard band can be used for deciding conformity, or whether  $u_{E1}$  is to be calculated separately for*

*each measurement of error of indication, in which case the z-statistic or Measurement Capability Index can be used. Reference to the present OIML Guide should be provided, along with discussion of how to use the z-statistic and/or Measurement Capability Index for the particular Recommendation.*

*Constructing PDFs and calculating areas under a PDF curve is in general a nontrivial matter, and so any advice and assistance to be provided in this regard in the Recommendation(s) should be carefully considered during its development.*

For the situation where it is decided that the risk of false acceptance approach is to be used, a decision must be made concerning what is the acceptable level of risk for false acceptance ( $p_{ca}$ , see 5.3.1), and a further analysis must be performed about whether the uncertainty of the error of indication can be taken as constant for each measurement, or whether it is necessary to recalculate it each time.

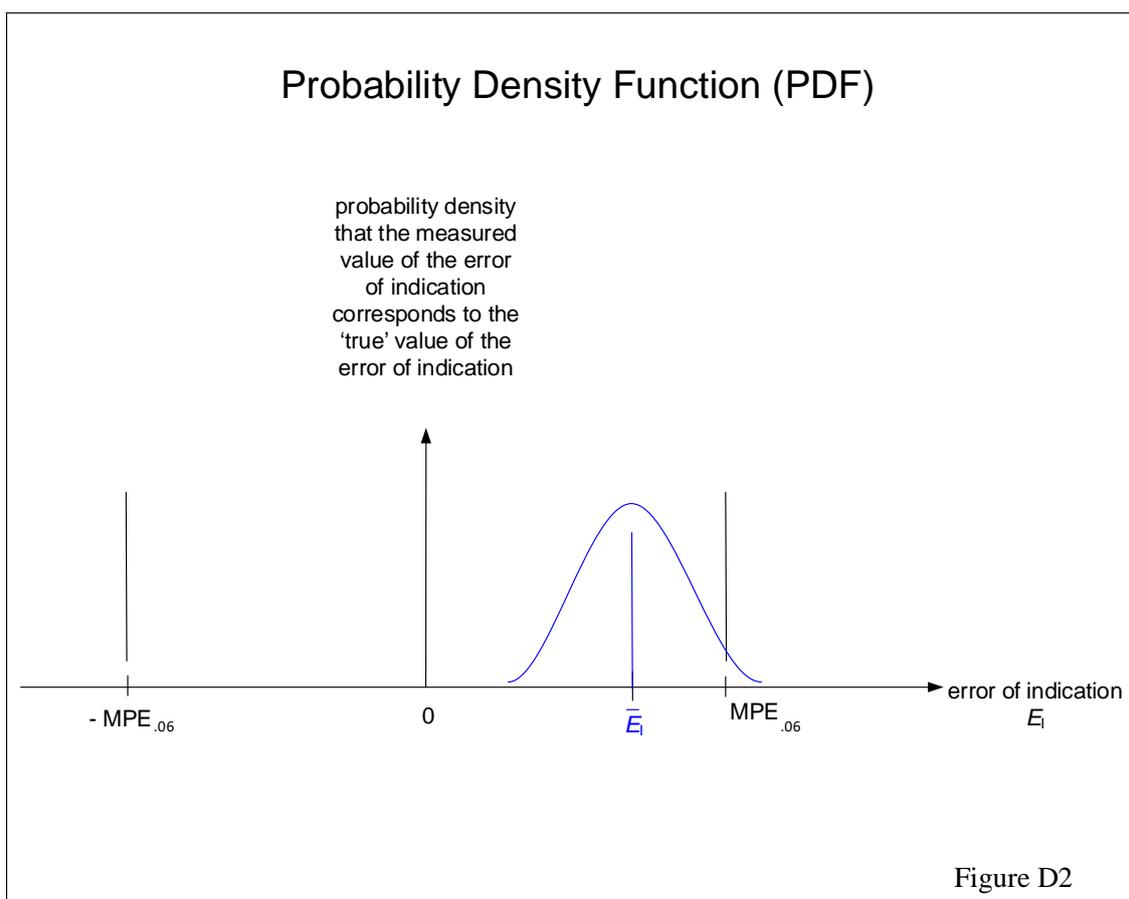
If  $u_{E_1}$  needs to be calculated each time, then it is necessary to either use the Z-table each time (e.g. see Annex B), or to calculate the Measurement Capability Index ( $C_m$ ) each time (e.g. see Annex E) and use the corresponding  $C_m$  table each time.

If  $u_{E_1}$  can be considered as a constant for a given type of measurement, and so does not need to be calculated each time, then a guard band can be constructed by shifting the MPE boundaries inward by a fixed amount (so as to keep the probability of false acceptance less than a specified value; see [5]). Pass-fail decisions are then made on the basis of whether the measured  $E_1$  lies within the new (reduced) MPE boundaries.

Returning to the type evaluation test for the example in Annex C, Figure C1 (and above), assume that it is decided that a 5 % level of risk of false acceptance (consumer's risk) will be used for this application of the IUT (i.e.  $p_{ca} = .05$ , and thus the probability of conformance is  $p_c = 0.95 = 95\%$ ). Since for this example it has been determined that  $MPE_{.06} = 600$  Pa and  $u_{E_1}$  (at maximum pressure of 1 MPa) = 105 Pa, the standard normal distribution table in Annex B can be used to determine the maximum value of the error of indication. Begin by locating the entry in that table for 0.9500 (or actually for 0.4500, since 0.5000 needs to be subtracted from 0.9500 in this case for the table in Annex B), which is between the entries 0.4495 ( $\alpha = 1.64$ ) and .4505 ( $\alpha = 1.65$ ). Using interpolation, the value of  $\alpha$  that will be used is then 1.645. Equation B.1 can then be used, in a slightly rearranged form:

$$\bar{E}_1 = MPE_{.06} - (u_{E_1} \cdot \alpha) \quad \text{D.1}$$

to obtain  $\bar{E}_1 = 425$  Pa, which is the maximum value that  $\bar{E}_1$  can have where there is no greater than a 5 % risk that the test should have been considered to fail even though it is considered to pass. This situation is demonstrated graphically in Figure D2.

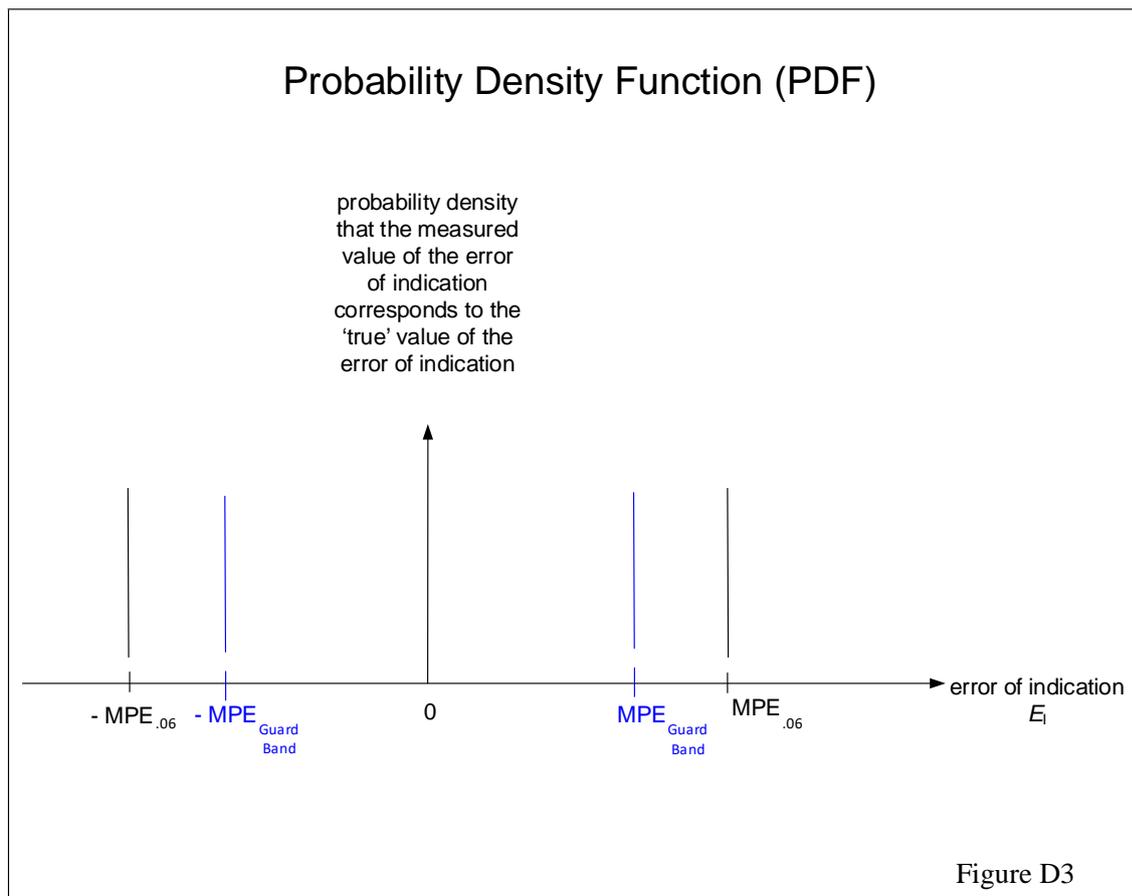


Rather than using the Z-table, it may be more convenient to use the measurement capability index chart to arrive at this same conclusion (see Annex E). In this case, the measurement capability index is first calculated using equation E.1 as  $C_m = \text{MPE}/[2 \cdot u_{E_I}] = 600/[2 \cdot 105] = 2.86$ . Using the 95 % chart in Annex E, the corresponding value of  $\hat{E}$  is about 0.85. Rearranging equation E.2,  $E_I = \text{MPE} (2 \cdot \hat{E} - 1) = 600 (1.7 - 1) = 420 \text{ Pa}$ , which is close to the 425 Pa obtained when using the more precise Z-table.

*While assessing the measurement uncertainty of the error of indication for an individual measurement for a specified type of measuring instrument may be somewhat complex, it is important to note that, once all of the derivation has been performed, and values and associated measurement uncertainties are obtained for typical measurement conditions, the process of obtaining a value of  $u_{E_I}$  for each subsequent individual measurement performed during a given type evaluation test should become relatively straightforward, since most components of measurement uncertainty will not change from one individual measurement to another. This aspect of the treatment of measurement uncertainty should be included in the discussion in each OIML Recommendation where measurement uncertainty is relevant.*

If it is determined experimentally that there is significant variation in  $u_{E_I}$  from one measurement to the next, then it will be necessary to use either the Z-table or measurement

capability index for each measurement of  $E_I$ . However, as indicated earlier, it is unlikely that  $u_{E_I}$  will vary appreciably for each measurement and, besides, it is sometimes more convenient to take a conservative approach and treat the  $u_{E_I}$  determined in Annex C as the likely upper bound of all of the  $u_{E_I}$ 's, and so treat it as a constant. In this case, a guard band can be created (where the new MPE is moved inward from 600 Pa to 425 Pa) and the decision-making becomes much simpler, where tests involving measured values of  $E_I$  less than 425 Pa are accepted, and those greater are rejected. This guard band approach is illustrated in Figure D3.



## Annex E

### Measurement Capability Index ( $C_m$ )

The “measurement capability index,” defined and discussed in [5], is a useful tool for quickly assessing whether a measured error of indication ( $E_I$ ), with associated combined standard uncertainty ( $u_{E_I}$ ), is considered to conform to the maximum permissible error (MPE) requirement within a specified conformance probability ( $p_c$ ).

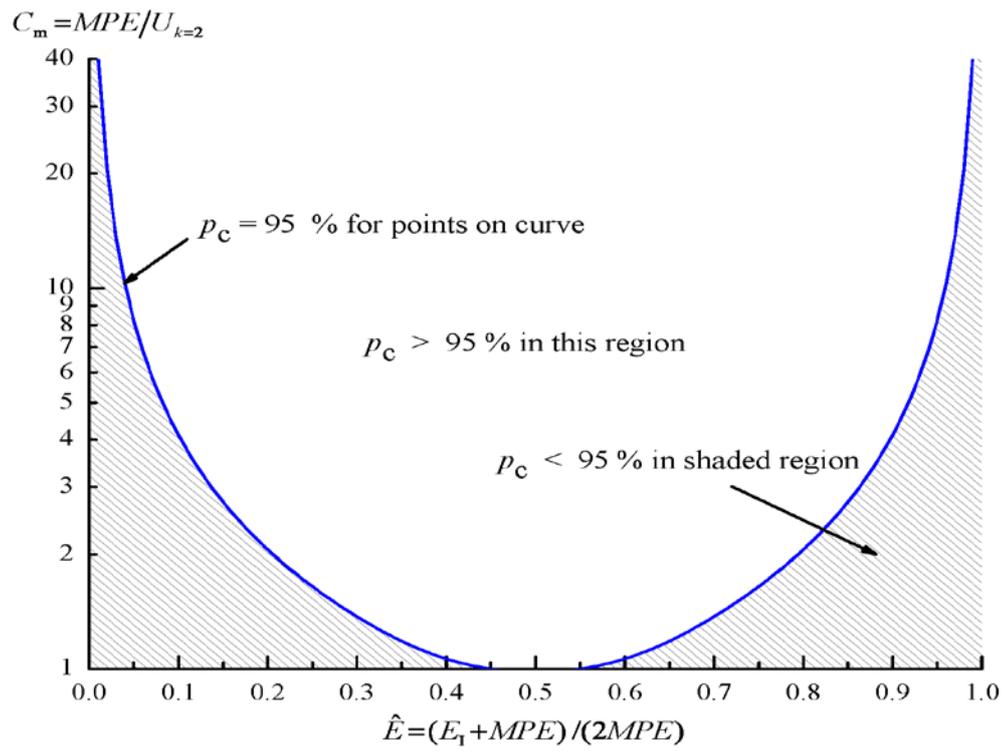
The measurement capability index is dimensionless, and defined for legal metrology as:

$$C_m = MPE/[2 \cdot u_{E_I}] = MPE/U_{k=2} \quad (\text{E.1})$$

In order to use the measurement capability index, it is first necessary to calculate another dimensionless parameter,  $\hat{E}$ , defined as:

$$\hat{E} = [E_I + MPE]/[2 MPE] \quad (\text{E.2})$$

Note that for  $-MPE < E_I < MPE$ , then  $0 < \hat{E} < 1$ . A chart such as the one below can then be constructed for a given  $p_c$  (shown here for  $p_c = 95\%$ ), where the intersection of  $\hat{E}$  and  $C_m$  can be found to see if it lies in the shaded region (test fails) or un-shaded region (test passes). (Figure courtesy of W. Tyler Estler).



**Annex F**  
**Establishing measurement uncertainty to use with conformity tested measuring instruments and systems**

Once a measuring instrument has passed an initial or subsequent verification test, it is sometimes used to perform a measurement where it is required that the measured value is accompanied by its associated measurement uncertainty. In such a situation, unless the instrument was not only verified but also calibrated, all that can be said about any measured value obtained when using the instrument is that the ‘true’ value of the measurand is believed to be best represented by the measured value (as given by the indication of the measuring instrument), but that the ‘true’ value could lie anywhere (with equal probability) in the range given by the measured value, plus or minus the MPE. This is the so-called ‘rectangular probability distribution’ - see GUM, 4.4.5.

According to that analysis, the measurement uncertainty that should be associated with the measured (indicated) value is

$$u = MPE / \sqrt{3} \tag{F.1}$$

where MPE is the maximum permissible error that was used when the measuring instrument was tested.