UKAS

A world of confidence

M3003

Edition 6, March 2024

The expression of uncertainty and confidence in measurement

Contents

Changes since last edition

- Various minor grammatical and readability changes throughout
- §1.5 moved to this section, *Changes since last edition*, for consistency with other UKAS publications
- §1.6 on editorial changes to the current edition of M3003 removed
- Minor numerical corrections and changes (most notably at K.8.3)
- New exercises: K9, K10, K11
- Revision of Appendix L
- New Appendix Q – input uncertainty expressed as a relative quantity
- Relabelling of original appendices Q and R to R and S

1 Introduction

1.1 The general requirements that testing and calibration laboratories must meet if they wish to demonstrate that they operate to a quality system, are technically competent and are able to generate technically valid results, are contained within ISO/IEC 17025 [5]. This international standard forms the basis for international laboratory accreditation and in cases of differences in interpretation always remains the authoritative document. M3003 is not intended as a prescriptive document and does not set out to introduce additional requirements to those in ISO/IEC 17025 but instead aims to provide amplification and guidance on the current requirements within the standard.

1.2 The purpose of these guidelines is to support policy on the evaluation and reporting of measurement uncertainty for testing and calibration laboratories. Related topics, such as evaluation of conformity with specifications, are addressed in UKAS LAB 48 [14]. Several worked uncertainty examples are included in M3003 to illustrate how practical implementation can be achieved. Further practical guidance for Medical Laboratories meeting requirements of ISO 15189 [6] is provided in ISO TS 20914 [9].

1.3 The guidance in this document is based on information in the *Guide to the Expression of Uncertainty in Measurement*, hereinafter referred to as the GUM [1]. M3003 is consistent with the GUM suite of documents both in methodology and terminology. It does not, however, have the same breadth of scope as the GUM, which also includes other methods of uncertainty evaluation that may be more appropriate to a specific discipline, for example, the use of Monte Carlo simulation [2].

1.4 M3003 is aimed both at the beginner and at those more experienced in the subject of measurement uncertainty. To address the needs of an audience with a wide spectrum of experience, the subject is introduced in relatively straightforward terms and describes the basic concepts involved. Cross-references are made to several Appendices in which more detailed information is presented for those seeking fuller guidance on the subject. For a more in-depth understanding of measurement uncertainty, courses such as the practitioner course offered by UKAS are recommended.

2 Overview

2.1 In many aspects of everyday life, we are accustomed to the doubt that arises when estimating how large or small things are. For example, if somebody asks, “what do you think the temperature of this room is?” we might say, “it is about 23 degrees Celsius”. The use of the word “about” implies that we know the room is not exactly 23 degrees but is somewhere near it. In other words, we recognise that there is some doubt about the value of the temperature that we have estimated.

2.2 We could, of course, be a bit more specific. We could say, “it is 23 degrees Celsius give or take a couple of degrees”. The term “give or take” implies that there is still doubt about the estimate, but now we are assigning limits to the extent of the doubt. We have given some quantitative information about the doubt, or **uncertainty**, of our estimate.

2.3 It is also quite reasonable to assume that we may be more certain that our estimate is within, say, 5 degrees of the “true” room temperature than we are that the estimate is within 2 degrees. The larger the uncertainty we assign, the more confident we are that it encompasses the “true” value. Hence, for a given situation, the uncertainty is related to the level of confidence.

2.4 So far, our estimate of the room temperature has been based on a subjective evaluation. This is not entirely a guess, as we may have experience of exposure to similar and known environments. However, to make a more objective measurement it is necessary to make use of a measuring instrument of some kind; in this case we can use a thermometer.

2.5 Even if we use a measuring instrument, there will still be some doubt, or uncertainty, about the result. For example, we could ask:

“Is the thermometer accurate?”

“How well can I read it?”

“Is the reading changing?”

“I am holding the thermometer in my hand. Am I warming it up?”

“The relative humidity in the room can vary considerably. Will this affect my results?”

“Does it matter whereabouts in the room I take the measurement?”

All these factors, and possibly others, may contribute to the uncertainty of our measurement of the room temperature.

2.6 In order to quantify the uncertainty of the room temperature measurement we will therefore have to consider all the factors that could influence the result. We will have to make estimates of the possible variations associated with these influences. Let us consider the questions posed above.

2.7 <u>Is the thermometer accurate?</u>

2.7.1 In order to find out, it will be necessary to compare it with a thermometer whose accuracy is better known. This thermometer, in turn, will have to be compared with an even better characterised one, and so on. This sequence leads to the concept of **traceability of measurements**, whereby measurements at all levels can be traced back to agreed references. In most cases, ISO/IEC 17025 [5] requires that measurements are traceable to SI units, which is usually achieved by an unbroken chain of comparisons originating at a national metrology institute.

In other words, we need a **traceable calibration**. This calibration itself will provide a source of uncertainty, as the calibrating laboratory will assign a calibration uncertainty to the reported values. When used in a subsequent evaluation of uncertainty, this is often referred to as an **imported uncertainty**.

2.7.2 In terms of the thermometer accuracy, however, a traceable calibration is not the end of the story. Measuring instruments change their characteristics as time goes by. Because they "drift" regular recalibration is necessary. It is therefore important to evaluate the likely change since the instrument was last calibrated.

If the instrument has a reliable history, it may be possible to predict what the measurement error will be at a given time in the future, based on past results, and apply a correction to the reading. This prediction will not be perfect and therefore an uncertainty on the corrected value will be present. In other cases, there may be insufficient past data, or it may not indicate a reliable trend, and a limit value may have to be assigned for the likely change since the last calibration. This value can be estimated from examination of changes that occurred in the past. Evaluations made using these methods yield the uncertainty due to secular stability, or changes with time, of the instrument. This change with time is commonly known as **drift**.

2.7.3 There are other possible influences relating to the thermometer accuracy. For example, suppose we have a traceable calibration, but only at 15 °C, 20 °C and 25 °C. What does this tell us about its indication error at 23 °C?

In such cases we will have to make an estimate of the applicable calibration error, often by interpolation between points where calibration data is available. The associated measurement uncertainty might usually be interpolated in the same fashion, with some additional allowance for uncertainty in the method of interpolation.

2.8 <u>How well can I read it?</u>

2.8.1 There will inevitably be a limit to which we can resolve the reading we observe on the thermometer. If it is a liquid-in-glass thermometer, this limit will often be imposed by our ability to interpolate between the scale graduations. If it is a thermometer with a digital readout, digital rounding will define the limit.

2.8.2 For example, suppose the last digit of a digital thermometer is rounded so that its displayed value changes in steps of 0.1 °C. The reading happens to be 23.4 °C.

The reading is a rounded representation of a larger series of values that the thermometer would indicate if it had more digits available. In the case of a reading of 23.4 °C, this represents all possible values in the range between 23.35 °C and 23.45 °C, which all round to 23.4 °C.

A reading of 23.4 °C therefore means that the value is somewhere between 23.35 °C and 23.45 °C. In other words, the 0.1 °C resolution of the display has caused a rounding error somewhere between 0.05 °C and -0.05 °C (corresponding to plus or minus half of the display resolution). As we have no way of knowing whereabouts in this range the value is located, we must assume the rounding error is zero with limits of ±0.05 °C. (Zero is the 'expectation' value – it is the best estimate based upon the available information).

2.8.3 It can therefore be seen that there will always be an uncertainty of ± half of the change represented by one increment of the last recorded digit. This rounding error does not only apply to digital displays; it applies *every time a number is recorded*. If we write down a rounded result of 123.456, we are imposing an identical effect by the fact that we have recorded this result to three decimal places, and an error not exceeding 0.0005 will arise.

2.8.4 This source of uncertainty is frequently referred to as “resolution”, however it is more correctly the numeric rounding *caused by* finite **resolution**.

2.9 Is the reading changing?

2.9.1 Yes, it probably is! Such changes may be due to variations in the room temperature itself, variations in the performance of the thermometer and variations in other influence quantities, such as the way we are holding the thermometer.

So, what can be done about this?

2.9.2 We could, of course, just record one reading and say that it is the measured temperature at a given moment and under particular conditions. This would have little meaning, as we know that the next reading, a few seconds later, could well be different. So, which is “correct”?

2.9.3 In practice, we will probably take an average of several measurements to obtain a more representative value. In this way, we can “smooth out” the effect of short-term variations in the thermometer indication. This average, or arithmetic mean, of several readings can often be closer to the “true” value than any individual reading is.

2.9.4 However, we can only take a finite number of measurements. This means that we will never obtain the “true” mean value that would be revealed if we could carry out an infinite (or very large) number of measurements. There will be an unknown error, and therefore an uncertainty, arising from the difference between our calculated mean value and the underlying “true” mean value.

2.9.5 This uncertainty cannot be evaluated using methods like those we have already considered. Up until now, we have looked for evidence, such as calibration uncertainty and secular stability, and we have considered what happens with finite resolution by logical reasoning. The effects of variation between readings cannot be evaluated like this, because there is no background information available upon which to base our evaluation.

2.9.6 The only information we might have is a series of readings and a calculated average, or mean value. We therefore need to use a statistical approach to determine how far our calculated mean could be away from the “true” mean. These statistics are quite straightforward… the so-called **repeatability uncertainty** is therefore estimated from the experimental standard deviation of the mean, often referred to as simply the **standard deviation of the mean**.

NOTE: Standard deviation of the mean is also known as *standard error.*

2.9.7 It is often convenient to regard the calculation of the standard deviation of the mean as a two-stage process. It can be performed easily by most scientific calculators or spreadsheet software.

2.9.8 First, we obtain the estimated **repeatability standard deviation**, s, e.g., using the values we have measured. This facility is indicated on most calculators by the function key $x\sigma_{n-1}$. On some calculators it is identified as $s(x)$ or simply s.

In Microsoft Excel the STDEV.S cell function can be used.

2.9.9 The standard deviation of the mean is then obtained by dividing the estimate obtained in 2.9.8 by the square root of the number of measurements that contributed to the mean value.

2.9.10 For example, suppose we record five consecutive readings with our thermometer. These are 23.0 °C, 23.4 °C, 23.1 °C, 23.6 °C and 22.9 °C, and we intend to report the mean 23.2 °C of these five values.

2.9.11 We obtain an estimated standard deviation of 0.2915 °C.

2.9.12 Five measurements contributed to the mean value, so we divide 0.2915 °C by the square root of 5, giving a repeatability estimate (standard deviation of the mean) equal to $\frac{0.2915}{\sqrt{5}} = \frac{0.2915}{2.236} = 0.1304\ ℃.$

2.9.13 Further information on the statistical processes used for evaluation of repeatability can be found in Section 4.

2.10 <u>I am holding the thermometer in my hand. Am I warming it up?</u>

2.10.1 Quite possibly. There may be heat conduction from the hand to the temperature sensor. There may be radiated heat from the body impinging on the sensor. These effects may or may not be significant, but we will not know until an evaluation is performed. In this case, special experiments may be required to determine the significance of the effect.

2.10.2 How could we do this? Some basic methods come to mind. For example, we could set up the thermometer in a temperature-stable environment and read it remotely, without the operator nearby. We could then compare this result with that obtained when the operator is holding it in the usual manner, or in a variety of manners. This would yield empirical data on the effects of heat conduction and radiation. If such effects turn out to be significant, we could either improve the method so that operator effects are eliminated, or we could include a contribution to measurement uncertainty based on the results of the experiment.

2.10.3 Consideration of the measurement method reveals several important issues. The measurement may not be independent of the operator and special consideration may have to be given to operator effects (we may have to train the operator to use the equipment in a certain way). Special experiments may be necessary to evaluate particular effects. Additionally, and significantly, evaluation of uncertainty may reveal ways in which the method can be improved, thus giving more reliable results.

2.11 <u>The relative humidity in the room can vary considerably. Will this affect my results?</u>

2.11.1 Maybe it will. If we are using a liquid in glass thermometer, it is difficult to see how the relative humidity could significantly affect the expansion of the liquid. However, if we are using a digital thermometer, it is possible that relative humidity could affect the electronics that amplify and process the signal from the sensor. The sensor itself could also be affected by relative humidity.

2.11.2 As with other influences, we need means of evaluating any such effects. In this case, we could expose the thermometer to an environment in which the temperature can be maintained at a constant level, but the relative humidity can be varied… which would reveal how sensitive the thermometer is to the quantity we are concerned about. Alternatively, we might rely upon information published by the equipment manufacturer.

2.11.3 This question also raises a general point that is applicable to all measurements. Every measurement we make must be carried out in an environment of some kind; it is unavoidable. So, we must consider whether any particular aspect of the environment could have an effect on the measured value and its uncertainty.

2.11.4 The significance of a particular aspect of the environment must be considered in the light of the specific measurement being made. For example, it is difficult to see how gravity could significantly influence the reading on a digital thermometer. However, it certainly will affect the results obtained on a precision weighing machine that might be right next to the thermometer!

2.11.5 The following environmental effects are amongst the most commonly encountered when considering measurement uncertainty:

Temperature
Relative humidity
Barometric pressure
Electric or magnetic fields
Gravity
Electrical supplies to measuring equipment
Air movement
Vibration
Light and optical reflections

Furthermore, some of these influences may have little effect as long as they remain constant but could affect measurement results when they start changing. Rate of change of temperature can be particularly important.

2.11.6 It should be apparent by now that understanding of a measurement system is important in order to identify and quantify the various uncertainties that can arise in a measurement situation. Conversely, analysis of uncertainty can often yield a deeper understanding of the system and reveal ways in which the measurement process can be improved, which leads on to the next question…

2.12 <u>Does it matter where in the room I make the measurement?</u>

2.12.1 It depends on what we are trying to measure! Are we interested in the temperature at a specific location, or in the average of the temperatures encountered at any location within the room, or the average temperature at bench height?

2.12.2 There may be further, related questions. For example, do we require the temperature at a certain time of day, or the average over a specific period of time?

2.12.3 Such questions must be asked, and answered, in order that we can devise an appropriate measurement method that gives us the information we require. Until we know the details of the method, we are not able to evaluate the uncertainties that will arise from that method.

2.12.4 This question and those preceding it are important questions to ask. But the most important question of all is one that should be asked before we even select a method and start our uncertainty evaluation:

2.13 <u>“What exactly is it that I am trying to measure?”</u>

2.13.1 Until this question is answered, we are not able to carry out a proper evaluation of the uncertainty. The quantity of interest is known as the **measurand**. To evaluate the uncertainty in a measurement we must define the measurand, otherwise we are not able to know how any particular influence quantity affects the value we obtain for it.

2.13.2 A consequence of this is that we need to establish a **measurement model**, which defines the assumed relationship between the influence (input) quantities and the measurand (output). This relationship can often described by a mathematical expression or **measurement equation**. Further details about establishing a measurement model can be found in JCGM GUM-6 [14] and

in Appendix D.
A proper analysis of this process also gives the answer to another important question…

2.14 “Am I actually measuring the quantity that I thought I was measuring?”

2.14.1 Most measurement processes are such that the end result would be only an approximation to the “true” value because of assumptions and approximations inherent in the chosen measurement method. The model should recognise any such assumptions and uncertainties that may arise from them should be accounted for in the analysis.

2.15 Summary

2.15.1 This section of M3003 has given an overview of uncertainty and some insights into how uncertainties might arise. It has shown that we must understand our measurement process and the way in which various influences can affect the result. It has also shown that analysis of uncertainty can have positive benefits in that it can reveal where enhancements can be made to measurement methods, hence improving the reliability of measurement results.

2.15.2 The following sections of M3003 explore the issues identified in this overview in more detail.

3 In more detail

3.1 The Overview section of M3003 has introduced the subject of uncertainty evaluation and has explored a number of the issues involved. This section provides a slightly more formal description of these processes, using terminology consistent with that in the GUM.

3.2 The *International vocabulary of metrology* (VIM) [4] defines a quantity (Q) as a property of a phenomenon, body or substance to which a magnitude (expressed as a number and a reference) can be assigned.

3.3 The purpose of a measurement is to assign a magnitude to the measurand; the quantity intended to be measured. The assigned magnitude is considered to be the best estimate of the value of the measurand. The uncertainty evaluation process will encompass several other 'influence' quantities that affect the result obtained for the measurand. These influence, or 'input', quantities are often referred to as X and the 'output' quantity, i.e., the measurand, is referred to as Y.

3.4 As there will usually be several influence quantities, they are differentiated from each other by the subscript i. So, there will be several input quantities called X_i, where i represents integer values from 1 to N, N being the number of such quantities. In other words, there will be input quantities of $X_1, X_2, \dots, X_N$.

3.5 Each of these input quantities will have a corresponding value. For example, one quantity might be the temperature of the environment – this will have a value, say 23 °C. A lower-case "x" represents the estimated values of the quantities. Hence the value of X_1 will be x_1, that of X_2 will be x_2, and so on.

3.6 The purpose of the measurement is to determine the best estimate of the measurand, Y. As for the input quantities, the estimated value of the measurand is represented by the lower-case letter, i.e., y. One of the first steps is to establish the mathematical relationship $Y = f(X_i)$ between the values of the input quantities, X_i, and that of the measurand, Y. This process is examined in Appendix D.

3.7 The values x_i of the input quantities X_i will generally all have an associated uncertainty. This can be expressed as $u(x_i)$, the standard uncertainty of x_i. The process of 'standardising' the available information about the uncertainty in x_i is described shortly. The uncertainty $u(y)$ associated with y will involve a combination of the input uncertainties $u(x_i)$.

3.8 Some uncertainties, particularly those associated with the determination of repeatability, must be evaluated by statistical methods. Others must be evaluated by examining other information, such as data in calibration certificates, evaluation of long-term drift, consideration of the effects of environment, etc.

3.9 The GUM [1] differentiates between statistical evaluations and those using other methods. It categorises them into two types – **Type A** and **Type B**.

3.10 A *Type A* evaluation of uncertainty is carried out using statistical analysis of a series of observations. Further details about Type A evaluations can be found in Section 4.

3.11 A *Type B* evaluation of uncertainty is carried out using methods other than statistical analysis of a series of observations. Further details about Type B evaluations can be found in Section 5.

3.12 In paragraph 3.3.4 of the GUM it is stated that the purpose of the Type A and Type B classification is to indicate the two different ways of evaluating uncertainty components, and the distinction between Type A and Type B is for convenience in discussion only. Whether components of uncertainty are classified as `random' or `systematic' in relation to a specific measurement process or described as Type A or Type B depending on the method of evaluation, all components regardless of classification are modelled by probability distributions, usually characterized by their limits, standard deviation (or variance) and degrees of freedom.

3.13 Therefore, any convention as to how uncertainty evaluations are classified does not affect the estimation of the combined uncertainty (defined in 3.36). In this guide, when the terms `random' and `systematic' are used they refer to the effects of uncertainty on a specific measurement process. It is the usual case that random components require Type A evaluations and systematic components require Type B evaluations, but there are exceptions.

3.14 For example, a random effect can produce a fluctuation in an instrument's indication, which is both noise-like in appearance and significant in terms of uncertainty. But it may only be possible to characterise it in terms of limits to the range of indicated values. This is not a common situation but when it occurs a Type B evaluation of the uncertainty component will be required. This is done by assigning limit values and an associated probability distribution, as in the case of other Type B evaluations.

3.15 The input uncertainties, associated with the values x_i of the influence quantities X_i, arise in a number of forms. Some may be characterised as limit values between which little is known about the most likely place (within the limits) where the "true" value may lie. A good example of this is the numeric rounding caused by finite resolution described in paragraph 2.8. In this example, it is equally likely that the underlying value is anywhere within the defined limits of ± half of the change represented by one increment of the last recorded digit. This concept is illustrated in Figure 1.

3.16

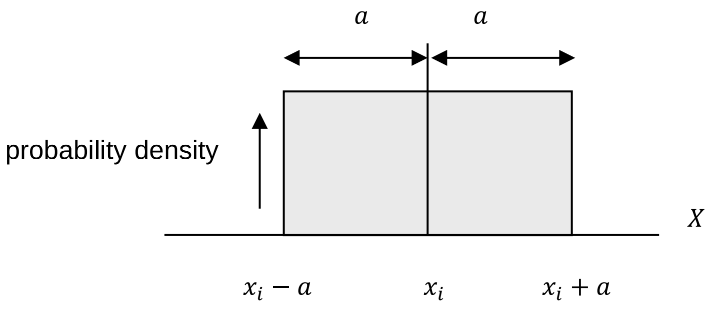


Figure 1

The expectation value x_i lies in the centre of a distribution of possible values with a half-width, or semi-range, of a.

3.17 In the resolution example, a = 0.5 (for the resolution of 1.0).

3.18 As all underlying values are presumed equally likely, we can say that there is equal probability of the value of x_i being anywhere within the range $x_i - a$ to $x_i + a$, and zero probability of it being outside these limits.

3.19 Thus, the uncertainty contribution associated with the value x_i is characterised by a **probability density function** (PDF), describing the range and relative likelihood of possible values of the measurand.

By the GUM definition, the **standard uncertainty** $u(x_i)$ is equal to the standard deviation of the corresponding PDF.

3.20 The probability distribution associated with an input quantity reflects the *available knowledge* about that particular quantity. In many cases, there will be insufficient information available to justify choosing a more 'informative' distribution than a uniform, or rectangular, probability distribution (as in Figure 1).

3.21 If more information is available, it may be possible to assign a different probability distribution to the value of a particular input quantity. For example, a measurement may be taken as the *difference* in readings on a digital scale – typically, the zero reading will be subtracted from a reading taken further up the scale. If the sensitivity is constant, both readings might have an associated rectangular distribution of identical size. If two identical rectangular distributions, each of magnitude $\pm a$, are combined then the resulting distribution will be triangular with a semi-range of $\pm 2a$.

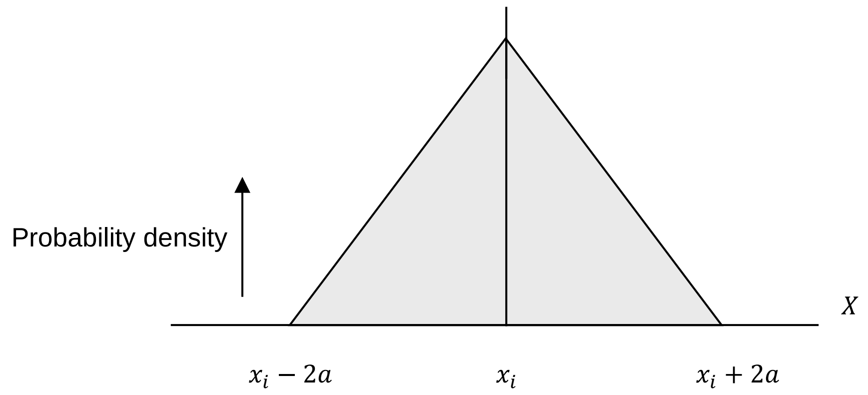


Figure 2
Combination of two identical rectangular distributions, each with semi-range limits of $\pm a$, yields a triangular distribution with a semi-range of $\pm 2a$.

3.22 There are other possible distributions that may be assigned. For example, when making measurements of radio-frequency power an uncertainty arises due to imperfect matching between the source and the termination. The imperfect match usually involves an unknown phase angle which means that a cosine function characterises the probability distribution for the uncertainty. Harris and Warner [17] have shown that a symmetrical U-shaped probability distribution arises from this effect.

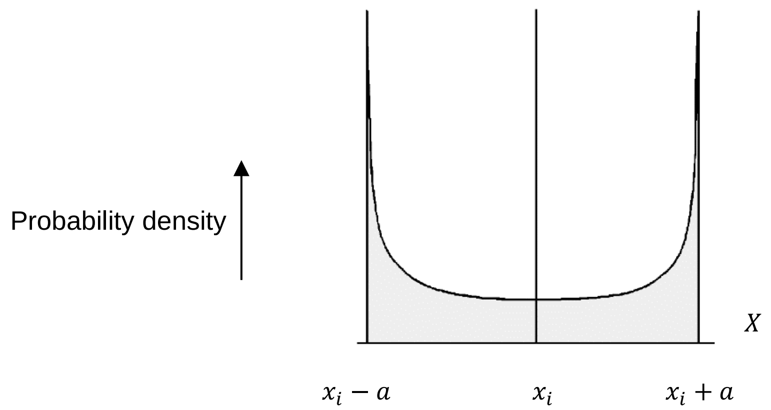


Figure 3
U-shaped distribution, associated with RF mismatch uncertainty. For this situation, x_i is more likely to be close to one or other of the edges of the distribution.

3.23 An evaluation of the effects of (non) repeatability, performed by statistical methods, will usually yield a Gaussian or normal distribution. Further details on this process can be found in Section 4.

3.24 When a number of distributions of whatever form are combined it can be shown that, apart from in exceptional cases, the resulting probability distribution tends to the normal form in accordance with the **Central Limit Theorem**.[16] The importance of this fact is that it makes it possible to use the well-known properties of the normal distribution to assign a **coverage probability** to the likelihood of the true value of the measurand being within a certain range of values, known as the **coverage interval**.

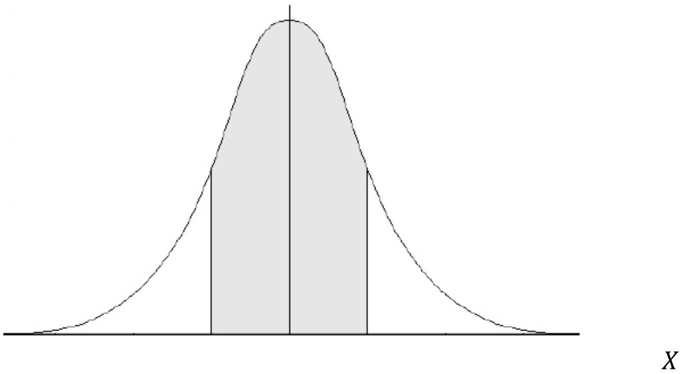


Figure 4

The normal, or Gaussian, probability distribution, obtained when a number of distributions, of any form, are combined and the conditions of the Central Limit Theorem are met. In practice, if three or more distributions of similar magnitude are present, they will usually combine to form a reasonable approximation to the normal distribution.

The size of the distribution is described in terms of a *standard deviation*. The shaded area bounds a region ±1 standard deviation from the centre of the distribution. This corresponds to approximately 68% of the area under the curve.

3.25 The exceptional case arises when one (or more) inputs to the combined uncertainty is dominant; in this circumstance, to varying degrees the resulting distribution resembles that of the dominant contribution(s).

NOTE 1: If the dominant contribution is normal, then clearly the resulting distribution will also be normal.

NOTE 2: The above statement and note may not be true when the measurement model is non-linear.

3.26 Whenever input uncertainties are expressed in terms of limit values (e.g., limits of a rectangular distribution) rather than standard deviations, some processing is needed to 'standardise' them to obtain $u(x_i)$, as described below.

3.27 When it is possible to assess *only* the upper and lower bounds of an error (as in the case of digital rounding) a rectangular probability distribution should be assumed for the uncertainty associated with this error. Then, if a_i is the semi-range limit, the standard uncertainty is given by $u(x_i) = \frac{a_i}{\sqrt{3}}$.

Table 1 gives the expressions for various situations.

Table 1 – Expression used to obtain the standard uncertainty for various probability distributions

Assumed probability distribution	Expression used to obtain the standard uncertainty	Comment or example
Rectangular	$u(x_i) = \frac{a_i}{\sqrt{3}}$	A digital thermometer gives readings to one decimal place, that is they are expressed to within 0.1 °C. The numeric rounding caused by finite resolution will have semi-range limits of 0.05 °C. Thus, the corresponding standard uncertainty will be $u(x_i) = \frac{a_i}{\sqrt{3}} = \frac{0.05\ ℃}{1.732} = 0.029\ ℃$
U-shaped	$u(x_i) = \frac{a_i}{\sqrt{2}}$	A mismatch uncertainty associated with the calibration of an RF power sensor has been evaluated as having semi-range limits of 1.3 %. Thus, the corresponding standard uncertainty will be $u(x_i) = \frac{a_i}{\sqrt{2}} = \frac{1.3\ \%}{1.414} = 0.92\ \%$
Triangular	$u(x_i) = \frac{a_i}{\sqrt{6}}$	A tensile testing machine is used in a testing laboratory where the air temperature can vary randomly but relatively quickly and does not depart from the nominal value by more than 3 °C. The machine has a large thermal mass and is therefore most likely to be at the mean air temperature, with no probability of being outside the 3 °C limits. It is reasonable to assume a triangular distribution; therefore the standard uncertainty for its temperature is $u(x_i) = \frac{a_i}{\sqrt{6}} = \frac{3\ ℃}{2.449} = 1.2\ ℃$
Normal *(from repeatability evaluation)*	$u(x_i) = \frac{s}{\sqrt{n}}$	A statistical evaluation of repeatability uncertainty is obtained in terms of repeatability standard deviation s and the number of values n contributing to the reported value.
Normal *(from a calibration certificate)*	$u(x_i) = \frac{U}{k}$	A calibration certificate normally quotes an expanded uncertainty U at a specified, high coverage probability. A coverage factor, k, will have been used to obtain this expanded uncertainty from the combination of standard uncertainties. It is therefore necessary to divide the expanded uncertainty by the same coverage factor to obtain the standard uncertainty. (See 3.42)
Normal *(from a specification, e.g. a manufacturer's specification)*	$u(x_i) = \frac{\text{Tolerance}}{k}$	Sometimes specifications are quoted at a given coverage probability (historically referred to as *confidence level*), e.g., 95 % or 99 %. In such cases, a normal distribution might be assumed, and the tolerance limit is divided by the coverage factor k for the stated coverage probability. (See 3.46) For a coverage probability of 95 %, k = 2 and for a coverage probability of 99 %, k = 2.58. If a coverage probability is not stated, then a rectangular distribution should be assumed.

3.28 The quantities X_i that affect the measurand Y may not have a direct, one to one, relationship with it. There may be a scaling factor, such as a multiplicative constant or different measurement units, or Y may not vary linearly with X_i (as in the relationship between area and radius of a circle).

3.29 For example, a dimensional laboratory may use steel gauge blocks for calibration of measuring tools. A significant influence quantity is temperature. Because the gauge blocks have a significant temperature coefficient of expansion, there is an uncertainty that arises in their length due to an uncertainty in temperature

3.30 To translate the temperature uncertainty into an uncertainty in length units, it is necessary to know how *sensitive* the length of the gauge block is to temperature. In other words, a **sensitivity coefficient** is required.

In this example, the steel used in the manufacture of gauge blocks has a temperature coefficient of expansion of approximately $+11.5 \times 10^{-6}$ per °C, which provides the value for the sensitivity coefficient.

3.31 The sensitivity coefficient associated with each input estimate x_i is represented by c_i . It is the partial derivative of the model function $f(X_i)$ with respect to X_i, evaluated at the input estimates x_i. It is given by

$$c_i = \left.\frac{\partial f}{\partial X_i}\right|_{X_i = x_i} \approx \frac{\partial y}{\partial x_i}$$

In other words, it describes how the output estimate y varies with a corresponding small change in an input estimate x_i.

3.32 If the functional relationship is not well known for a given measurement system, or it cannot easily be differentiated, the sensitivity coefficients can usually be obtained by the practical approach of changing one of the input variables by a known amount, whilst keeping all other inputs constant, and noting the change in the output estimate.

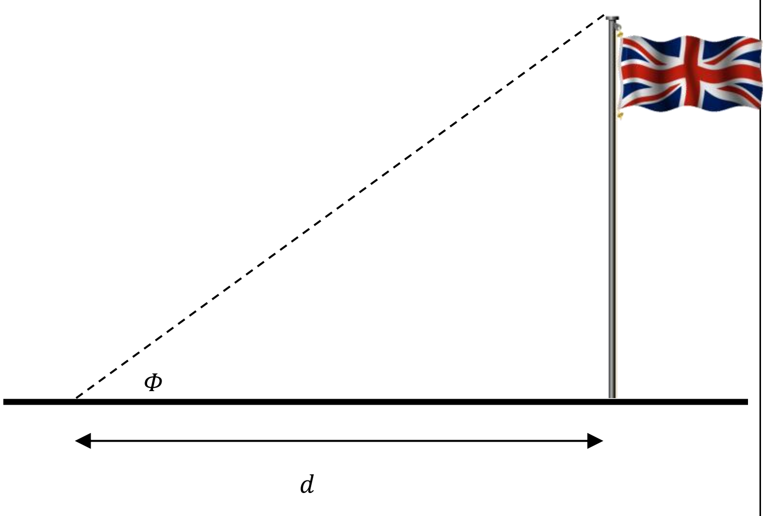
3.33 In effect, this 'numerical' approach approximates the partial derivative $\frac{\partial y}{\partial x_i}$ by the quotient $\frac{\Delta y}{\Delta x_i}$, where Δy is the change in $y = f(x_i)$ resulting from a change Δx_i in x_i. It is important to choose the magnitude of the change Δx_i around x_i carefully. It should be balanced between being sufficiently large to obtain adequate numerical accuracy in Δy and sufficiently small to provide a mathematically sound approximation to the partial derivative. The following example illustrates this approach.

Example

The height h of a flagpole is determined by measuring the angle obtained when observing the top of the pole at a specified distance d. Thus $h = d\ \tan \Phi$

Both h and d are in units of length but are related by $\tan \Phi$. In other words, $h = f(d, \Phi) = d\ \tan \Phi$.

If the measured distance is 7.0 m and the measured angle is 37°, the estimated height is
$h = 7.0 \tan{(37°)}\ \mathrm{m} = 5.275\ \mathrm{m}$.



3.34 If the maximum error in d is, say, $0.1\ \mathrm{m}$ then the estimate of h could be anywhere between $(7.0 - 0.1).\tan(37°)\mathrm{m}$ and $(7.0 + 0.1).\tan(37°)\mathrm{m}$, i.e., between 5.200 m and 5.350 m. So, a change of ±0.1 m in the input quantity x_i has resulted in a change of ±0.075 m in the output estimate y. The sensitivity coefficient is therefore estimated to be $c_d = \frac{0.075}{0.1} = 0.75$.

3.35 Similar reasoning can be applied to the uncertainty in the angle Φ. If the maximum error in Φ is 0.5°, then the estimate of h could be anywhere between $7.0\ \tan(36.5°)\mathrm{m}$ and $7.0\ \tan(37.5°)\mathrm{m}$, i.e., between 5.179 m and 5.371 m. A change of ±0.5° in the input quantity x_i has resulted in a change of ±0.096 m in the output estimate y. The sensitivity coefficient is therefore estimated to be
$c_\Phi = \frac{0.096\ \mathrm{m}}{0.5°} = 0.192$ metre per degree.

3.36 Once the standard uncertainties $u(x_i)$ and the sensitivity coefficients c_i have been evaluated, the uncertainties can be combined in order to give a single value of uncertainty to be associated with the estimate y of the measurand Y. That value is known as the **combined standard uncertainty** and is represented by the symbol $u_c(y)$.

NOTE: The subscript "c" in $u_c(y)$ is superfluous and can be omitted. It is retained here for consistency with the GUM [1].

3.37 The combined standard uncertainty is usually calculated from:

$$u_c(y) = \sqrt{\sum_{i=1}^{N} c_i^2 u^2(x_i)} = \sqrt{\sum_{i=1}^{N} u_i^2(y)} \qquad (1)$$

where

$u_i(y) = |c_i| u(x_i)$

is the standard uncertainty corresponding to the i^{th} input quantity, expressed in terms of the measurand.

NOTE: Equation (1) only applies when all x_i are independent otherwise GUM equation 13 should be used.

3.38 In other words, the individual standard uncertainties, expressed in terms of the measurand, are squared; these squared values are summed, and the square root is taken.

3.39 An example of this process is presented below, using the data from the measurement of the flagpole described above. For the purposes of the example, it is assumed that the repeatability of the process has been evaluated by making repeated measurements of the flagpole height, giving an estimated standard deviation of the mean of 0.05 metres. See Section 4 for further details about the evaluation of repeatability.

Note that there is no standardised format for presenting the content of an uncertainty budget and many variations will be encountered in practice. For example. In this table, to save on space, the standard uncertainties $u(x_i)$ have not been separately evaluated and reported. Instead, all calculations are performed in a single stage which is summarised in the final column.

Source of uncertainty	Uncertainty	Probability distribution	Divisor	Sensitivity coefficient, c_i	Standard uncertainty $u_i(y)$
Distance from flagpole	$0.1\ \mathrm{m}$	Rectangular	$\sqrt{3}$	0.75	$\frac{0.1\ \mathrm{m}}{\sqrt{3}} 0.75 = 0.0433\ \mathrm{m}$
Angle measurement	$0.5°$	Rectangular	$\sqrt{3}$	$0.192\ \mathrm{m}/°$	$\frac{0.5°}{\sqrt{3}} 0.192\ \mathrm{m}/°$ $= 0.0554\ \mathrm{m}$
Measurement Repeatability	$0.05\ \mathrm{m}$	Normal	1	1	$\frac{0.05\ \mathrm{m}}{1} 1 = 0.05\ \mathrm{m}$
Combined standard uncertainty $u_c(y) = \sqrt{0.0433^2 + 0.0554^2 + 0.05^2}$					$= 0.0863\ \mathrm{m}$

NOTE 1: The columns headed “Uncertainty” and “Probability distribution” represent the known information about the corresponding input.
The term “Uncertainty” is used here in a general sense and might, as in the case of the first two terms, correspond to the ‘half-width’ for a range of possible values (e.g., for a range $\pm 0.1\ \mathrm{m}$, the half-width is $0.1\ \mathrm{m}$). In the case of the final input for this example (measurement repeatability) it represents a standard deviation.
The “Probability distribution” summarises the nature of the information known about the respective inputs and, in association with the “Uncertainty” information determines the relevant “Divisor”. In this example the ‘Rectangular’ distributions reflect the lack of all information other than the limit values
The “Divisor” serves to *standardize* the information to establish the standardised input uncertainty $u(x_i)$.

NOTE 2: As is the case for all uncertainty evaluations, the combined standard uncertainty is a consequence of applying GUM principles to a measurement model. In this example the model is $h = d\ \tan\Phi + \delta h_r$. See Appendix D Measurement Equations or references [2] and [3] for a detailed explanation of processes for establishing a measurement model.

3.40 In accordance with the Central Limit Theorem, the PDF for y is a normal distribution with standard deviation equal to $u_c(y)$, as illustrated in Figure 5.

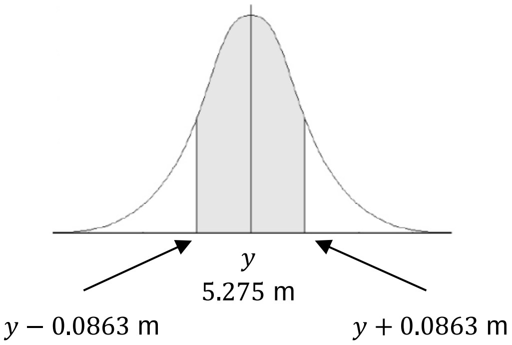


Figure 5
The measured value y is at the centre of a normal distribution with a standard deviation equal to $u_c(y)$. The numerical values relate to the example discussed above.

3.41 For a normal distribution, ± one standard deviation encompasses 68.3 % of the area under the curve. This means that there is about 68 % probability that the measurand lies within these limits.

3.42 The GUM recognises the need for providing a coverage interval with a higher coverage probability and achieves this by defining the coverage interval in terms of **expanded uncertainty**, U, which is obtained by multiplying the combined standard uncertainty by a **coverage factor**. The coverage factor is given the symbol k, thus the expanded uncertainty is given by

$$U = k\,u_{\mathrm{c}}(y) \qquad (2)$$

Where necessary to avoid ambiguity, subscripts can be attached to both U and k to reflect the corresponding coverage probability, e.g., U_p, k_p, $U_{95\%}$, $k_{95\%}$, ...

3.44 In accordance with generally accepted international practice, it is recommended that a coverage factor of $k = 2$ is used to calculate the expanded uncertainty. This value of k will give a coverage probability of approximately 95 %, assuming a normal distribution.

NOTE: A coverage factor of $k = 2$ provides a coverage probability of 95.45 % for a normal distribution. For convenience this is approximated to 95 % (which actually corresponds to a coverage factor of $k = 1.96$). However, the difference is not generally found to be significant when model assumptions and the reliability of input quantities is taken into consideration.

3.45 For example:
The measurement of the height of the flagpole had a combined standard uncertainty $u_{\mathrm{c}}(y)$ of 0.0863 m. Hence the expanded uncertainty
$U = k\,u_{\mathrm{c}}(y) = 2 \times 0.0863\ \mathrm{m}\ =\ 0.17\ \mathrm{m}.$

3.46 There may however be situations where a different coverage probability is required. For example, in safety-critical situations a higher coverage probability may be more appropriate. The table below gives the coverage factor necessary to obtain various levels of coverage for a normal distribution.

Coverage probability	Coverage factor
p	k
90 %	1.64
95 %	1.96
95.45 %	2.00
99 %	2.58
99.73 %	3.00

3.47 There may also be situations where a normal distribution cannot be assumed, and a different coverage factor may be needed to obtain a coverage probability of approximately 95 %. Such situations are described in Appendix B and Appendix C.

3.48 It is sometimes the case that a laboratory may wish to overestimate input values as a form of 'safety factor' or may wish to apply a 'comfort factor' to the calculated value of the expanded uncertainty.
This is not consistent with the principles of the GUM or M3003. If the principles of this document are followed when constructing the uncertainty budget, the resulting expanded uncertainty should be a realistic estimate of uncertainty. If not, it will be incorrect to claim, for example, that the coverage probability of an associated coverage interval is 95 %.

Decisions about 'safety', i.e., risk associated with measurement results, can only reliably be made after obtaining a realistic estimate of the measurement uncertainty (not beforehand by manipulating or applying a safety factor to the inputs) and by applying the correct coverage factor for the stated or desired coverage probability.

4 Type A evaluation of standard uncertainty

4.1 If an uncertainty is evaluated by statistical analysis of a series of observations, it is known as a Type A evaluation.

4.2 A Type A evaluation will normally be used to obtain a value for the repeatability uncertainty of a measurement process. For some measurements, this ‘random’ component of uncertainty may not be significant in relation to other contributions to uncertainty. It is nevertheless desirable for any measurement process that the relative importance of random effects be established.

When there is a spread in a sample of measurement results, the arithmetic mean (average) of the results should be calculated. If there are m independent repeated values for a quantity, Q then the mean value $\bar{q}$ is given by:

$$\bar{q} = \frac{1}{m}\sum_{j=1}^{m} q_j = \frac{q_1 + q_2 + \cdots + q_m}{m} \quad (3)$$

4.3 The values obtained (q_j) are considered to be a random, finite sample arising from a measurement process whose underlying variability is characterised by a standard deviation σ.

It is instructive to ask - if we repeated the set of measurements… would we obtain the same mean value? I.e., would we get the same value for $\bar{q}$? This seems unlikely (except in the case when the measurements are limited by poor resolution). In most cases, we would actually see a distribution of values for $\bar{q}$.
For samples of size n the standard deviation of the distribution of these sample means is $\sigma/\sqrt{n}$, known as the standard deviation of the mean (sometimes referred to as the standard error).

4.4 In practice however, it is not usually possible to obtain the value of σ and an estimate s is instead used, thus standard repeatability uncertainty is

$$u_{\mathrm{rep}} = \frac{s}{\sqrt{n}} \quad (4)$$

4.5 The dataset used to evaluate the estimate $\bar{q}$ could be used to obtain an estimate s for the standard deviation σ using

$$s = s(q_j) = \sqrt{\frac{1}{(m-1)}\sum_{j=1}^{m}(q_j - \bar{q})^2} \quad (5)$$

4.6 Example: Four measurements were made to estimate the value of a quantity q and the repeatability for the value. The results obtained were 3.42, 3.88, 2.99 and 3.17.
The mean value,

$$\bar{q} = \frac{1}{m}\sum_{j=1}^{m} q_j = \frac{3.42 + 3.88 + 2.99 + 3.17}{4} = 3.365$$

The estimated standard deviation,

$$s = s(q_j) = \sqrt{\frac{1}{(m-1)}\sum_{j=1}^{m}(q_j - \bar{q})^2} = 0.386$$

The repeatability uncertainty is therefore,

$$u_{\text{rep}} = \frac{s}{\sqrt{n}} = \frac{0.386}{\sqrt{4}} = 0.193$$

where in this case $n = m$.

4.7 It may not always be practical or possible to repeat the measurement many times during a test or a calibration. In these cases, a more reliable estimate of the standard deviation may sometimes be obtainable from previously obtained data, based on a larger number of readings.

This approach must be treated with caution – it relies on the reliability of the previously obtained data to represent the variation in the present measurements, i.e., it assumes that the underlying standard deviation σ is the same in both cases. A previous estimate of standard deviation can only be used if there has been no subsequent change in the measuring system or procedure that could have an effect on the repeatability. If an apparently excessive spread in measurement values is found, the cause should be investigated and resolved before proceeding further.

4.8 Example: Suppose that two measurements were made to estimate the value of a quantity x, i.e., $n = 2$.
However, the repeatability is to be estimated from $m = 20$ previously obtained measurements, with standard deviation

$$s = 0.247$$

The repeatability uncertainty is therefore,

$$u_{\text{rep}} = \frac{s}{\sqrt{n}} = \frac{0.247}{\sqrt{2}} = 0.175$$

NOTE: The degrees of freedom under such circumstances are $m-1$, where m is the number of measurements in the prior evaluation. Indeed, this is the reason that a large number of readings in a prior evaluation can give a more reliable estimate when only a few measurements can be made during the routine procedure. Degrees of freedom are discussed further in Appendix B.

5 Type B evaluation of standard uncertainty

5.1 If an uncertainty is evaluated by non-statistical analysis, it is known as a Type B evaluation.

5.2 Type B measurement uncertainties usually arise from fixed but unknowable (or poorly known) measurement errors. In evaluating the components of uncertainty, it is necessary to consider and include at least the following possible sources:

(a) For measuring instruments - the imported uncertainties associated with their calibration and any drift or instability in their values or readings.

(b) The reported uncertainty assigned to reference materials and any drift or instability in their values.

(c) Effects arising from the use of ancillary equipment, including items such as connecting leads, pipework, heaters etc., and any drift or instability in their values or readings.

(d) The equipment or item being measured, for example its resolution and any instability during the measurement. It should be noted that the anticipated long-term performance of an item being calibrated is not normally included in the uncertainty evaluation for that calibration.

(e) The operational procedure.

(f) The effects of environmental conditions on any or all of the above.

5.3 Having identified all the possible Type B components of uncertainty based as far as possible on experimental data or on theoretical grounds, they should be characterised in terms of standard uncertainties based on assigned probability distributions. The probability distribution of an uncertainty obtained from a Type B evaluation can take a variety of forms, but it is generally acceptable to assign a well-defined distribution for which the standard uncertainty can be obtained from a simple calculation. These distributions and sample calculations are presented in paragraphs 3.15 to 3.22 and in more detail elsewhere, e.g., JCGM-101 [2].

NOTE: It is a basic feature of the GUM framework that standard uncertainty is taken as the standard deviation of the assigned probability distribution.

5.4 Whenever possible, corrections should be made for known errors revealed by calibration or other sources. (It is not possible to make corrections for (random) repeatability errors)
The convention is that an error is given a positive sign if the measured value is greater than the expected value. The correction for error therefore involves *subtracting* the error from the measured value. On occasions, to simplify the measurement process, it may be convenient to treat such an error, when it is small compared with other uncertainties (and when doing so has an insignificant effect upon the overall evaluation) as if it were a systematic uncertainty of the uncorrected error magnitude.

5.5 Measurement errors should not be confused with mistakes.
Common examples of mistakes are incorrectly applied corrections, transcription errors, and faults in software designed to control or report on a measurement process. The effects of such mistakes cannot readily be included in the evaluation of uncertainty and care is needed to avoid them.

6 Reporting of results

6.1 After the expanded uncertainty has been calculated, usually for a coverage probability of 95 %, the measured value and associated expanded uncertainty should be reported as a coverage interval
$y \pm U$ and accompanied by the following statement:

6.2 "The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements".

6.3 In cases where the procedure of Appendix B has been followed the actual value of the coverage factor should be substituted for $k = 2$ and the following statement used:

6.4 "The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = XX$, which for a *t*-distribution with $\nu_{\mathrm{eff}} = YY$ effective degrees of freedom corresponds to a coverage probability of 95.45 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements".

6.5 In circumstances where a dominant non-Gaussian Type B contribution occurs, and the procedures described in Appendix C have been followed the following statements should be used:

6.6 "The reported expanded uncertainty is based on a standard uncertainty obtained by combining a dominant Type B uncertainty with other smaller uncertainties. The standard uncertainty has been multiplied by a coverage factor $k = XX$ which corresponds to a coverage probability of 95.45 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

For the purpose of further propagation, the measurement uncertainty can be imported into subsequent uncertainty budgets in terms of two independent quantities described by:
1. a rectangular distribution with half width of $a_{\mathrm{R}} = YY$, and
2. a normal distribution with a standard uncertainty $u_{\mathrm{N}} = ZZ$"

6.7 If uncertainty is being reported as a formula, refer also to Appendix L.

6.8 For the purposes of this guide "approximately" is interpreted as meaning sufficiently close that any difference may be considered insignificant.

6.9 Uncertainties are usually expressed in units of the measurand or as relative values, for example as a percentage (%), parts per million (ppm), parts in 10^x, $\frac{\mu\mathrm{V}}{\mathrm{V}}$, etc.

6.10 Measurement uncertainties should generally be reported to two significant figures, as it is seldom justified to report more. The numerical form of the measured value in the final statement should be reported with the same number of decimal places as the measurement uncertainty.

6.11 Rounding should always be carried out at the end of the process (to avoid the effects of cumulative rounding errors).

6.12 In situations where the PDF describing inputs to an uncertainty evaluation are asymmetric, or where the measurement model is non-linear, the resulting PDF for the measurand may also be asymmetric. In such cases Monte Carlo Simulation (as described, for example, in JCGM-101 [2]) offers a more suitable approach to evaluation of measurement uncertainty and coverage intervals.

7 Step by step procedure for evaluation of measurement uncertainty

The following is a step-by-step guide to the use of this guide for the treatment of uncertainties. The left-hand column in the table gives the general case while the right-hand column indicates how this relates to example K.4 in Appendix K. Although this example describes a calibration activity, the process is quite general and applies to most measurement situations.

	General case	**Example K.4: Calibration of a weight of nominal value 10 kg of OIML Class M1**
7.1	Determine the mathematical relationship between values of the input quantities and that of the measurand: $Y = f(X_1, X_2, \ldots X_N)$ See Appendix D for details.	It will be assumed that the unknown weight, W_X, can be obtained from the following relationship: $W_X = W_S + \delta D_S + \delta I_d + \delta C + \Delta A_b + \delta W_r$
7.2	Identify all corrections that have to be applied to the results of measurements of a quantity (measurand) for the stated conditions of measurement.	It is not normal practice to apply corrections for this class of weight and the comparator has no measurable linearity error. Estimates for these values are therefore taken to be: Drift of standard mass since last calibration $\delta D_S = 0$ Effect of least significant digit resolution $\delta I_d = 0$ Comparator linearity correction $\delta C = 0$ Correction for air buoyancy $\Delta A_b = 0$ Repeatability error $\delta W_r = 0$ However, finite (i.e., non-zero) uncertainty contributions have been determined for each of these inputs.

	General case	Example K.4: Calibration of a weight of nominal value 10 kg of OIML Class M1
7.3	List components of uncertainty associated with Type B inputs, including corrections and uncorrected systematic errors treated as uncertainties. Seek prior experimental work or theory as a basis for assigning uncertainties and probability distributions to the systematic components of uncertainty. Calculate the standard uncertainty for each component of uncertainty, obtained from Type B evaluations, as in Table 1. For assumed rectangular distributions: $u(x_i) = \frac{a_i}{\sqrt{3}}$ For assumed triangular distributions: $u(x_i) = \frac{a_i}{\sqrt{6}}$ For assumed normal distributions: $u(x_i) = \frac{U}{k}$ or consult [2] if the assumed probability distribution is not covered in this publication.	Source of uncertainty / uncertainty (mg) / Distribution* W_S Calibration of std. mass / 30 / N ($k = 2$) δD_S Drift of standard mass / 30 / R δI_d Resolution effects / 10 / T δC Comparator linearity / 3 / R ΔA_b Air buoyancy / 10 / R Then, in standardised form: $u(x_1) = u(W_S) = \frac{30\ \text{mg}}{2} = 15\ \text{mg}$ $u(x_2) = u(\delta D_S) = \frac{30\ \text{mg}}{\sqrt{3}} = 17.3\ \text{mg}$ $u(x_3) = u(\delta I_d) = \frac{10\ \text{mg}}{\sqrt{6}} = 4.08\ \text{mg}$ $u(x_4) = u(\delta C) = \frac{3\ \text{mg}}{\sqrt{3}} = 1.73\ \text{mg}$ $u(x_5) = u(\Delta A_b) = \frac{10\ \text{mg}}{\sqrt{3}} = 5.77\ \text{mg}$ *see glossary for explanation of labels, N, R, T.
7.4	Use prior knowledge or make trial measurements and calculations to determine if there is to be a random component of uncertainty that is significant compared with the effect of the other components of uncertainty.	From previous knowledge of the measurement process it is known that there is a significant random component of uncertainty.
7.5	If a random component of uncertainty is significant make repeated measurements to obtain the mean from equation (3): $\bar{q} = \frac{1}{m}\sum_{j=1}^{m} q_j = \frac{q_1 + q_2 + \cdots + q_m}{m}$	Three measurements were made of the difference between the unknown weight and the standard weight, from which the mean difference was calculated: $\overline{W}_S = \frac{0.015 + 0.025 + 0.020}{3} = 0.020\ \text{g}$

	General case	Example K.4: Calibration of a weight of nominal value 10 kg of OIML Class M1
7.6	If correlation is suspected use the guidance in paragraph D.3 or consult other referenced documents.	None of the input quantities is considered to be correlated to any significant extent; therefore Equation (1) can be used to calculate the combined standard uncertainty
7.7	Either calculate the standard deviation of the mean value from equations (4) and (5): $$s = s(q_j) = \sqrt{\frac{1}{(m-1)}\sum_{j=1}^{m}(q_j - \bar{q})^2}$$ $$u_{\mathrm{rep}} = \frac{s}{\sqrt{n}} = \frac{s(q_j)}{\sqrt{n}}$$ or refer to the results of previous repeatability evaluations with an estimate s_p based on a larger number of readings: $s = s_p$ $u_{\mathrm{rep}} = \frac{s}{\sqrt{n}} = \frac{s_p}{\sqrt{n}}$ where m is the number of readings used in the evaluation of s and n is the number of readings that contribute to the evaluation of the mean value.	A previous Type A evaluation had been made to determine the repeatability of the comparison using the same type of 10 kg weights. The standard deviation was determined from $m = 10$ measurements using the conventional bracketing technique and was calculated, to be $s(\delta W_r) = 8.7 \text{ mg}$ Since the number of determinations taken when calibrating the unknown weight was 3 this is the value of n that is used to calculate the standard deviation of the mean using equation (4): $u(x_6) = u_{\mathrm{rep}} = \frac{s(\delta W_r)}{\sqrt{n}} = \frac{8.7}{\sqrt{3}} = 5.0 \text{ mg}$

	General case	Example K.4: Calibration of a weight of nominal value 10 kg of OIML Class M1
7.8	Calculate the combined standard uncertainty for uncorrelated input quantities using equation (1) if absolute values are used: $$u_c(y) = \sqrt{\sum_{i=1}^{N} c_i^2 u^2(x_i)} = \sqrt{\sum_{i=1}^{N} u_i^2(y)}$$ where c_i is the partial derivative $c_i = \left.\frac{\partial f}{\partial X_i}\right\|_{X_i = x_i} \approx \frac{\partial y}{\partial x_i}$, or a known sensitivity coefficient. Alternatively use equation (8) if the standard uncertainties are all relative values and $Y = f(X_i)$ is a pure product of terms $$\frac{u_c(y)}{y} = \sqrt{\sum_{i=1}^{N} \left[\frac{p_i u(x_i)}{\|x_i\|}\right]^2}$$ where p_i are known positive or negative exponents in the functional relationship.	The units of all standard uncertainties are in terms of those of the measurand, i.e., milligrams, and the functional relationship between the input quantities and the measurand is a linear summation; therefore, all the sensitivity coefficients are unity ($c_i = 1$). Applying equation (1) gives $u_c(W_X) = \sqrt{15^2 + 17.32^2 + 4.08^2 + 1.73^2 + 5.77^2 + 5^2}$ $u_c(W_X) = 25 \text{ mg}$
7.11	If there is a dominant Type A contribution (i.e., a significant value obtained from a small number of readings), use Appendix B to calculate a suitable alternative value for coverage factor k and use this value to calculate the expanded uncertainty. Similarly, if there is a dominant Type B contribution, use Appendix C to calculate a suitable alternative value for coverage factor k and use this value to calculate the expanded uncertainty. Then, calculate the expanded uncertainty $U_p = k_p\, u_c(y)$	$U_{95\%} = 2 \times 24.55 \text{ mg} = 49 \text{ mg}$ A coverage probability of approximately 95 % is obtained with $k = 2$. It was not necessary to use an alternative value for k as there are no dominant Type A or Type B contributions. NOTE: The evaluation includes a single Type A component for which the standard deviation, $s(\delta W_r)$ is based upon $m = 10$ values, for which there are $\upsilon_{\text{rep}} = (m-1) = 9$ degrees of freedom. All other components are Type B, for which $\upsilon = \infty$, therefore $\upsilon_{\text{eff}} = \frac{u_c^4(y)}{\frac{u_{\text{rep}}^4(y)}{(m-1)}} = 5145$, i.e., the Type A input is not dominant Also, the r-value for the largest Type B component is 0.99, much less that a threshold value of 1.42 for dominant Type B
7.12	Report the result and the expanded uncertainty in accordance with Section 6.	The measured value of the 10 kg weight is $10\,000.025 \text{ g} \pm 0.049 \text{ g}$. The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

Appendix A Calibration and Measurement Capability

A.1 The Calibration and Measurement Capability (CMC), defines the measurement capabilities, ranges, and boundaries of a calibration activity. In particular, it defines the lowest measurement uncertainty that can be achieved during a calibration under normal conditions.

A.2 For an accredited calibration laboratory, the CMC is described in the Schedule of Accreditation. The associated measurement uncertainty relates to the calibration of real items for which the laboratory has been accredited, using processes that were the subject of assessment.

A.3 The measurement uncertainty is calculated according to the procedures described in this guide, and in the GUM [1], and is normally stated as an expanded uncertainty at a coverage probability of 95 %, which usually requires a coverage factor of $k = 2$.

A.4 An accredited laboratory is not permitted to quote a smaller uncertainty in certificates issued under its accreditation but may report an equal or larger uncertainty if necessary. For example, if a particular item under calibration itself contributes significantly to the uncertainty (e.g., through limited resolution or significant non-repeatability) then the uncertainty reported on a calibration certificate will naturally be larger than the CMC uncertainty to account for such factors.

A.5 It is sometimes the case that a laboratory may request to be accredited for a CMC uncertainty that is larger than it can achieve and may wish to report values that are larger than are obtained by its uncertainty evaluations.
The implication is that the laboratory is in some way uncomfortable about the magnitude of the expanded uncertainty. If this is the case, the measurement model and contributions to the uncertainty budget should be reviewed and amended if justified. However, neither the inputs nor the output should simply be inflated (see 3.4.8)

A.6 Refer to ILAC-P14 [11] and to UKAS LAB 45 [15] for further explanation of Calibration and Measurement Capability.

NOTE: The term CMC also applies to the measurement capabilities of National Metrology Institutes that are published in the BIPM key comparison database (KCDB) of the CIPM MRA.

Appendix B Coverage factor when there is a single dominant Type A input

B.1 In some cases, it may not be practical to base a Type A evaluation on a large number of readings. In these situations, if it is also the case that the Type A input is a significant part of the combined uncertainty, this could result in the coverage probability being significantly less than 95 % if a coverage factor of $k = 2$ is used. In these situations, the value of k, or more precisely k_p, where p is the confidence probability, should be based on a ***t*-distribution** rather than a normal distribution. This value of k_p will give an expanded uncertainty, U_p, that maintains the coverage probability at approximately the required level p.

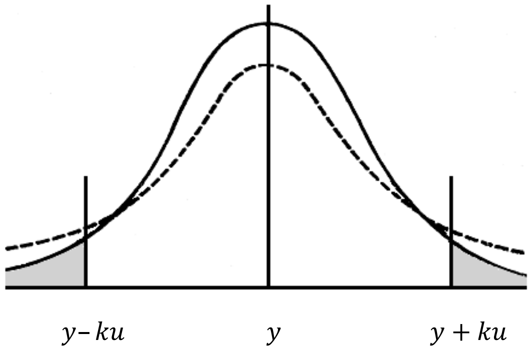


Figure 6

In Figure 6, the solid line depicts a normal distribution with standard deviation (standard uncertainty) u.

A specified proportion p of the values under the curve are encompassed by the interval between $y - ku$ and $y + ku$.

An example of the *t*-distribution is superimposed, using dashed lines. For the *t*-distribution, a greater proportion of the values lies outside the interval $y - ku$ to $y + ku$, and a smaller proportion lies inside this region. An increased value of k is therefore required to restore the original coverage probability. This new coverage factor, k_p, is obtained by evaluating the effective degrees of freedom of $u_c(y)$ and obtaining the corresponding value t_p, e.g., from a *t*-distribution table. The required coverage factor is then $k_p = t_p$

B.2 In order to obtain a value for k_p it is necessary to obtain an estimate of the effective degrees of freedom, υ_{eff} , of the combined standard uncertainty $u_c(y)$. The GUM [1] recommends that the Welch-Satterthwaite equation is used to calculate a value for υ_{eff} based on the degrees of freedom, υ_i , of the individual standard uncertainties $u_i(y)$; where

$$\upsilon_{\text{eff}} = \frac{{u_c}^4(y)}{\sum_{i=1}^{N} \frac{{u_i}^4(y)}{\upsilon_i}} \qquad (6)$$

B.3 The degrees of freedom, υ_i, for contributions obtained from Type A evaluations are $m - 1$, where m is the number of values used to evaluate $u_i(y)$.

B.4 It is often possible to take the degrees of freedom, υ_i, of Type B uncertainty contributions as being infinite, that is, their value is known with a very high degree of reliability. If this is the case then the calculation simplifies, as all the terms relating to the Type B uncertainties become zero. This case is illustrated in the example in paragraph B.10.

B.5 However, it is possible for a Type B contribution to come from a calibration certificate, in the form of an uncertainty based on a *t*-distribution (as is described in this Appendix) rather than a normal distribution. This is an example of a Type B contribution that does not have infinite degrees of freedom. In this eventuality the degrees of freedom will be as quoted on the calibration certificate.

B.6 Having obtained a value for υ_{eff} , the corresponding value from the *t*-distribution can be obtained, either from tables (such as the table below), or by calculation.

B.7 Unless otherwise specified, the values corresponding to a coverage probability of $p = 95.45\ \%$ should be used.

B.8 Normally υ_{eff} will not be an integer and, when using tabulated data, it will be necessary to interpolate between the values given in the table. Linear interpolation will suffice for $\upsilon_{\mathrm{eff}} > 3$ in the table provided below, higher-order interpolation should be used otherwise, or else the next lower value of υ_{eff} may be used.

To calculate t_p directly the Excel spreadsheet function T.INV.2T$(1-p, \upsilon_{\mathrm{eff}})$ can be used.

B.9 The required coverage factor is then $k_p = t_p$. This is the coverage factor required to calculate the expanded uncertainty, U_p, from $U_p = k_p\, u_{\mathrm{c}}(y)$.
Unless otherwise specified, the coverage probability p will usually be 95.45 %.

B.10 <u>Example</u>

B.10.1 In a measurement system a Type A evaluation, based on 4 observations, gave a value of $u_i(y)$ of 3.5 units. There were 5 other contributions all based on Type B evaluations for each of which infinite degrees of freedom had been assumed. The combined standard uncertainty, $u_{\mathrm{c}}(y)$, had a value of 5.7 units.
Then, using the <u>Welch-Satterthwaite equation</u>:

$$\upsilon_{\mathrm{eff}} = \frac{5.7^4}{\frac{3.5^4}{4-1} + 0 + 0 + 0 + 0 + 0} = \frac{5.7^4}{3.5^4} \times 3 = 21.1$$

B.10.2 In the <u>*t*-distribution table</u>, the value of υ_{eff} for a coverage probability $p = 95.45\ \%$, immediately lower than 21.1 is 20. This gives a value for k_p of 2.13 and this is the coverage factor that should be used to calculate the expanded uncertainty…
The expanded uncertainty is $5.7 \times 2.13 = 12.14$ units.

Degrees of freedom, υ	Values of $t_p(\upsilon)$ from the t-distribution for υ degrees of freedom					
	$p = 68.27\,\%$	$p = 90\,\%$	$p = 95\,\%$	$p = 95.45\,\%$	$p = 99\,\%$	$p = 99.73\,\%$
1	1.84	6.31	12.71	13.97	63.66	235.78
2	1.32	2.92	4.30	4.53	9.92	19.21
3	1.20	2.35	3.18	3.31	5.84	9.22
4	1.14	2.13	2.78	2.87	4.60	6.62
5	1.11	2.02	2.57	2.65	4.03	5.51
6	1.09	1.94	2.45	2.52	3.71	4.90
7	1.08	1.89	2.36	2.43	3.50	4.53
8	1.07	1.86	2.31	2.37	3.36	4.28
9	1.06	1.83	2.26	2.32	3.25	4.09
10	1.05	1.81	2.23	2.28	3.17	3.96
11	1.05	1.80	2.20	2.25	3.11	3.85
12	1.04	1.78	2.18	2.23	3.05	3.76
13	1.04	1.77	2.16	2.21	3.01	3.69
14	1.04	1.76	2.14	2.20	2.98	3.64
15	1.03	1.75	2.13	2.18	2.95	3.59
16	1.03	1.75	2.12	2.17	2.92	3.54
17	1.03	1.74	2.11	2.16	2.90	3.51
18	1.03	1.73	2.10	2.15	2.88	3.48
19	1.03	1.73	2.09	2.14	2.86	3.45
20	1.03	1.72	2.09	2.13	2.85	3.42
25	1.02	1.71	2.06	2.11	2.79	3.33
30	1.01	1.70	2.04	2.09	2.75	3.27
35	1.01	1.70	2.03	2.07	2.72	3.23
40	1.01	1.68	2.02	2.06	2.70	3.20
45	1.01	1.68	2.01	2.06	2.69	3.18
50	1.01	1.68	2.01	2.05	2.68	3.16
100	1.005	1.660	1.984	2.025	2.626	3.077
∞	1.000	1.645	1.960	2.000	2.576	3.000

Appendix C Coverage factor when there is a single dominant Type B input

C.1 It is quite common in some measurement processes, particularly calibrations, for there to be a component of uncertainty derived from a Type B evaluation that is dominant in magnitude compared all other components. In these circumstances it can no longer be assumed that the Central Limit Theorem applies (resulting in a normal distribution for the output, y.)

C.2 A commonly encountered example arises from the resolution of a digital indicating instrument. This will have a rectangular distribution which, if the half-width a is large, may dominate the shape of the distribution for output y. Consequently, the properties of the normal distribution can no longer be used to establish the coverage factor for a required coverage probability.

C.3 As calibration results are required to be presented in terms of a coverage interval, a different approach must be taken to obtain a suitable coverage factor. A method based upon the use of tabulated values is described below.

C.4 The presence of a dominant *rectangular* input quantity can be detected by a 'rule of thumb'… for a 95.45 % coverage probability, it will 'dominate' (i.e., the error if using a coverage factor of $k = 2$ will be more than 5 %) if its standard uncertainty $u_{\mathrm{R}}(y)$ is more than about 1.42 times the combined standard uncertainty $u_{\mathrm{N}}(y)$ for the remaining $N - 1$ inputs (assuming that these result in a normal distribution).

I.e., a rectangular term 'dominates' when $\frac{u_{\mathrm{R}}(y)}{u_{\mathrm{N}}(y)} \gtrsim 1.42$, or inversely $\frac{u_{\mathrm{N}}(y)}{u_{\mathrm{R}}(y)} \lesssim 0.70$

where

$$u_{\mathrm{R}}(y) = c_{\mathrm{R}} \frac{a}{\sqrt{3}} = \frac{a_{\mathrm{R}}}{\sqrt{3}}$$

which defines effective half width $a_{\mathrm{R}} = c_{\mathrm{R}} a$,

and (*excluding* the dominant term)

$$u_{\mathrm{N}}(y) = \sqrt{\sum_{i=1}^{N-1} u_i^2(y)}$$

which as expected combine to give

$$u_c(y) = \sqrt{u_{\mathrm{R}}^2(y) + u_{\mathrm{N}}^2(y)}$$

C.5 <u>Example</u>
A digital voltmeter is calibrated with an applied reference voltage V_{ref} of 1.00000 V; the resulting indication is 1.001 V.
The expanded uncertainty of the applied voltage is 0.19 mV ($k = 2$). A normal distribution is assumed.
The indicator can display readings in steps of 0.001 V; therefore, there will be a possible maximum rounding error of ±0.5 mV. A rectangular probability distribution is assumed.
There are some minor drift and temperature effects, characterised by a limit of ±50 μV, for which a rectangular probability distribution is assumed.
There are no other significant influences.

The standard uncertainty associated with the potentially dominant rectangular term is

$$u_{\mathrm{R}}(y) = c_R \frac{a}{\sqrt{3}} = 1 \times \frac{0.5\ \mathrm{mV}}{\sqrt{3}} = 0.289\ \mathrm{mV}$$

The standard uncertainty associated with the remaining terms is

$$u_{\mathrm{N}}(y) = \sqrt{\sum_{i=1}^{N-1} u_i^2(y)} = \sqrt{\left(1 \times \frac{0.19}{2}\right)^2 + \left(1 \times \frac{0.05}{\sqrt{3}}\right)^2}\ \mathrm{mV} = 0.099\ \mathrm{mV}$$

The value of the corresponding ratio is $\frac{u_{\mathrm{N}}(y)}{u_{\mathrm{R}}(y)} = 0.344$

The value is less than 0.70 which indicates that the term dominates, and a normal distribution cannot be assumed for y

C.6 A suitable coverage factor can be obtained from tables of factors derived from a convolution between the dominant distribution and the distribution assumed for the combined remainder of the terms (usually a normal distribution)

C.7 If a rectangular distribution and a normal distribution are convolved, the coverage factor k for a coverage probability of 95.45 % may be obtained from the following table:

$\frac{u_{\mathrm{N}}(y)}{u_{\mathrm{R}}(y)}$	$k_{95.45\%}$	$\frac{u_{\mathrm{N}}(y)}{u_{\mathrm{R}}(y)}$	$k_{95.45\%}$	$\frac{u_{\mathrm{N}}(y)}{u_{\mathrm{R}}(y)}$	$k_{95.45\%}$
0.00	1.65	0.50	1.84	0.95	1.95
0.10	1.66	0.55	1.85	1.00	1.95
0.15	1.68	0.60	1.87	1.10	1.96
0.20	1.70	0.65	1.89	1.20	1.97
0.25	1.72	0.70	1.90	1.40	1.98
0.30	1.75	0.75	1.91	1.80	1.99
0.35	1.77	0.80	1.92	2.00	1.99
0.40	1.79	0.85	1.93	2.50	2.00
0.45	1.82	0.90	1.94	∞	2.00

C.8 For this example the coverage factor $k = 1.77$,

the standard uncertainty is

$$u_{\mathrm{c}}(y) = \sqrt{u_{\mathrm{R}}^2(y) + u_{\mathrm{N}}^2(y)} = \sqrt{0.289^2 + 0.099^2}\ \mathrm{mV} = 0.305\ \mathrm{mV}$$

and the corresponding expanded uncertainty is

$$U = k\, u_{\mathrm{c}}(y) = 1.77 \times 0.305\ \mathrm{mV} = 0.54\ \mathrm{mV}$$

C.9 This approach provides a suitable coverage factor for establishing the coverage interval, however for the purposes of propagation of uncertainties it is more helpful to describe the uncertainty in two parts - corresponding to the dominant distribution, and the distribution representing the balance of the terms (usually taken as a normal distribution)

"The reported expanded uncertainty is based on a standard uncertainty obtained by combining a dominant Type-B uncertainty with other smaller uncertainties. The standard uncertainty has been multiplied by a coverage factor $k = XX$ which, corresponds to a coverage probability of 95.45 %.

For the purpose of further propagation, the measurement uncertainty can be imported into subsequent uncertainty budgets in terms of two independent quantities described by:
1. a rectangular distribution with half width of $a_{\mathrm{R}} = YY$, and
2. a normal distribution with a standard uncertainty $u_{\mathrm{N}} = ZZ$"

C.10 The example described at C.5 can be summarised and reported as follows

$\Delta V = V_{\mathrm{ind}} - V_{\mathrm{ref}} + \delta I_T$

Quantity	Source of uncertainty	Value / mV	Probability distribution	Divisor	c_i	$u_i(\Delta V)$ / mV
V_{ref}	Applied voltage	0.19	Normal	2	1	0.095
δI_T	Drift and temperature effects	0.05	Rectangular	$\sqrt{3}$	1	0.029
V_{ind}	Digital rounding of indicator	0.5	Rectangular	$\sqrt{3}$	1	0.289
$u_c(\Delta V)$	Combined standard uncertainty		Convolved $\frac{0.99}{0.289} = 0.344$			0.305
$U(\Delta V)$	Expanded uncertainty		Convolved $k = 1.77$			0.54

Reported result:

For an applied voltage of $1.00000\ \mathrm{V}$ the voltmeter reading was $1.001\ \mathrm{V}$

The measurement error was $\Delta V = V_{\mathrm{ind}} - V_{\mathrm{ref}} = 0.00100\ \mathrm{V} \pm 0.54\ \mathrm{mV}$

The reported expanded uncertainty is based on a standard uncertainty obtained by combining a dominant Type B uncertainty with other smaller uncertainties. The standard uncertainty has been multiplied by a coverage factor $k = 1.77$ which corresponds to a coverage probability of 95.45 %.

For the purpose of further propagation, the measurement uncertainty can be imported into subsequent uncertainty budgets in terms of two independent quantities described by:
1. a rectangular distribution with half width of $a_{\mathrm{R}} = 0.5$ mV, and
2. a normal distribution with a standard uncertainty $u_{\mathrm{N}} = 0.099$ mV"

The table below demonstrates how this uncertainty might be imported into a subsequent uncertainty budget for the *use* of the voltmeter, seen in this partially completed example as the entries in the first two rows:

Quantity	Source of uncertainty	Value / mV	Probability distribution	Divisor	c_i	$u_i(y)$ / mV
ΔV	Calibration error, uncertainty term 1, a_{R}	0.5	Rectangular	$\sqrt{3}$		
ΔV	Calibration error, uncertainty term 2, u_{N}	0.095	Normal	1		
$\vdots$	Other sources relating to use…	$\vdots$	$\vdots$	$\vdots$		
V_{ind}	Rounding of indication in use	0.5	Rectangular	$\sqrt{3}$		
$\vdots$	$\vdots$		$\vdots$	$\vdots$		

C.11 In the subsequent budget, the rectangular contribution that dominated in the preceding budget is less likely to be as dominant, if at all. Otherwise this process can simply be repeated for the latest budget.

C.12 The same situation may be encountered with other distributions associated with Type B uncertainties. An example is the U-shaped distribution associated with mismatch uncertainty in RF and microwave systems. Similar reasoning applies here, and the suggested coverage probability statement can be modified accordingly.

Caution:

The ‘rules of thumb for identifying a single dominant input quantity detect the situation where the error in using a coverage factor of $k = 2$ will be more than 5 %.

In applications where expanded measurement uncertainty is an input to a quantitative process, such as the calculation of guard bands for a conformity decision [14], an error of this size may be unacceptable, in which case a suitable coverage factor can be established by following this process regardless of the ‘rule of thumb’.

C.13 If a U-shaped distribution and a normal distribution are convolved, the coverage factor k for a coverage probability of 95.45 % may be obtained from the following table:

$\frac{u_{\mathrm{N}}(y)}{u_{\mathrm{U-shape}}(y)}$	$k_{95.45\%}$	$\frac{u_{\mathrm{N}}(y)}{u_{\mathrm{U-shape}}(y)}$	$k_{95.45\%}$	$\frac{u_{\mathrm{N}}(y)}{u_{\mathrm{U-shape}}(y)}$	$k_{95.45\%}$
0.00	1.41	0.50	1.77	0.95	1.93
0.10	1.47	0.55	1.80	1.00	1.93
0.15	1.51	0.60	1.82	1.10	1.95
0.20	1.55	0.65	1.84	1.20	1.96
0.25	1.60	0.70	1.86	1.40	1.97
0.30	1.64	0.75	1.88	1.80	1.99
0.35	1.67	0.80	1.89	2.00	1.99
0.40	1.71	0.85	1.90	2.50	2.00
0.45	1.74	0.90	1.92	∞	2.00

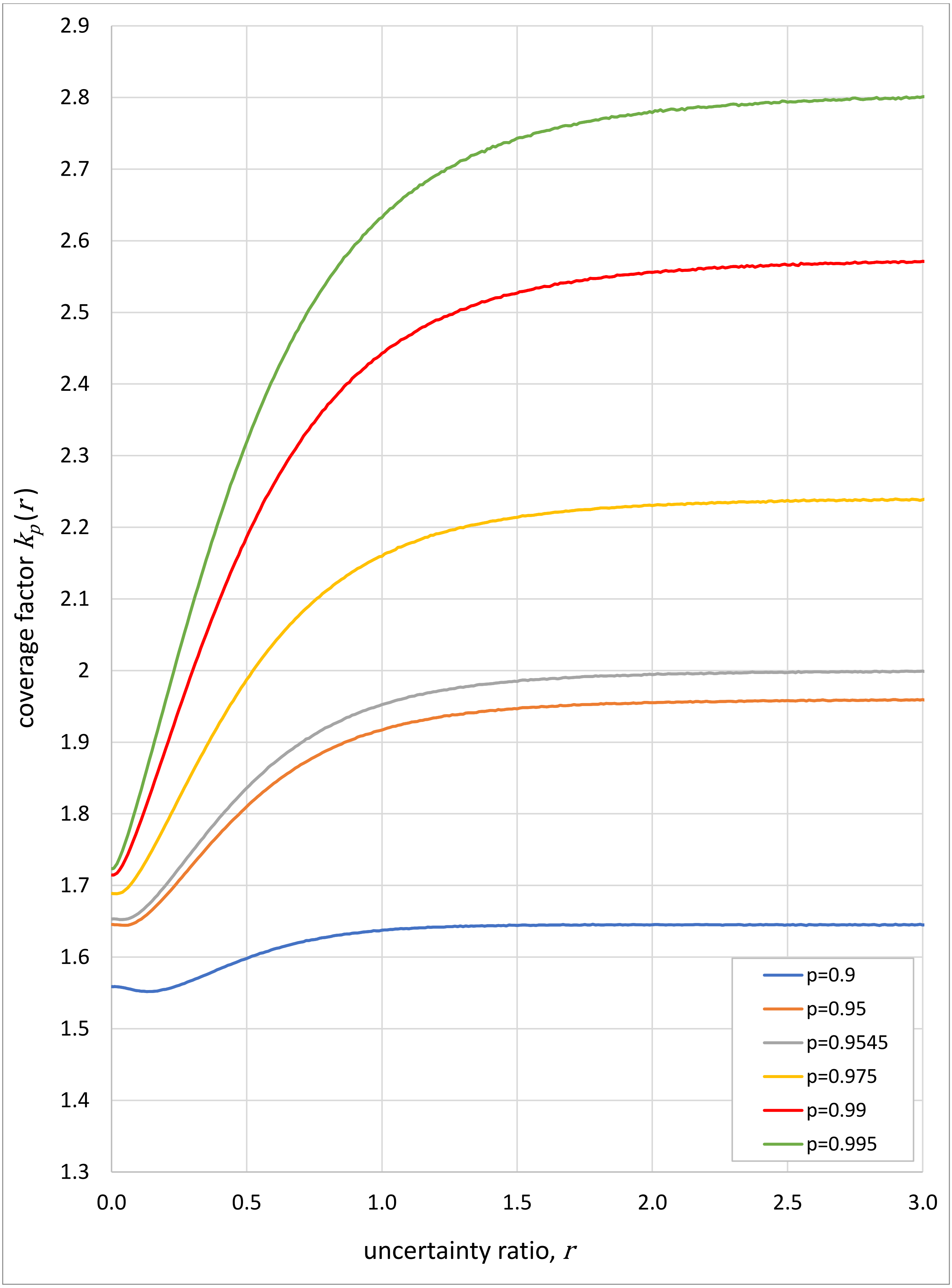
C.14 If a U-shaped distribution and a rectangular distribution are convolved, the coverage factor k for a coverage probability of 95.45 % may be obtained from the following table:

$\frac{u_{\mathrm{R}}(y)}{u_{\mathrm{U-shape}}(y)}$	$k_{95.45\%}$	$\frac{u_{\mathrm{R}}(y)}{u_{\mathrm{U-shape}}(y)}$	$k_{95.45\%}$	$\frac{u_{\mathrm{R}}(y)}{u_{\mathrm{U-shape}}(y)}$	$k_{95.45\%}$
0.00	1.41	0.45	1.75	3.0	1.80
0.10	1.48	0.50	1.78	4.0	1.75
0.15	1.53	0.60	1.82	5.0	1.72
0.20	1.57	0.70	1.86	6.0	1.70
0.25	1.62	0.80	1.88	7.5	1.68
0.30	1.66	0.90	1.89	10	1.66
0.35	1.69	1.0	1.90	20	1.65
0.40	1.73	2.0	1.86	∞	1.65

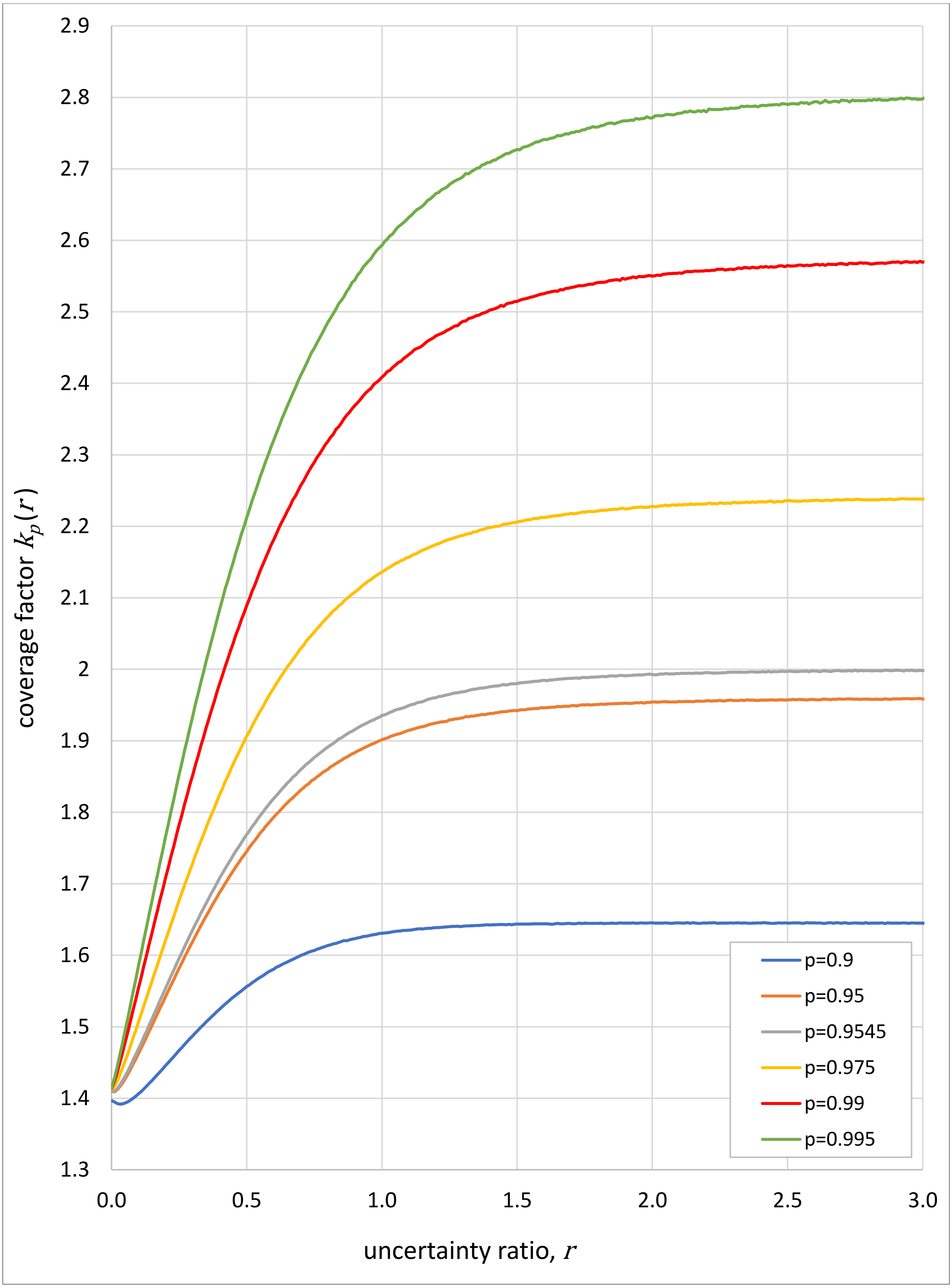
C.15 If two distributions of identical form, either rectangular or U-shaped, are convolved, the coverage factor k for a coverage probability of 95.45 % may be obtained from the following table:

$\frac{u_{\mathrm{smaller}}(y)}{u_{\mathrm{larger}}(y)}$	k for stated ratio 2 Rectangular Distributions	k for stated ratio 2 U-shaped Distributions
0.00	1.65	1.41
0.05	1.65	1.44
0.10	1.66	1.49
0.15	1.69	1.53
0.20	1.71	1.58
0.25	1.74	1.62
0.30	1.77	1.66
0.35	1.79	1.69
0.40	1.82	1.72
0.45	1.84	1.75
0.50	1.86	1.77
0.60	1.89	1.81
0.70	1.91	1.83
0.80	1.92	1.85
0.90	1.93	1.86
1.00	1.93	1.86

C.16 Coverage factors for various other coverage probabilities are shown graphically on the following pages.



Coverage factors for the combination of normal and rectangular PDFs at several different coverage probabilities for different uncertainty ratio: $r = \frac{u_N(y)}{u_R(y)}$



Coverage factors for the combination of normal and U-shaped PDFs at several different coverage probabilities for different uncertainty ratio: $r = \frac{u_{\mathrm{N}}(y)}{u_{\mathrm{U-shape}}(y)}$

Appendix D Measurement equations

D.1 Establishing a measurement equation

D.1.1 A measurement model describes the relationship between an output quantity Y and input quantities X_i.

The 'bottom-up' approach of the GUM [1] law of propagation of uncertainties, as covered in M3003, requires this to be a functional relationship, $Y = f(X_1, \ldots, X_N)$, which also relates the estimates of the input quantities x_i to the estimate of the output, $y = f(x_1, \ldots, x_N)$.
This relationship is referred to in M3003 as a measurement equation.

A common approach to establishing a measurement model [2,3] involves first defining a 'basic model' which relates the output to the 'physical' input quantities. For example, this might be an equation such as $p = \rho g h$, or $Q = \frac{\pi(P_1 - P_2)r^4}{\eta L}$, or it might be an equation defining a measurement error $\Delta I = I_{\mathrm{obs}} - I_{\mathrm{ref}}$

The equation is then extended by the addition of 'metrological' terms representing possible quantities that are not already part of the basic equation. These usually correspond to error quantities for which an exact value is unknown. The best estimate of their value (in additive terms) is zero, however each has a finite uncertainty.

D.1.2 For example, suppose that the voltage output of a transducer (the measurand) is directly measured using a calibrated meter. A basic model could be written as

$$V = V_{\mathrm{ave}} + \Delta V_{\mathrm{cal}}$$

where

V	is the estimate of the voltage output,
V_{ave}	is the average of n observations $V_{j=1\ \mathrm{to}\ n}$ (obtained under repeatability conditions of measurement),
ΔV_{cal}	is an additive correction to the observed values, established from calibration of the meter.

The estimate V is further influenced by other factors, each of which corresponds to a poorly known or unknowable measurement error δV_i (each with a best estimate of zero value, but finite uncertainty).
The measurement equation is therefore extended to account for these 'metrological terms' e.g.,

$$V = V_{\mathrm{ave}} + \Delta V_{\mathrm{cal}} + \delta V_{\mathrm{res}} + \delta V_{\mathrm{drift}} + \delta V_{\mathrm{T}}$$

where say

δV_{res}	is the error due to finite resolution of the observed values V_j
$\delta V_{\mathrm{drift}}$	is the error due to drift in the correction ΔV_{cal} since the meter was last calibrated
δV_{T}	is the error due to possible temperature effects

D1.3 There is no single standard or 'correct' way to construct a measurement model. The process is dictated by the nature of the information available and by the knowledge of the person performing the evaluation, nevertheless all valid approaches should lead to a similar result

Similarly, there is no universally accepted standard for the choice of symbols to represent different quantities in a measurement equation. Choices should be made primarily to ensure clarity of meaning.

However, a useful convention is to use a δ symbol to represent any term whose best estimated value is zero and to use a Δ symbol to represent a difference, a measurement error or correction.

D.1.4 In some cases, the measurement equation can be written as a pure product of terms (i.e., the output quantity is obtained from *only* the multiplication or division of the input quantities)

$$y = c x_1^{p_1} x_2^{p_2} \dots x_N^{p_N} \tag{7}$$

where the exponents p_i are known positive or negative numbers

For example,

$$Q = Q_{\text{ave}} \times f_{\text{cal}} \times f_{\text{res}} \times f_{\text{drift}} \times f_{\text{T}}$$

where say

f_{cal}	is a multiplicative correction established from calibration of the meter
f_{res}	is a factor representing error due to finite resolution of the observed values Q_j.
f_{drift}	is a factor representing error due to drift in the correction f_{cal} since the meter was last calibrated
f_{T}	is a factor representing error due to possible temperature effects

If the multiplicative factor is poorly known or unknowable, the best estimate of its value is one, analogous to the value of zero attributed to poorly known or unknowable values in the additive case.

D1.4.1 In this special case the uncertainty evaluation can be performed in terms of relative values (e.g., in % terms, or in parts per million) and the relative standard uncertainty will then be given by,

$$\frac{u_c(y)}{|y|} = \sqrt{\sum_{i=1}^{N} \left(\frac{p_i u(x_i)}{x_i}\right)^2} \tag{8}$$

Some examples are given below

$$P = f(I, V) = IV, \qquad \frac{u(P)}{|P|} = \sqrt{\left(\frac{u(I)}{I}\right)^2 + \left(\frac{u(V)}{V}\right)^2}$$

$$E = f(m, v) = \tfrac{1}{2} m v^2, \qquad \frac{u(E)}{|E|} = \sqrt{\left(\frac{u(m)}{m}\right)^2 + \left(\frac{2.u(v)}{v}\right)^2}$$

$$V = f(P, Z) = (PZ)^{\frac{1}{2}}, \qquad \frac{u(V)}{|V|} = \sqrt{\left(\frac{u(P)}{2P}\right)^2 + \left(\frac{u(Z)}{2Z}\right)^2}$$

Use of relative uncertainties can often simplify the calculations and is particularly helpful when the input quantities and the uncertainties are already available in relative terms.

Equation (8) should *not* be used when the functional relationship includes any addition or subtraction of quantities.

D.1.5 More generally there will be both additive and multiplicative terms in the measurement equation. For example, a measurement model might be written as:

$$Q = (Q_{\mathrm{ave}} \times f_{\mathrm{cal}}) + \delta Q_{\mathrm{res}} + \delta Q_{\mathrm{drift}} + \delta Q_{\mathrm{T}}$$

where say

Q_{ave}	is the average of n observations $Q_{j=1\ \mathrm{to}\ n}$ (obtained under repeatability conditions of measurement)
f_{cal}	is a multiplicative correction established from calibration of the meter
δQ_{res}	is the error due to finite resolution of the observed values Q_j
$\delta Q_{\mathrm{drift}}$	is the error due to drift in the correction f_{cal} since the meter was last calibrated
δQ_{T}	is the error due to possible temperature effects

D.1.6 In all cases, to evaluate the measurement uncertainty, the contribution associated with each ‘input’ quantity in the measurement equation must be considered. For uncertainty budgets that are represented in tabular format this will usually involve creating a separate line for each term

D.1.7 In situations where a suitable measurement equation cannot readily be established a ‘top-down’ approach is often adopted, resulting in a form of statistical model for the measurement. This approach is not covered in M3003. The reader is referred to other guidance documents such as EURACHEM/CITAG Guide CG4 [10] and ISO 21748 [8] for further information.

D.1.8 More detailed guidance on the development of measurement models of all types can be found in JCGM GUM-6 [3].

D.2 Measurement repeatability in measurement equations

D.2.1 Measurement repeatability can be incorporated into a model in several different ways. The choice will largely depend on how its value is to be estimated.

D.2.2 For example, the measurement equation

$$V = V_{\mathrm{ave}} + \Delta V_{\mathrm{cal}} + \delta V_{\mathrm{res}} + \delta V_{\mathrm{drift}} + \delta V_{\mathrm{T}}$$

might associate the measurement repeatability with the quantity V_{ave} .

If the budget is presented in a table, it might appear as shown in the example below:

Quantity X_i	Estimate x_i /volt	Source of uncertainty	Uncertainty /volt	PDF	Divisor	$u(x_i)$	c_{i}	$u_i(V)$ /volt
V_{ave}	100.001	Repeatability	0.0051					
ΔV_{cal}	0.001	Calibration	0.0035					
δV_{res}	0	Resolution	0.00005					
$\delta V_{\mathrm{drift}}$	0	Calibration drift	0.0001					
δV_{T}	0	Thermal effects	0.00012					

D.2.3 Alternatively, an additional error term, say δV_{rep} might be used.

$$V = V_{\mathrm{ave}} + \Delta V_{\mathrm{cal}} + \delta V_{\mathrm{res}} + \delta V_{\mathrm{drift}} + \delta V_{\mathrm{T}} + \delta V_{\mathrm{rep}}$$

in which case V_{ave} no longer features in the table in the example below, because all associated errors are taken into account by the other quantities, in effect it is treated as a constant

Quantity X_i	Estimate x_i /volt	Source of uncertainty	Uncertainty /volt	PDF	Divisor	$u(x_i)$	c_{i}	$u_i(V)$ /volt
δV_{rep}	0	Repeatability	0.0051					
ΔV_{cal}	0.001	Calibration	0.0035					
δV_{res}	0	Resolution	0.00005					
$\delta V_{\mathrm{drift}}$	0	Calibration drift	0.0001					
δV_{T}	0	Thermal effects	0.00012					

D.2.4 For a model that has several measured input quantities the repeatability can be treated as a separate input for each individual quantity. Such a treatment is most likely to be useful when the inputs are obtained separately.

For example, the outer surface area of a smooth regular cylinder can be modelled as

$$A = \pi \times d_{\mathrm{ave}} \times h_{\mathrm{ave}}$$

which, with extension to include metrological terms might be written as

$$A = \pi\left(d_{\mathrm{ave}} + \delta d_{\mathrm{rep}} + \delta d_{\mathrm{cal}} + \cdots\right)\left(h_{\mathrm{ave}} + \delta h_{\mathrm{rep}} + \delta h_{\mathrm{cal}} + \cdots\right)$$

where terms for inputs such as drift and resolution have been omitted for the sake of clarity.

If the budget is presented in a table, it might appear as shown in the example below.

Quantity X_i	Estimate x_i	Source of uncertainty	Uncertainty	PDF	Divisor	$u(x_i)$	c_i	$u_i(A)$ /m^2
δd_rep	$0\ \mathrm{m}$	Repeatability of diameter	$0.0081\ \mathrm{m}$					
δd_cal	$0\ \mathrm{m}$	Calibration of diameter gauge	$0.00061\ \mathrm{m}$					
⋮	⋮	⋮	⋮					
δh_rep	$0\ \mathrm{m}$	Repeatability of height	$0.0012\ \mathrm{m}$					
δh_cal	$0\ \mathrm{m}$	Calibration of height gauge	$0.00044\ \mathrm{m}$					
⋮	⋮	⋮	⋮					

Note that d_ave and h_ave do not need to appear in the table if all uncertainty associated with these quantities is accounted for by other terms.

D.2.5 Alternatively, repeatability is often evaluated and modelled in terms of the output quantity. This is more likely to be the case in situations where reliable individual repeatability estimates are difficult to obtain (as is very often the case), and repeatability is instead estimated from multiple realisations of the measurand (either during the measurement or during a previous repeatability assessment).

For example, the outer surface area of the smooth regular cylinder can be estimated from the average of several repeat values

$$A = \frac{1}{n}\sum_{i=1}^{n} A_i = \frac{1}{n}\sum_{i=1}^{n} \pi \times d_i \times h_i$$

which, with extension to include metrological terms might be written as

$$A = \frac{1}{n}\sum_{i=1}^{n} \pi(d_i + \delta d_\mathrm{cal} + \cdots)(h_i + \delta h_\mathrm{cal} + \cdots) + \delta A$$

again, terms such as drift and resolution have been omitted for the sake of clarity.

If the budget is presented in a table, it might appear as shown in the example below

Quantity X_i	Estimate x_i	Source of uncertainty	Uncertainty	PDF	Divisor	$u(x_i)$	c_i	$u_i(V)$ /m^2
δd_cal	$0\ \mathrm{m}$	Calibration of diameter gauge	$0.00061\ \mathrm{m}$					
δh_cal	$0\ \mathrm{m}$	Calibration of height gauge	$0.00044\ \mathrm{m}$					
δA	$0\ \mathrm{m}^2$	Repeatability	$0.0012\ \mathrm{m}^2$					
⋮	⋮	⋮	⋮					

D.3 Correlated input quantities

D.3.1 The expressions for the standard uncertainty of the output given in equations (1) and (8) only apply when the input quantities are independent of each other, in other words, when there is no **correlation** between any of the input estimates.

It may however be the case that some input quantities are affected by a common factor which introduces correlation into the equation. For example, common temperature effects, or common measurement errors from an instrument that is used to measure several of the inputs to a process introduces correlation.

D.3.2 Sometimes it is possible to construct the equation in such a way as to avoid correlated inputs.

For example, weights taken from the same set are usually assumed to have correlated values. Suppose that 3 weights m_1, m_2, and m_3 from such a set are combined to form a load that is applied to a piston. The applied pressure can be written as:

$$P = \frac{(m_1 + m_2 + m_3)g}{A}$$

where A is the cross-sectional area over which the weight acts and g is the acceleration due to gravity.

This equation includes three correlated terms, m_1, m_2, and m_3, so it would be **incorrect** to combine the uncertainties of all items using equation (1)

$$u_c(p) = \sqrt{u_{m_1}^2(p) + u_{m_1}^2(p) + u_{m_2}^2(p) + u_{m_3}^2(p) + u_g^2(p) + u_A^2(p)} \qquad \textbf{(incorrect)}$$

A common approach to this problem would be to evaluate separately the uncertainty for the total load, W and incorporate this into an evaluation involving only independent terms

$$u_c(p) = \sqrt{u_W^2(p) + u_g^2(p) + u_A^2(p)}$$

where $W = m_1 + m_2 + m_3$

When there is full correlation between several input quantities, x_j, their combined standard uncertainty is found by summation rather than combining them using equation (1)

$$\textstyle\sum_j c_j u(x_j) \qquad (9)$$

It is common practice to assume that there is full correlation between the values of weights in a set; therefore, the uncertainty $u(W)$ is found from

$$u(W) = u(m_1) + u(m_2) + u(m_3)$$

D.3.3 In cases where measurement errors combine so as to increase uncertainty, as in the above example, this is referred to as positive correlation. In other cases, the effects of correlated input quantities may serve to reduce the combined uncertainty, such as when an instrument is used as a comparator between a standard and an unknown - this is referred to as negative correlation.

D.3.4 A further example of the treatment of correlated contributions can be found in paragraph K.6.4. GUM (Annex F.1.2) should be consulted for a more detailed description of approaches for dealing with correlation.

Appendix E Some sources of error and uncertainty in electrical calibrations

The following is a description of the more common sources of systematic error and uncertainty (after correction) in electrical calibration work, with brief comments about their nature. Further, more detailed, advice is given in specialised technical publications and manufacturers' application notes, as well as other sources.

E.1 Imported uncertainty

E.1.1 The uncertainties assigned to the values on a calibration certificate for the calibration of an instrument, whether measuring equipment or a reference standard, are all contributors to the uncertainty budget.

E.2 Secular stability

E.2.1 The performance of all instruments, and the values of reference standards, must be expected to change to some extent with the passage of time. Passive devices such as standard resistors or high-grade RF and microwave attenuators may be expected to drift slowly with time. An estimate of such a drift must be assessed on the basis of values obtained from previous calibrations. It cannot be assumed that a drift will be linear. Data can be assimilated readily if displayed in a graphical form. A curve fitting procedure that gives a progressively greater weight to each of the more recent calibrations can be used to allow the most probable value at the time of use to be assessed. The degree of complexity in curve fitting is a matter of judgement; in some cases, drawing a smooth curve through the chosen data points by hand can be quite satisfactory. Whenever a new calibration is obtained the drift characteristic will need re-assessment. The corrections that are applied for drift are subject to uncertainty based on the scatter of data points about the drift characteristic. The magnitude of the drift and the random instability of an instrument, and the accuracy required will determine the calibration interval.

E.2.2 With complex electronic equipment it is not always possible to follow this procedure as changes in performance can be expected to be more random in nature over relatively long periods. Checks against passive standards can establish whether conformity to specification is being maintained or whether a calibration with subsequent equipment adjustment is needed. The manufacturer's specification can be a good starting point for assigning the uncertainty due to instrument drift but should be confirmed by analysis of quality control and calibration data.

E.3 Environmental conditions

E.3.1 The laboratory measurement environment can be one of the most important considerations when performing electrical calibrations. Ambient temperature is often the most important influence and information on the temperature coefficient of, for example, resistance standards must be sought or determined. Variations in relative humidity can also affect the values of unsealed components. The influence of barometric pressure on certain electrical measurement standards can also be significant. At RF and microwave frequencies, ambient temperature can affect the performance of, for example, attenuators, impedance standards that depend on mechanical dimensions for their values and other precision components. Devices that incorporate thermal sensing, such as power sensors, can be affected by rapid temperature changes that can be introduced by handling or exposure to sunlight or other sources of heat.

E.3.2 It is also necessary to be aware of the possible effects of electrical operating conditions, such as power dissipation, harmonic distortion, or level of applied voltage being different when a device is in use from when it was calibrated. Resistance standards, resistive voltage dividers and attenuators at any frequency are examples of devices being affected by self-heating and/or applied voltage. It should also be ensured that all equipment is operating within the manufacturer's stated range of supply voltages.

E.3.3 The effects of harmonics and noise on ac calibration signals may have an influence on the apparent value of these signals. Similarly, the effects of any common-mode signals present in a measurement system may have to be accounted for.

E.4 Interpolation of calibration data

E.4.1 When an instrument with a broad range of measurement capabilities is calibrated, there are practical and economic factors that limit the number of calibration points. Consequently, the value of the quantity to be measured and/or its frequency may be different from any of the calibration points. When the value of the quantity lies between two calibration values, consideration needs to be given to systematic errors that arise from, for example, scale non-linearity.

E.4.2 If the measurement frequency falls between two calibration frequencies, it will also be necessary to assess the additional uncertainty due to interpolation that this can introduce. One can only proceed with confidence if:

(a) a theory of instrument operation is known from which one can predict a frequency characteristic, or there is additional frequency calibration data from other models of the same instrument,

and wherever reasonable,

(b) the performance of the actual instrument being used has been explored with a swept frequency measurement system to verify the absence of resonance effects or aberrations due to manufacturing or other performance limitations.

E.5 Resolution

E.5.1 The limit to the ability of an instrument to indicate small changes in the quantity being measured, referred to as resolution or “digital rounding error”, is treated as a systematic component of uncertainty.

E.5.2 Many instruments with a digital display use an analogue-to digital converter (ADC) to convert the analogue signal under investigation into a form that can be displayed in terms of numeric digits. The last recorded digit will be a rounded representation of the underlying analogue signal. The error introduced by this process will be from -0.5 digit (else the last digit would be one lower) to +0.5 digit (else the last digit would be one higher). A maximum quantisation error of ±0.5 digit is therefore present. As there is no way of knowing where within this range the underlying value is, the resulting error is assumed to be zero with limits of ±0.5 digit.

E.5.3 A maximum error of ±0.5 digit may not apply in all instances and an understanding of instrument operation is needed if the assigned uncertainty is to be realistic. For example, a direct-gating frequency counter has a maximum digital rounding error of ±1 digit, due to the random relationship between the signal being measured and the internal clock. Some instruments may also display hysteresis that, although not necessarily a property of the display itself, may result in further uncertainties amounting to several digits.

E.5.4 In an analogue instrument the effect of resolution is determined by the practical ability to read the position of a pointer on a scale. In either case, the last digit actually recorded will always be subject to an uncertainty of at least ±0.5 digit. The presence of electrical noise causing fluctuations in instrument readings will commonly determine the usable resolution, however it is possible to make a good estimate of the mean position of a fluctuating pointer by eye.

E.6 Apparatus layout

E.6.1 The physical layout of one item of equipment with respect to another and the relationship of these items to the earth plane can be important in some measurements. Thus a different arrangement between calibration and subsequent use of an instrument may be the source of systematic errors. The main effects are leakage currents to earth, interference loop currents, and electromagnetic leakage fields. In inductance measurements it is necessary to define connecting lead configuration and be conscious of the possible effects of an earth plane or adjacent ferromagnetic material. The effect of mutual heating between apparatus may also need to be considered.

E.7 Thermoelectric voltages

E.7.1 If an electrical conductor passes through a temperature gradient, then a potential difference will be generated across that gradient. This is known as the *Seebeck effect* and these unwanted, parasitic voltages can cause errors in some measurement systems – in particular, where small dc voltages are being measured.

E.7.2 They can be minimised by design of connections that are thermally symmetrical, so that the Seebeck voltage in one lead is cancelled by an identical and opposite voltage in the other. In some situations, e.g., ac/dc transfer measurements, the polarity of the dc supply is reversed, and an arithmetic mean is taken of two sets of dc measurements.

E.7.3 Generally, an allowance must be made as a Type B component of uncertainty for the presence of thermal EMFs.

E.8 Loading and cable impedance

E.8.1 The finite input impedance of voltmeters, oscilloscopes and other voltage sensing instruments may so load the circuit to which they are connected as to cause significant systematic errors. Corrections may be possible if impedances are known. In particular, it should be noted that some multi-function calibrators can exhibit a slightly inductive output impedance. This means that when a capacitive load is applied, the resulting resonance may cause the output voltage to increase with respect to its open-circuit value.

E.8.2 The impedance and finite electrical length of connecting leads or cables may also result in systematic errors in voltage measurements at any frequency. The use of four-terminal connections minimises such errors in some dc and ac measurements.

E.8.3 For capacitance measurements, the inductive properties of the connecting leads may be important, particularly at higher values of capacitance and/or frequency. Similarly, for inductance measurements the capacitance between connecting leads may be important.

E9 RF mismatch errors and uncertainty

E.9.1 At RF and microwave frequencies the mismatch of components to the characteristic impedance of the measurement system transmission line can be one of the most important sources of error and of the systematic component of uncertainty in power and attenuation measurements. This is because the phases of voltage reflection coefficients are not usually known and hence corrections cannot be applied.

E.9.2 In a power measurement system, the power, P_0, that would be absorbed in a load equal to the characteristic impedance of the transmission line has been shown (by Harris and Warner [17]) to be related to the actual power, P_L, absorbed in a wattmeter terminating the line by the equation

$$P_0 = \frac{P_L}{1 - |\Gamma_L|^2}(1 - 2|\Gamma_G||\Gamma_L| \cos\phi + |\Gamma_G|^2|\Gamma_L|^2) \quad \text{E1}$$

where ϕ is the relative phase of the generator and load voltage reflection coefficients Γ_G and Γ_L. When Γ_G and Γ_L are small, this becomes

$$P_0 = \frac{P_L}{1 - |\Gamma_L|^2}(1 - 2|\Gamma_G||\Gamma_L| \cos\phi) \quad \text{E2}$$

E.9.3 This expression for absorbed power can have limits:

$$P_0(\text{limits}) = \frac{P_L}{1 - |\Gamma_L|^2}(1 \pm 2|\Gamma_G||\Gamma_L|) \quad \text{E3}$$

E.9.4 The calculable mismatch error is $(1 - |\Gamma_L|^2)$ and is accounted for in the calibration factor, while the limits of mismatch uncertainty are $(\pm 2|\Gamma_G||\Gamma_L|)$. Because a cosine function characterises the probability distribution for the uncertainty, Harris and Warner show that the distribution is U-shaped with a standard deviation given by

$$u(\text{mismatch}) = \frac{2|\Gamma_G||\Gamma_L|}{\sqrt{2}} = 1.414\,|\Gamma_G||\Gamma_L| \quad \text{E4}$$

E.9.5 When a measurement is made of the attenuation of a two-port component inserted between a generator and load that are not perfectly matched to the transmission line, Harris and Warner have shown that the standard deviation of mismatch, M, expressed in dB is approximated by

$$M = \frac{8.686}{\sqrt{2}}[M_G + M_L + M_{GL}]^{0.5} \quad \text{E5}$$

where

$$M_G = |\Gamma_G|^2\,(|s_{11a}|^2 + |s_{11b}|^2)$$

$$M_L = |\Gamma_L|^2\,(|s_{22a}|^2 + |s_{22b}|^2)$$

$$M_{GL} = |\Gamma_G|^2|\Gamma_L|^2\,(|s_{21a}|^4 + |s_{21b}|^4)$$

and Γ_G and Γ_L are the source and load voltage reflection coefficients respectively and s_{11}, s_{22}, s_{21} are the scattering coefficients of the two-port component with the suffix a referring to the starting value of the attenuator and b referring to the finishing value of the attenuator. Harris and Warner concluded that the distribution for M would approximate to that of a normal distribution due to the combination of its component distributions.

E.9.6 The values of Γ_G and Γ_L used in equations E(4) and E(5) and the scattering coefficients used in equation E(5) will themselves be subject to uncertainty because they are derived from measurements. This uncertainty has to be considered when calculating the mismatch uncertainty and it is recommended that this is done by adding it in quadrature with the measured or derived value of the reflection coefficient; for example, if the measured value of Γ_L is 0.03 ± 0.02 then the value of Γ_L that should be used to calculate the mismatch uncertainty is $\sqrt{0.03^2 + 0.02^2}$ i.e. 0.036.

E.10 Directivity

E.10.1 When making voltage reflection coefficient (VRC) measurements at rf and microwave frequencies, the finite directivity of the bridge or reflectometer gives rise to an uncertainty in the measured value of the VRC, if only the magnitude and not the phase of the directivity component is known. The uncertainty will be equal to the directivity, expressed in linear terms, e.g., a directivity of 30 dB is equivalent to an uncertainty of 0.0316 VRC.

E.10.2 As with E9.6 above it is recommended that the uncertainty in the measurement of directivity is taken into account by adding the measured value in quadrature with the uncertainty, in linear quantities; for example, if the measured directivity of a bridge is 36 dB (0.016) and has an uncertainty of +8 dB -4 dB (±0.01) then the directivity to be used is $\sqrt{0.016^2 + 0.01^2} = 0.019$, (i.e. 34.4 dB).

E.11 Test port match

E.11.1 The test port match of a bridge or reflectometer used for reflection coefficient measurements will give rise to an error in the measured VRC due to re-reflection. The uncertainty, $u(TP)$, is calculated from
$u(TP) = TP.|\Gamma_X|^2$, where TP is the test port match, expressed as a VRC, and Γ_X is the measured reflection coefficient. When a directional coupler is used to monitor incident power in the calibration of a power meter it is the effective source match of the coupler that defines the value of Γ_G referred to in E9. As with E9.6 and E10, the measured value of test port match will have an uncertainty that should be taken into account by using quadrature summation.

E.12 RF connector repeatability

The lack of repeatability of coaxial pair insertion loss and, to a lesser extent, voltage reflection coefficient is a problem when calibrating devices in a coaxial line measurement system and subsequently using them in some other system. Repeatedly connecting and disconnecting the device can evaluate the repeatability of given connector pairs in use.

Appendix F Some sources of error and uncertainty in mass calibrations

This Appendix describes the more common sources of errors and uncertainties in mass calibration with brief comments about their nature. They may not all be significant at all levels of measurement, but their effect should at least be considered when estimating the overall uncertainty of a measurement. Further information about mass calibration and the calibration of weighing machines can be found in UKAS LAB 14 [13] and EURAMET CG-18 [12].

F.1 Reference weight calibration

F.1.1 The uncertainties assigned to the values on a calibration certificate for the calibration of the reference weights are all contributors to the uncertainty budget.

F.2 Secular stability of reference weights

F.2.1 It is necessary to take into account the likely change in mass of the reference weights since their last calibration. This change can be estimated from the results of successive calibrations of the reference weights. If such a history is not available, then it is usual to assume that they may change in mass by an amount equal to their uncertainty of calibration between calibrations. The stability of weights can be affected by the material and quality of manufacture (e.g., ill-fitting screw knobs), surface finish, unstable adjustment material, physical wear and damage and atmospheric contamination. The figure adopted for stability will need to be reconsidered if the usage or environment of the weights changes. The calibration interval for reference weights will depend on the stability of the weights.

F.3 Weighing machine/weighing process

F.3.1 The performance of the weighing machine used for the calibration should be assessed to estimate the contribution it makes to the overall uncertainty of the weighing process. The performance assessment should cover those attributes of the weighing machine that are significant to the weighing process. For example, the length of arm error (assuming it is constant) of an equal arm balance need not be assessed if the weighing process only uses substitution techniques (Borda's method). The assessment may include some or all of the following:

(a) repeatability of measurement;

(b) linearity within the range used;

(c) digit size/weight value per division, i.e., resolution;

(d) eccentricity (off centre load), especially if groups of weights are placed on the weighing pan simultaneously;

(e) magnetic effects (e.g., magnetic weights, or the effect of force balance motors on cast iron weights);

(f) temperature effects, e.g., differences between the temperature of the weights and the weighing machine;

(f) length of arm error.

F.4 Air buoyancy effects

F.4.1 The accuracy with which air buoyancy corrections can be made depends on how well the density of the weights is known, and how well the air density can be determined. Some laboratories can determine the density of weights, but for most mass work assumed figures are used. The air density is usually calculated from an equation (e.g., see [13]) after measuring the air temperature, pressure and humidity. For the highest levels of accuracy, it may also be necessary to measure the carbon dioxide content of the air. The figures that follow are based upon an air density range of 1.079 kg m^{-3} to 1.291 kg m^{-3} which can be produced by ranges of relative humidity from 30 % to 70 %, air temperature from 10 °C to 30 °C and barometric pressure from 950 mbar to 1050 mbar.

F.4.2 For mass comparisons a figure of ±1 part in 10^6 of the applied mass is typical for common weight materials such as stainless steel, plated brass, German silver and gunmetal. For cast iron the figure may be up to ±3 parts in 10^6 and for aluminium up to ±30 parts in 10^6. The uncertainty can be reduced if the mass comparisons are made within suitably restricted ranges of air temperature, pressure and humidity. If corrections are made for the buoyancy effects the uncertainty can be virtually eliminated, leaving just the uncertainty of the correction.

F.4.3 Certain weighing machines display mass units directly from the force they experience when weights are applied. It is common practice to reduce the effects of buoyancy on such devices by the use of an auxiliary weight, known as a spanning weight, which is used to normalise the readings to the prevailing conditions, as well as compensating for changes in the machine itself. This spanning weight can be external or internal to the machine. If such machines are not spanned at the time of use the calibration may be subject to an increased uncertainty due to the buoyancy effects on the loading weights. For weighing machines that make use of stainless steel, plated brass, German silver or gunmetal weights this effect may be up to ±16 parts in 10^6. For cast iron weights the figure may be up to ±18 parts in 10^6 and for aluminium weights up to ±45 parts in 10^6.

F.4.4 For the ambient conditions stated above the uncertainty limits due to buoyancy effects may be ±110 parts in 10^6 and ±140 parts in 10^6 respectively for comparing water and organic solvents with stainless steel mass standards, and ±125 parts in 10^6 and ±155 parts in 10^6 respectively for direct weighing.

F.4.5 Apart from air buoyancy effects, the environment in which the calibration takes place can introduce other uncertainties. Temperature gradients can give rise to convection currents in the balance case, which will affect the reading, as will draughts from air conditioning units. Rapid changes of temperature in the laboratory can affect the weighing process. Changes in the level of humidity in the laboratory can make short-term changes to the mass of weights, while low levels of humidity can introduce static electricity effects on some comparators. Dust contamination also introduces errors in calibrations. The movement of weights during the calibration causes disturbances to the local environment.

Appendix G Some sources of error and uncertainty in temperature calibrations

The more common sources of systematic error and uncertainty in the measurement of temperature are described in this section. Each source may have several uncertainty components.

G.1 Reference thermometer calibration

G.1.1 The uncertainties assigned to the values on a calibration certificate for the calibration of the reference thermometer are all contributors to the uncertainty budget.

G.2 Measuring instruments

G.2.1 The uncertainty assigned to the calibration of any electrical or other instruments used in the measurements, e.g., standard resistors, measuring bridges and digital multimeters.

G.3 Further influences

G.3.1 Additional uncertainties in the measurement of the temperature using the reference thermometers:

(a) Drift since the last calibration of the reference thermometers and any associated measuring instruments;

(b) Resolution of reading; this may be very significant in the case of a liquid-in-glass thermometer or digital thermometers;

(c) Instability and temperature gradients in the thermal environment, e.g., the calibration bath or furnace, including any contribution due to difference in immersion of the reference standard from that stated on its certificate of calibration;

(d) When platinum resistance thermometers are used as reference standards any contribution to the uncertainty due to self-heating effects should be considered. This will mainly apply if the measuring current is different from that used in the original calibration and/or the conditions of measurement e.g., `in air' or in stirred liquid.

G.4 Contributions associated with the thermometer to be calibrated

G.4.1 These may include factors associated with electrical indicators as well as some of the further influences already mentioned. When partial immersion liquid-in-glass thermometers are to be calibrated an additional uncertainty contribution to account for effects arising from differences in depth of immersion should be included even when the emergent column temperature is measured.

G.4.2 When thermocouples are being calibrated any uncertainty introduced by compensating leads and reference junctions should be taken into account. Similarly, any thermal EMFs introduced by switches or scanner units should be investigated. Unknown errors arising from inhomogeneity of the thermocouple being calibrated can give rise to significant uncertainties. Ideally this should be evaluated at the time of calibration, possibly by varying the immersion depth of the thermocouple in an isothermal enclosure. For many calibrations, however, this will not be practical. In such cases, a figure of 20 % of the maximum permissible error for the particular thermocouple type is considered reasonable.

G.5 Mathematical interpretation

G.5.1 Uncertainty arising from mathematical interpretation, e.g., in applying scale corrections or deviations from a reference table, or in curve-fitting to allow for scale non-linearity, should be assessed.

Appendix H Some sources of error and uncertainty in dimensional calibrations

The more common sources of systematic error and uncertainty in dimensional measurements are described in this section.

H.1 Reference standards and Instrumentation

H.1.1 The uncertainties assigned to the reference standards and those for the measuring instruments used to make the measurements.

H.2 Secular stability of reference standards and instrumentation

H.2.1 The changes that occur over time must be taken into account, usually by reference to the calibration history of the equipment. This is particularly important when the equipment may be exposed to physical wear as part of normal operation.

H.3 Temperature effects

H.3.1 The uncertainties associated with differences in temperature between the gauge being calibrated and the reference standards and measuring instruments used should be accounted for. These will be most significant over the longer lengths and in cases involving dissimilar materials. Whilst it may be possible to make corrections for temperature effects there will be residual uncertainties resulting from uncertainty in the values used for the coefficients of expansion and the calibration of the measuring thermometer.

H.4 Elastic compression

H.4.1 These are uncertainties associated with differences in elastic compression between the materials from which the gauge being calibrated and the reference standards were manufactured. They are likely to be most significant in the more precise calibrations and in cases involving dissimilar materials. They will relate to the measuring force used and the nature of stylus contact with the gauge and reference standard. Whilst mathematical corrections can be made there will be residual uncertainties resulting from the uncertainty of the measuring force and in the properties of the materials involved.

H.5 Cosine errors

H.5.1 Any misalignment of the gauge being calibrated, or reference standards used, with respect to the axis of measurement, will introduce errors into the measurements. Such errors are often referred to as cosine errors and can be minimised by adjusting the attitude of the gauge with respect to the axis of measurement to find the relevant turning points that give the appropriate maximum or minimum result. Small residual errors can still result where, for instance, incorrect assumptions are made concerning any features used for alignment of the datums.

H.6 Geometric errors

H.6.1 Errors in the geometry of the gauge being calibrated, any reference standards used, or critical features of the measuring instruments used to make the measurements can introduce additional uncertainties. Typically, these will include small errors in the flatness or sphericity of stylus tips, the straightness, flatness, parallelism, or squareness of surfaces used as datum features, and the roundness or taper in cylindrical gauges and reference standards. Such errors are often most significant in cases where perfect geometry has been wrongly assumed and where the measurement methods chosen do not capture, suppress or otherwise accommodate the geometric errors that prevail in a particular case.

Appendix J Some sources of error and uncertainty in pressure calibrations using dead weight testers

The more common sources of systematic error and uncertainty in the generation of known pressures, using dead weight testers (DWT), are described in this section.

J.1 Reference DWT

J.1.1 The uncertainties assigned to the values on a calibration certificate for the reference dead weight tester are all contributors to the uncertainty budget. These include the following:

(a) Area uncertainty including any uncertainty in the distortion. This uncertainty will often vary with pressure;

(b) Piston and weight carrier mass.

J.2 Secular stability of the reference dead weight tester

J.2.1 It is necessary to account for likely changes in the area and mass of the reference DWT since the last calibration. This change can be estimated from successive calibrations of the reference DWT. The secular stability uncertainty for the area will depend on the calibration interval and can be larger than the calibration uncertainty. It may also vary with pressure and should be evaluated over the range of use of the DWT. The variation between calibrations in the area of a DWT will depend on its usage, design, and material composition and is therefore a best estimate from actual data. Where this is not available it is recommended that a pessimistic estimate is made, and a short calibration interval set.

J.2.2 The drift of the piston mass will be larger in oil DWTs as this will reflect the difficulties in repeat weighting of pistons that have been immersed in oil. These difficulties arise from incomplete cleaning processes and possible instability due to the evaporation of solvents.

J.3 Reference DWT mass set uncertainty

J.3.1 The uncertainties assigned to the values on a calibration certificate for the weights in the reference dead weight tester mass set are all contributors to the uncertainty budget. The uncertainty of the mass stack used to generate pressure should be evaluated over the range of the DWT. The relative uncertainty is often higher at lower pressures.

J.4 Secular stability of the reference DWT mass set

J.4.1 It is necessary to account for likely changes the mass set of the reference DWT since the last calibration. Paragraph F.2.1 addresses the subject of secular stability of reference weights.

J.5 Uncertainty of local gravity determination

J.5.1 The pressure generated by a DWT is directly affected by the local acceleration due to gravity, g. With care, this can be measured with an uncertainty of less than 1 ppm. It is possible for an estimate of the g value to be obtained from a reputable geological survey organisation based on a grid reference; this would attract an uncertainty of around 3 ppm. It can also be calculated from knowledge of latitude and altitude; however, the uncertainty will be much larger - around 50 ppm in the UK. Some knowledge of the Bouguer anomalies is required to achieve these levels of uncertainty from such calculations.

J.6 Air buoyancy effect

J.6.1 Air buoyancy affects the mass set of a DWT in the same way as described in paragraph F.4

J.7 Temperature effect on DWT area

J.7.1 The area on a DWT changes with temperature; its temperature coefficient of expansion being related to the particular materials that the piston and cylinder are made from. Consideration must be given to any variation in temperature from the reference temperature when the DWT was calibrated, variation in temperature during a calibration and uncertainty in the determination of the piston temperature.

J.8 Uncertainty due to head correction

J.8.1 Any difference between the height of the reference DWT datum level and that of the item being calibrated will affect the pressure generated at that item. For pneumatic calibrations this effect is proportional to pressure and normally equates to about 116 ppm/m. For hydraulic calibrations the effect is a fixed pressure effect that will depend on the density of the fluid used, local acceleration due to gravity and the height difference (fluid head pressure = $\rho g h$). For most DWT oils the effect is between 8 Pa/mm and 9 Pa/mm.

J.8.2 The float height position of the piston will also contribute to the head correction uncertainty. This effect will be related to the fall rate of the piston and the particular measurement procedure in use.

J.9 Effects of fluid properties

J.9.1 For hydraulic calibrations the effect of the fluid properties on fluid head corrections, buoyancy volume corrections and surface tension corrections will also need to be considered. These figures are usually reported on calibration certificates for DWTs. However, care must be taken to convert any quoted correction to the actual oil used if different from that used during the calibration of the reference DWT. In most circumstances the uncertainty of these influence quantities can be treated as negligible.

J.10 Non-verticality of the DWT piston

An uncertainty arises because the piston may not be perfectly vertical. If it were, then all the force would act on the area. Any departure from vertical will reduce the force and therefore the generated pressure. The effect in terms of generated pressure is proportional to the cosine of the angle from true vertical.

J.11 Uncertainties arising from the calibration process

J.11.1 Any uncertainty arising from the calibration process will need to be evaluated. These could include the resolution and repeatability of the unit being calibrated and the effects of the environment on it. Uncertainties due to calibration or data fitting of the calibration results may also have to be considered.

Appendix K Examples of application

Introduction:

(a) This Appendix presents a number of example uncertainty budgets in various fields of measurement. The examples are not intended as preferred or mandatory requirements. They are presented to illustrate the principles involved in uncertainty evaluation and to show how the common sources of uncertainty in the various fields can be analysed in practice. They are, however, believed to be realistic for the measurements described.

(b) An uncertainty budget is a statement of measurement uncertainty evaluation, of the components of that uncertainty, and of their calculation and combination. It should include the measurement model, estimates and uncertainties associated with the quantities in the measurement model, covariances, type of assumed probability density functions, degrees of freedom, type of evaluation, coverage interval, coverage probability and coverage factors. An uncertainty budget is not simply a summary table; it must include all these factors. The example uncertainty budgets presented in this Appendix comply with this definition.

(c) These examples may also be used for the purpose of software validation. If an uncertainty budget has been prepared using a spreadsheet, the configuration of the spreadsheet can be verified by entering the same values and comparing the output of the spreadsheet with the results shown in the examples.

K.1 Measurement of a 10 kΩ resistor by voltage intercomparison

K.1.1 A high-resolution digital voltmeter is used to measure the voltages developed across a standard resistor and an unknown resistor of the same nominal value as the standard when the series-connected resistors are supplied from a constant current dc source. Both resistors are immersed in a temperature-controlled oil bath maintained at 20.0 °C. The value of the unknown resistor, R_X, is given by

$$R_X = (R_S + \delta R_D + R_{TC}\Delta t)\frac{V_X}{V_S}$$

where

R_S = calibration value for the standard resistor,
δR_D = drift in R_S since the previous calibration,
R_{TC} = temperature coefficient of resistance for R_S,
Δt = maximum variation in oil bath from nominal temperature,
V_X = voltage across R_X,
V_S = voltage across R_S.

K.1.2 The calibration certificate for the standard resistor reported an uncertainty of ±0.5 ppm at a coverage probability of approximately 95% ($k = 2$).

K.1.3 No correction was made for drift in the value of R_S i.e., the drift is assumed to be $\delta R_D = 0$. Records indicate that the relative drift in R_S is unlikely to exceed ±0.5 ppm.

K.1.4 The temperature coefficient of resistance for the standard resistor was obtained from a graph of temperature versus resistance. Such curves are normally parabolic in nature, however using a linear approximation over the small range of temperature variation encountered in the bath, a value of ±2.5 ppm per ºC was assigned. This value was included in the uncertainty budget as a sensitivity coefficient.

K.1.5 Records of evaluation of the oil bath characteristics showed that the maximum temperature deviation from the set point did not exceed ±0.1 ºC at any point within the bath.

K.1.6 The same voltmeter is used to measure V_X and V_S and although the uncertainty contributions will be correlated the effect is to reduce the uncertainty and it is only necessary to consider the relative difference in the voltmeter readings due to linearity and resolution, which was estimated to have limits of ±0.2 ppm for each reading. Each of these is assigned a rectangular distribution.

K.1.7 Type A evaluation: The repeatability for R_X can be estimated from the repeatability of the measured voltage ratio, V_X/V_S. Five measurements were made to record the departure from unity in the ratio The measured departures were:

+10.4 ppm, +10.7 ppm, +10.6 ppm, +10.3 ppm, +10.5 ppm

From equation (3), the mean departure from unity = +10.50 ppm

The repeatability standard deviation $s(R_X)$ is estimated from the five measured departure values. Applying equation (5), gives $s(R_X)$ = 0.158 ppm.

The reported measurement result is to be calculated using the mean of the $n = 5$ measurements. So, from equation (4), the repeatability uncertainty is

$$u_{\text{rep}}(R_X) = \frac{s(R_X)}{\sqrt{n}} = \frac{0.158}{\sqrt{5}} = 0.0707 \text{ ppm}$$

K.1.8 Summary table for $R_X = (R_S + \delta R_D + R_{TC}\Delta t)\frac{V_X}{V_S}$

Quantity	Source of uncertainty	Uncertainty	Probability distribution	Divisor	c_i	$u_i(R_X)$ ppm	ν_i or ν_{eff}
R_S	Calibration of standard resistor	0.5 ppm	Normal	2	1	0.25	∞
δR_D	Uncorrected drift since last calibration	0.5 ppm	Rectangular	√3	1	0.289	∞
Δt	Temperature effects	0.1 ºC	Rectangular	√3	2.5 ppm/ºC	0.144	∞
V_S	Voltmeter across R_S	0.2 ppm	Rectangular	√3	1	0.115	∞
V_X	Voltmeter across R_X	0.2 ppm	Rectangular	√3	1	0.115	∞
R_X	Repeatability of indication	0.071 ppm	Normal	1	1	0.071	4
$u_c(R_X)$	Combined standard uncertainty		Normal			0.445	>500
$U_{95\%}(R_X)$	Expanded uncertainty		Normal ($k = 2$)			0.891	>500

K.1.9 Reported result

The measured value of the 10 kΩ resistor at 20 ºC ± 0.1 ºC was **10 000.1050 Ω ± 0.0089 Ω**

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

NOTE: The temperature coefficient of the resistor being calibrated is not included here, as it is an “unknown” quantity. The relevant temperature conditions are included in the reporting of the result. Best practice would be to estimate a value together with a suitable uncertainty, and to include these details with the reported results.

K.2 Calibration of a coaxial power sensor at a frequency of 18 GHz

K.2.1 The measurement involves the calibration of an unknown power sensor against a standard power sensor by substitution on a stable, monitored source of defined source impedance. The measurement is made in terms of Calibration Factor, defined as,

$$K_X = \frac{\text{Incident power at reference frequency}}{\text{Incident power at calibration frequency}}$$

for the same power sensor response. It is determined from the following:

Calibration Factor,

$$K_X = (K_S + \delta D_S) \times f_{\mathrm{DC}} \times f_{\mathrm{M}} \times f_{\mathrm{REF}}$$

where

K_S = Calibration Factor of the standard sensor,
δD_S = error due to drift in standard sensor since the previous calibration,
f_{DC} = DC voltage output nonlinearity factor,
f_{M} = ratio of Mismatch Losses,
f_{REF} = ratio of reference power source (short-term stability of 50 MHz reference).

K.2.2 Four separate measurements were made which involved disconnection and reconnection of both the unknown sensor and the standard sensor on a power transfer system. All measurements were made in terms of voltage ratios that are proportional to calibration factor.

K.2.3 None of the uncertainty contributions are considered to be correlated to any significant extent.

K.2.4 There will be mismatch uncertainties associated with the source/standard sensor combination and with the source/unknown sensor combination. These will be $200.\Gamma_G.\Gamma_S$ % and $200.\Gamma_G.\Gamma_X$ % respectively, where

Γ_G = 0.02 at 50 MHz and 0.07 at 18 GHz,
Γ_S = 0.02 at 50 MHz and 0.10 at 18 GHz,
Γ_X = 0.02 at 50 MHz and 0.12 at 18 GHz.

These values include the uncertainty in the measurement of Γ as described in paragraph E.9.6.

K.2.5 The standard power sensor was calibrated by an accredited laboratory 6 months before use; the expanded uncertainty of 1.1 % was quoted for a coverage factor $k = 2$.

K.2.6 The long-term stability of the standard sensor was estimated from the results of 5 annual calibrations. No predictable trend could be detected so drift corrections could not be made. The error due to secular stability was therefore assumed to be zero with limits, in this case, not greater than ±0.4 % per year. A value of ±0.2 % was used as the previous calibration was within 6 months.

K.2.7 The instrumentation linearity uncertainty was estimated from measurements against a reference attenuation standard. The expanded uncertainty for $k = 2$ of 0.1 % applies to ratios up to 2:1.

K.2.8 Type A evaluation: The four measurements resulted in the following values of Calibration Factor:

93.45 %, 92.20 %, 93.95 %, 93.02 %

From equation (3), the mean value $\overline{K_X} = 93.16$ %

The repeatability standard deviation $s(K_X)$ is estimated from the four measured values. Applying equation (5), gives $s(K_X) = 0.7415$ %.

The reported measurement result is to be calculated using the mean of $n = 4$ measurements. So, from equation (4), the repeatability uncertainty is

$$u_{\text{rep}}(K_X) = \frac{s(K_X)}{\sqrt{n}} = \frac{0.7415}{\sqrt{4}} = 0.3707\ \%$$

K.2.9 Summary table for $K_X = (K_S + \delta D_S) \times f_{\text{DC}} \times f_{\text{M}} \times f_{\text{REF}}$

Symbol	Source of uncertainty	Uncertainty %	Probability distribution	Divisor	c_i	$u_i(K_X)$ %	ν_i or ν_{eff}
K_S	Calibration factor of standard	1.1	Normal	2.0	1.0	0.55	∞
δD_S	Drift since last calibration	0.2	Rectangular	√3	1.0	0.116	∞
f_{DC}	Instrumentation linearity	0.1	Normal	2.0	1.0	0.05	∞
f_{REF}	Stability of 50 MHz reference	0.2	Rectangular	√3	1.0	0.116	∞
	Mismatch:						
M_1	Standard sensor at 50 MHz	0.08	U-shaped	√2	1.0	0.06	∞
M_2	Unknown sensor at 50 MHz	0.08	U-shaped	√2	1.0	0.06	∞
M_3	Standard sensor at 18 GHz	1.40	U-shaped	√2	1.0	0.99	∞
M_4	Unknown sensor at 18 GHz	1.68	U-shaped	√2	1.0	1.19	∞
K_X	Repeatability of indication	0.37	Normal	1.0	1.0	0.37	3
$u_c(K_X)$	Combined standard uncertainty		Normal			1.69	>500
$U_{95\%}(K_X)$	Expanded uncertainty		Normal ($k = 2$)			3.39	>500

K2.10 Reported result

The measured calibration factor at 18 GHz was **93.2 % ± 3.4 %.**

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95%. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

NOTES:

1 For the measurement of calibration factor, the uncertainty in the absolute value of the 50 MHz reference source need not be included if the standard and unknown sensors are calibrated using the same source, within the timescale allowed for its short-term stability.

2 This example illustrates the significance of mismatch uncertainty in measurements at relatively high frequencies.

3 In a subsequent use of a sensor further uncertainty contributions may arise due to the use of different connector pairs.

K.3 Measurement of a 30 dB coaxial attenuator

K.3.1 The measurement involves the calibration of a coaxial step attenuator at a frequency of 10 GHz using a dual channel 30 MHz IF substitution measurement system. The measurement is made in terms of the attenuation in dB between a matched source and load from:

$$A_X = A_b - A_a + \delta A_{\mathrm{IF}} + \delta D_{\mathrm{IF}} + \delta L_{\mathrm{M}} + \delta R_{\mathrm{D}} + \delta M + \delta A_{\mathrm{L}}$$

where

A_X = value for the attenuator,

A_b = indicated attenuation with unknown attenuator set to zero,

A_a = indicated attenuation with unknown attenuator set to 30 dB,

δA_{IF} = error due to calibration of reference IF attenuator,

δD_{IF} = drift in reference IF attenuator since last calibration,

δL_M = departure from linearity of mixer,

δR_D = error due to resolution of detection system when measuring A_b or A_a ,

δM = mismatch error,

δA_L = effect of signal leakage.

K.3.2 The result is corrected for the calibrated value of the IF attenuator using the results from a calibration certificate, which gave an uncertainty of ±0.005 dB at a coverage probability of 95 % $(k = 2)$.

K.3.3 No correction is made for the drift of the IF attenuator. The limits of ±0.002 dB were estimated from the results of previous calibrations.

K.3.4 No correction is made for mixer non-linearity. The uncertainty was estimated from a series of linearity measurements over the dynamic range of the system to be ±0.002 dB / 10 dB. An uncertainty of ±0.006 dB was therefore assigned at 30 dB. The probability distribution is assumed to be rectangular.

K.3.5 The resolution of the detection system was estimated to cause possible rounding errors of up to one-half of one least significant recorded digit i.e., ±0.005 dB. This occurs twice - once for the 0 dB reference setting and again for the 30 dB measurement. Two identical rectangular distributions with semi-range limits of *a* combine to give a triangular distribution with semi-range limits of $2a$. The uncertainty due to resolution is therefore 0.01 dB with a triangular distribution.

K.3.6 No correction is made for mismatch error. The mismatch uncertainty is calculated from the scattering coefficients using the equation given at E.9.5. The values used were as follows:

$\Gamma_L = 0.03 \qquad \Gamma_G = 0.03$

$s_{11a} = 0.05 \qquad s_{11b} = 0.05 \qquad s_{22a} = 0.05 \qquad s_{22b} = 0.01 \qquad s_{21a} = 1 \qquad s_{21b} = 0.31$

K.3.7 Special experiments were performed to determine whether signal leakage had any significant effect on the measurement system. No effect greater than ±0.001 dB could be observed for attenuation values up to 70 dB. The probability distribution is assumed to be rectangular.

K.3.8 Type A evaluation: Four measurements were made which involved setting the reference level with the step attenuator set to zero and then measuring the attenuation for the 30 dB setting. The results were as follows:

30.04 dB, 30.07 dB, 30.03 dB, 30.06 dB

From equation (3), the mean value $\overline{A_X}$ = 30.050 dB

The repeatability standard deviation $s(A_X)$ is estimated from the four measured values. Applying equation (5), gives $s(A_X) = 0.018$ dB.

The reported measurement result is to be calculated using the mean of $n = 4$ measurements. So, from equation (4), the repeatability uncertainty is

$$u_{\text{rep}}(A_X) = \frac{s(A_X)}{\sqrt{n}} = \frac{0.018}{\sqrt{4}} = 0.009 \text{ dB}$$

K.3.9 Summary table for $A_X = A_b - A_a + \delta A_{\text{IF}} + \delta D_{\text{IF}} + \delta L_{\text{M}} + \delta R_{\text{D}} + \delta M + \delta A_{\text{L}}$

Quantity	Source of uncertainty	Uncertainty / dB	Probability distribution	Divisor	c_i	$u_i(K_X)$ / dB	ν_i or ν_{eff}
δA_{IF}	Calibration of reference attenuator	0.005	Normal	2.0	1.0	0.0025	∞
δD_{IF}	Drift since last calibration	0.002	Rectangular	√3	1.0	0.0012	∞
δL_{M}	Mixer non-linearity	0.006	Rectangular	√3	1.0	0.0035	∞
δR_{D}	Resolution of indication	0.010	Triangular	√6	1.0	0.0041	∞
δM	Mismatch	0.022	Normal	1	1.0	0.022	∞
δA_{L}	Signal leakage effects	0.001	Rectangular	√3	1.0	0.0006	∞
A_X	Repeatability of indication	0.009	Normal	1.0	1.0	0.009	3
$u_{\text{c}}(A_X)$	Combined standard uncertainty		Normal			0.0245	>150
$U_{95\%}(A_X)$	Expanded uncertainty		Normal ($k = 2$)			0.0491	>150

Note that A_b and A_a do not need to appear in the table as all uncertainty associated with these quantities is accounted for by other terms.

K.3.10 Reported result

The measured value of the 30 dB attenuator at 10 GHz was **30.050 dB ± 0.049 dB**.

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

NOTES:

1 Combination of relatively small uncertainties expressed in dB is permissible since $\log_e(1+x) \approx x$ when x is small and $2.303\log_{10}(1+x) \approx x$. For example: 0.1 dB corresponds to a power ratio of 1.023 and $2.303\log_{10}(1+0.023) = 0.0227$.

Thus, relatively small uncertainties expressed in dB may be combined in the same way as those expressed as linear relative values, e.g., percentage.

2 For attenuation measurements, the probability distribution for RF mismatch uncertainty is dependent on the combination of at least three mismatch uncertainties and can be treated as having a normal distribution. For further details see paragraph E.9.5.

3 In a subsequent use of an attenuator further uncertainty contributions may arise due to the use of different connector pairs.

K.4 Calibration of a weight of nominal value 10 kg of OIML Class M1

K.4.1 The calibration is carried out using a mass comparator whose performance characteristics have previously been determined, and a weight of OIML Class F2. The unknown weight is obtained from:

$$W_X = W_S + \Delta W + \delta D_S + \delta I_d + \Delta C + \Delta A_b + \delta W_r$$

where

W_X	=	value of unknown weight,
W_S	=	calibration of the standard weight,
ΔW	=	measured weight difference weight,
W_S	=	calibration of the standard weight,
δD_S	=	drift of standard weight since last calibration,
δI_d	=	the rounding error of the value of the least significant digit of the two indications,
ΔC	=	comparator non-linearity,
ΔA_b	=	correction for air buoyancy,
δW_r	=	repeatability error.

K.4.2 The calibration certificate for the standard mass gives an uncertainty of 30 mg at a coverage probability of approximately 95 % ($k = 2$).

K.4.3 The allowed monitored drift limits for the standard mass have been set equal to the expanded uncertainty of its calibration, and are ±30 mg. A rectangular probability distribution has been assumed.

K.4.4 The least significant digit I_d for the mass comparator represents 10 mg. Digital rounding has limits of $\pm 0.5\, I_d$ for the indication of the values of both the standard and the unknown weights. Combining these two rectangular distributions gives a triangular distribution, with uncertainty limits of $\pm I_d$, that is ±10 mg.

K.4.5 The linearity error of the comparator over the 2.5 g range permitted by the laboratory's procedures for the comparison was estimated from previous measurements to have limits of ±3 mg. A rectangular probability distribution has been assumed.

K.4.6 No correction is made for air buoyancy, for which limits were estimated to be ±1 ppm of nominal value i.e., ±10 mg. A rectangular probability distribution has been assumed.

K.4.7 A previous Type A evaluation of the repeatability of the measurement process, comprising $m = 10$ comparisons between the standard and unknown weight, gave a repeatability standard deviation, $s(\delta W_r)$, of 8.7 mg, with $\upsilon = (m - 1) = 10$ degrees of freedom.

This evaluation replicates the normal variation in positioning single weights on the comparator, and therefore includes effects due to eccentricity errors.

K.4.8 Three results were obtained for the unknown weight using the conventional technique of bracketing the reading with two readings for the standard. The results were as follows:

No.	Weight on pan	Comparator reading	Standard mean	unknown - standard
	standard	+ 0.01 g		
1	unknown	+ 0.03 g	+ 0.015 g	+ 0.015 g
	standard	+ 0.02 g		
2	unknown	+ 0.04 g	+ 0.015 g	+ 0.025 g
	standard	+ 0.01 g		
3	unknown	+ 0.03 g	+ 0.010 g	+ 0.020 g
	standard	+ 0.01 g		
			Mean difference:	$\Delta W = +0.020$ g

From the calibration certificate, the mass of the standard is $10\ 000.005$ g. The calibrated value of the unknown is therefore $W_X = 10\ 000.005\text{ g} + 0.020\text{ g} = 10\ 000.025\text{ g}$.

K.4.9 The reported measurement result is to be calculated using the mean of $n = 3$ measurements (since *three* comparisons between standard and unknown are made). So, from equation (4), the repeatability uncertainty is

$$u_{\text{rep}}(\delta W_r) = \frac{s(\delta W_r)}{\sqrt{n}} = \frac{8.7}{\sqrt{3}} = 5.0\text{ mg}$$

K.4.10 Summary table for $W_X = W_S + \Delta W + \delta D_S + \delta I_d + \Delta C + \Delta A_b + \delta W_r$

Quantity	Source of uncertainty	Uncertainty /mg	Probability distribution	Divisor	c_i	$u_i(W_X)$ /mg	ν_i or ν_{eff}
W_S	Calibration of standard weight	30.0	Normal	2.0	1.0	15.0	∞
D_S	Drift since last calibration	30.0	Rectangular	√3	1.0	17.32	∞
δI_d	Digital rounding error, comparison	10.0	Triangular	√6	1.0	4.08	∞
ΔC	Comparator non-linearity	3.0	Rectangular	√3	1.0	1.73	∞
ΔA_b	Air buoyancy (1 ppm of nominal value)	10.0	Rectangular	√3	1.0	5.77	∞
δW_r	Repeatability of indication	5.0	Normal	1.0	1.0	5.02	9
$u_c(W_X)$	Combined standard uncertainty		Normal			24.56	>500
U	Expanded uncertainty		Normal ($k = 2$)			49.12	>500

Note that all uncertainty associated with ΔW is accounted for by δI_d and δW_r.

K.4.11 Reported result

The measured value of the 10 kg weight was **10 000.025 g ± 0.049 g**.

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

NOTE: The degrees of freedom shown in the uncertainty budget are derived from a previous evaluation of repeatability, for which 10 readings were used (see paragraph B4).

K.5 Calibration of a weighing machine of 205 g capacity by 0.1 mg digit

K.5.1 The calibration is carried out using weights of OIML Class E2. Checks will normally be carried out for linearity of response across the nominal capacity of the weighing machine, eccentricity effects of the positioning of weights on the load receptor, and repeatability of the machine for repeated weighing near full load. The span of the weighing machine has been adjusted using its internal weight before calibration. The following uncertainty evaluation is carried out for a near full loading of 200 g.
The machine indication errors are obtained from

$$\Delta I_X = I_X - W_S + \delta D_S + \delta I_{d0} + \delta I_d + \Delta A_b + \delta I_r$$

where

ΔI_X	=	measurement error ('error of indication') for indication I_X,
I_X	=	indication (for a weight of 'unknown' value),
W_S	=	weight of the standard,
δD_S	=	drift of standard since last calibration,
δI_{d0}	=	the rounding error of the value of one digit at the zero reading,
δI_d	=	the rounding error of the value of one digit of the indicated value,
ΔA_b	=	correction for air buoyancy,
δI_r	=	repeatability error of the indication.

NOTE: In practice, there may be other sources of error to consider, for example:
a) when it is not possible to apply the load centrally, see EURAMET CG-18, 7.1.1.4
b) when the test load is partially made up of substitution loads, see EURAMET CG-18, 7.1.2.6

K.5.2 The calibration certificate for the stainless steel 200 g standard mass gives an uncertainty of 0.1 mg at a coverage probability of approximately 95 % ($k = 2$).

K.5.3 No correction is made for drift, but the calibration interval is set so as to limit the drift to ±0.1 mg. The probability distribution is assumed to be rectangular.

K.5.4 No correction can be made for the rounding due to the resolution of the digital display of the machine when zeroing. The least significant digit on the range being calibrated corresponds to 0.1 mg and there is therefore a possible rounding error of ±0.05 mg. The probability distribution is assumed to be rectangular.

NOTE: It is often the case that when a weighing machine is zeroed, or tared, it may do so to a greater resolution than that provided by the digital readout. The above contribution may be reduced where justified, for example by determining the error at zero By OIML R76-1, A.4.2.3.2.

K.5.5 No correction can be made for the rounding due to the resolution of the digital display of the machine when loaded. The least significant digit on the range being calibrated corresponds to 0.1 mg and there is therefore a possible rounding error of ±0.05 mg. The probability distribution is assumed to be rectangular.

K.5.6 No correction is made for air buoyancy. As the span of the machine was adjusted with its internal weight before calibration, the uncertainty limits were estimated to be ±1 ppm of the nominal value, i.e., ±0.2 mg.

NOTE: - 1 ppm (1 mg/kg) is only appropriate where the air density is in the range 1.1 to 1.3 kg/m^3 and the weights are within the range specified in OIML R111, table B.7 for stainless steel. The input for weights of other materials may differ, for example cast iron – 3 ppm and aluminium - 30 ppm.

K.5.7 The repeatability of the machine was established from a series of $m = 10$ readings (Type A evaluation), which gave a repeatability standard deviation, $s(I_r)$, of 0.05 mg, with $\upsilon = (m - 1) = 9$ degrees of freedom.

K.5.8 Only one reading was taken to establish the weighing machine indication for each linearity and eccentricity point. For this calibration point the indication, I_X , was 199.9999 g when the 200 g standard mass was applied.

The reported measurement result is to be this single, $n = 1$ measurement. So, from equation (4), the repeatability uncertainty is:

$$u_{\text{rep}}(\delta I_r) = \frac{s(\delta I_r)}{\sqrt{n}} = \frac{0.05}{\sqrt{1}} = 0.05 \text{ mg.}$$

K.5.9 Summary table for $\Delta I_X = I_X - W_S + \delta D_S + \delta I_{d0} + \delta I_d + \Delta A_b + \delta I_r$

Quantity	Source of uncertainty	Uncertainty / mg	Probability distribution	Divisor	c_i	$u_i(I_X)$ / mg	υ_i or υ_{eff}
W_S	Calibration of standard weight	0.1	Normal	2.0	1.0	0.05	∞
δD_S	Drift since last calibration	0.1	Rectangular	√3	1.0	0.058	∞
δI_{d0}	Digital rounding error (at zero)	0.05	Rectangular	√3	1.0	0.029	∞
δI_d	Digital rounding error (for indicated value)	0.05	Rectangular	√3	1.0	0.029	∞
ΔA_b	Air buoyancy (1 ppm of nominal value)	0.2	Rectangular	√3	1.0	0.115	∞
δI_r	Repeatability of indication	0.05	Normal	1.0	1.0	0.05	9
$u_c(I_X)$	Combined standard uncertainty		Normal			0.150	>500
U	Expanded uncertainty		Normal ($k = 2$)			0.300	>500

K5.10 Reported result

For an applied weight of 200 g:
the error of indication of the weighing machine was **0.10 mg ± 0.30 mg**

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

K.6 Calibration of a grade 2 gauge block of nominal length 10 mm

K.6.1 The calibration was carried out using a comparator with reference to a grade K standard gauge block of similar material. The length of the unknown gauge block, L_X, was determined from

$$L_X = L_S + \delta L_D + \Delta L - [L(\alpha\,\Delta t + \Delta\alpha\,\Delta T\,)] + \Delta D_C + \delta C + L_{V(X)} + \delta L_r$$

where

L_S	=	certified length of the standard gauge block at 20ºC,
δL_D	=	drift with time of certified length of standard gauge block,
ΔL	=	measured difference in length,
α	=	mean thermal expansion coefficient of the standard and unknown gauge blocks,
Δt	=	difference in temperature between the standard and unknown gauge blocks,
$\Delta\alpha$	=	difference in thermal expansion coefficients of the standard and unknown gauge blocks,
ΔT	=	difference in mean temperature of gauge blocks and reference temperature of 20 ºC when ΔL is determined,
ΔD_C	=	discrimination and linearity of the comparator,
δC	=	difference in coefficient of compression of standard and unknown gauge blocks,
$L_{V(X)}$	=	variation in length with respect to the measuring faces of the unknown gauge block,
δL_r	=	repeatability error of measurement.

K.6.2 The value of L_S was obtained from the calibration certificate for the standard gauge block. The associated uncertainty was 0.03 µm ($k = 2$).

K.6.3 The change in value L_D of the standard gauge block with time (drift) was estimated from previous calibrations to be zero with an uncertainty of ±15 nm. From experimental evidence and prior experience, the value of zero was considered the most likely, with diminishing probability that the value approached the limits. A triangular distribution was therefore assigned to this uncertainty contribution.

K.6.4 The coefficient of thermal expansion applicable to each gauge block was assumed to have a value, α, of 11.5 µm m^{-1} ºC^{-1} with limits of ±1 µm m^{-1} ºC^{-1}. Combining these two rectangular distributions, the difference in thermal expansion coefficient between the two blocks, is ±2 µm m^{-1} ºC^{-1} with a triangular distribution. For L = 10 mm this corresponds to ±20 nm ºC^{-1}. This difference will have two influences:

(a) The difference in temperature, δt, between the standard and unknown gauge blocks was estimated to be zero with limits of ±0.08 ºC, giving rise to a length uncertainty of ±1.6 nm.

(b) The difference, δT, between the mean temperature of the two gauge blocks and the reference temperature of 20 ºC was measured to be zero and was assigned limits of ±0.2 ºC, giving rise to a length uncertainty of ±4 nm.

As the influence of the expansion coefficient appears directly in both of these uncertainty contributions they are considered to be correlated and, in accordance with paragraph D.3.2, the corresponding uncertainties have been added before being combined with the remaining contributions. This is included in the uncertainty budget as $\delta T_{S,X}$.

K.6.5 The error due to the discrimination and non-linearity of the comparator, ΔD_C, was taken as zero with limits of ±0.05 µm, assessed from previous measurements. Similarly, the difference in elastic compression δC between the standard and unknown gauge blocks was estimated to be zero with limits of ±0.005 µm.

K.6.6 The variation in length of the unknown gauge block, $L_{V(X)}$, was considered to comprise two components:

(a) Effect due to incorrect central alignment of the probe; assuming this misalignment was within a circle of radius 0.5 mm, calculations based on the specifications for grade C gauge blocks indicted uncertainty limits of ±17 nm.

(b) Effects due to surface irregularities such as scratches or indentations; such effects have a detection limit of approximately ±25 nm when examined by experienced staff.

Little is known about the PDF for either effect, but it is considered very unlikely that the extremes of both effects will be encountered. Their combined effect is therefore modelled by a triangular PDF with limits of ±42 nm.

K.6.7 The repeatability of the calibration process was established from previous measurements using gauge blocks of similar type and nominal length. This Type A evaluation, based upon $m = 11$ measurements and using equation (5), yielded a repeatability standard deviation $s(\delta L_r)$ of 16 nm with $\upsilon = (m-1) = 10$ degrees of freedom.

K.6.8 The calibration of the unknown gauge block was established from a single measurement; however, as the conditions were the same as for the previous evaluation of repeatability, the standard uncertainty due to repeatability can be obtained from equation (4) using the previous estimate of repeatability standard deviation $s(\delta L_r)$, and $n = 1$ (because only one reading is made for the actual calibration).

So, from equation (4), the repeatability uncertainty is

$$u_{\mathrm{rep}}(L_r) = \frac{s(\delta L_r)}{\sqrt{n}} = \frac{16}{\sqrt{1}} = 16\ \mathrm{nm}$$

K.6.9 The measured length of the unknown gauge block was $L_X = L_S + \Delta L = 9.999\,940\ \mathrm{mm}$.

K.6.10 Summary table for $L_X = L_S + \delta L_D + \Delta L - [L(\alpha\,\Delta t + \Delta\alpha\,\Delta T)] + \Delta D_C + \delta C + L_{V(X)} + \delta L_r$

Quantity	Source of uncertainty	Uncertainty /nm	Probability distribution	Divisor	c_i	$u_i(L_X)$ /nm	υ_i or υ_{eff}
L_S	Calibration of the standard gauge block	30	Normal	2.0	1	15.0	∞
δL_D	Drift since last calibration	15	Triangular	√6	1	6.1	∞
ΔD_C	Comparator	50	Rectangular	√3	1	28.9	∞
δC	Difference in elastic compression	5.0	Rectangular	√3	1	2.9	∞
$\delta T_{S,X}$	Temperature effects	5.6	Triangular	√6	1	2.3	∞
$L_{V(X)}$	Length variation of unknown gauge block	42	Triangular	√6	1	17.2	∞
δL_r	Repeatability	16	Normal	1.0	1	16.0	10
$u_c(L_X)$	Combined standard uncertainty		Normal			40.7	>350
U	Expanded uncertainty		Normal ($k = 2$)			81.5	>350

Note that all uncertainty associated with ΔL is accounted for by ΔD_C and δL_r.

K.6.11 Reported result

The measured length of the gauge block was **9.999 940 mm ± 0.081 µm**.

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

K.7 Calibration of an N-type thermocouple at 1000 ºC

K.7.1 An N-type thermocouple is calibrated against two reference standard R-type thermocouples in a horizontal furnace with all three measuring (hot) junctions at a temperature of 1000 ºC. The EMFs generated by the thermocouples are measured using a digital microvoltmeter via a selector/reversing switch. All the thermocouples have their reference (cold) junctions at 0 ºC. The unknown thermocouple is connected to the reference point using compensating cables.

K.7.2 The temperature t_X of the hot junction of the unknown thermocouple is given by

$$t_X = t_S\left(V_{iS} + \delta V_{iS1} + \delta V_{iS2} + \delta V_R - \frac{\delta t_{0S}}{C_{S0}}\right) + \delta t_D + \delta t_F + \delta t_r$$

$$\approx t_S(\,V_{iS}) + C_S\,\delta V_{iS1} + C_S\,\delta V_{iS2} + C_S\,\delta V_R - \frac{C_S}{C_{S0}}\delta t_{0S} + \delta t_D + \delta t_F + \delta t_r$$

The voltage $V_X(t)$ across the thermocouple wires with the reference junction at 0 ºC during the calibration is

$$V_X(t) \approx V_X(t_X) + \frac{\Delta t}{C_X} - \frac{\delta t_{0X}}{C_{X0}} + \delta V_r = V_{iX} + \delta V_{iX1} + \delta V_{iX2} + \delta V_{LX} + \delta V_{TH} + \delta V_R + \frac{\delta t}{C_X} - \frac{\delta t_{0X}}{C_{X0}} + \delta V_r$$

where

$t_S(V)$	=	temperature of the reference thermometer in terms of voltage with the cold junction at 0 ºC. The function is given in the calibration certificate,
V_{iS}, V_{iX}	=	indication of the microvoltmeter,
δV_{iS1}, δV_{iX1}	=	voltage corrections due to the calibration of the microvoltmeter,
δV_{iS2}, δV_{iX2}	=	rounding errors due to the resolution of the microvoltmeter,
δV_R	=	voltage error due to contact effects of the reversing switch,
δt_{0S}, δt_{0X}	=	temperature corrections associated with the reference junctions,
C_S, C_X	=	sensitivity coefficients of the thermocouples for voltage at the measurement temperature of 1000 ºC,
C_{S0}, C_{X0}	=	sensitivity coefficients of the thermocouples for voltage at the reference temperature of 0 ºC,
δt_D	=	drift of the reference thermometers since the last calibration,
δt_F	=	temperature correction due to non-uniformity of the furnace,
t	=	temperature at which the unknown thermocouple is to be calibrated,
$\Delta t = t - t_X$	=	deviation of the temperature of the calibration point from the temperature of the furnace,
δV_{LX}	=	voltage error due to the compensation leads,
δV_{TH}	=	error due to inhomogeneity of the unknown thermocouple,
δt_r	=	repeatability error for established reference temperatures,
δV_r	=	repeatability error of voltage measurements.

K.7.3 The reported result is the output EMF of the test thermocouple at the temperature of the hot junction. The measurement process consists of two parts - determination of the temperature of the furnace and determination of the EMF of the test thermocouple. The evaluation of uncertainty has therefore been split into two parts to reflect this situation.

K.7.4 The R-type reference thermocouples are supplied with calibration certificates that relate the temperature of their hot junctions with their cold junctions at 0 ºC to the voltage across their wires. The expanded uncertainty is 0.3 ºC with a coverage factor $k = 2$.

K.7.5 No correction is made for drift of the reference thermocouples since the last calibration but an uncertainty of ±0.3 ºC has been estimated from previous calibrations. A rectangular probability distribution has been assumed.

K.7.6 The voltage sensitivity coefficients of the reference and unknown thermocouples have been obtained from reference tables as follows:

Thermocouple	Sensitivity coefficient at temperatures of	
	0 ºC	1000 ºC
Reference (R-type)	c_{S0} = 0.189 ºC/μV	c_S = 0.077 ºC/μV
Unknown (N-type)	c_{X0} = 0.038 ºC/μV	c_X = 0.026 ºC/μV

K.7.7 The least significant digit of the microvoltmeter corresponds to a value of 1 μV. This results in possible rounding errors, δV_{iS2} and δV_{iX2}, of ±0.5 μV for each indication.

K.7.8 Corrections were made to the microvoltmeter readings by using data from its calibration certificate. Drift and other influences were all considered negligible, therefore only the calibration uncertainty of 2.0 μV ($k = 2$) is to be included in the uncertainty budget.

K.7.9 Residual parasitic offset voltages due to the switch contacts were estimated to be zero within ±2.0 μV.

K.7.10 The temperature of the reference junction of each thermocouple is known to be 0 ºC within ±0.1 ºC. For the 1000 ºC measurements, the sensitivity coefficient associated with the uncertainty in the reference junction temperature is the ratio of those at 0 ºC and 1000 ºC, i.e., $\frac{c_s}{c_{so}} = -0.407$.

K.7.11 The temperature gradients inside the furnace had been measured. At 1000 ºC deviations from non-uniformity of temperature in the region of measurement are within ±1 ºC.

K.7.12 The compensation leads had been tested in the range 0 ºC to 40 ºC. Voltage differences between the leads and the thermocouple wires were estimated to be less than ±5 μV.

K.7.13 The error due to inhomogeneity of the unknown thermocouple was determined during the calibration by varying the immersion depth. Corrections are not practical for this effect therefore the error was assumed to be zero within ±0.3 ºC.

K.7.14 The sequence of measurements is as follows:

1. First standard thermocouple
2. Unknown thermocouple
3. Second standard thermocouple
4. Unknown thermocouple
5. First standard thermocouple

The polarity is then reversed, and the sequence is repeated. Four readings are thus obtained for all the thermocouples. This sequence reduces the effects of drift in the thermal source and parasitic thermocouple voltages. The results were as follows:

Thermocouple:	First standard	Unknown	Second standard
Corrected voltages:	+ 10500 µV + 10503 µV – 10503 µV – 10504 µV	+ 36245 µV + 36248 µV – 36248 µV – 36251 µV	+ 10503 µV + 10503 µV – 10505 µV – 10505 µV
Absolute mean values:	10502.5 µV	36248 µV	10504.0 µV
Temperature of hot junctions:	1000.4 ºC		1000.6 ºC
Mean temperature of furnace:		1000.5 ºC	

K.7.15 The thermocouple output EMF is corrected for the difference between the nominal temperature of 1000 ºC and the measured temperature of 1000.5 ºC. The reported thermocouple output is $V_X = 36248 \times \frac{1000}{1000.5}\ \mu\text{V} = 36230\ \mu\text{V}$.

K.7.16 In this example it is assumed that the procedure requires that the difference between the two standards must not exceed 0.3 ºC. If this is the case, then the measurement must be repeated and/or the reason for the difference investigated.

K.7.17 From the four readings on each thermocouple, one observation of the mean voltage of each thermocouple was deduced. The mean voltages of the reference thermocouples are converted to temperature observations by means of temperature/voltage relationships given in their calibration certificates. These temperature values are highly correlated. By taking the mean they are combined into one observation of the temperature of the furnace at the location of the test thermocouple. In a similar way one observation of the voltage of the test thermocouple is extracted.

K.7.18 To determine the repeatability standard deviation associated with these measurements Independent, Type A evaluations had been carried out on a previous occasion. In each case a series of $m = 10$ measurements had been undertaken at the same temperature of operation. Using equation (5), this gave estimates of the repeatability standard deviations for the temperature of the furnace, $s(\delta t_r)$ of 0.10 ºC and for the voltage of the thermocouple to be calibrated, $s(\delta V_r)$, of 1.6 µV.
In each case there are $\upsilon = (m - 1) = 9$ degrees of freedom

Only one calibration measurement is made, so from equation (4), the repeatability uncertainties are

$$u_{\text{rep}}(\delta t_r) = \frac{s(\delta t_r)}{\sqrt{n}} = \frac{0.10}{\sqrt{1}} = 0.10\ ^\circ\text{C}$$

and

$$u_{\text{rep}}(\delta V_r) = \frac{s(\delta V_r)}{\sqrt{n}} = \frac{1.6}{\sqrt{1}} = 1.6\ \mu\text{V}$$

(The value of $n = 1$ is used to calculate the standard uncertainty because in the normal procedure only one sequence of measurements is made at each temperature.)

K.7.19 Summary table - temperature of the furnace

$$t_X = t_S\, V_{iS} + c_S\, \delta V_{iS1} + c_S\, \delta V_{iS2} + c_S\, \delta V_R - \frac{c_S}{c_{S0}} \delta t_{0S} + \delta t_D + \delta t_F + \delta t_r$$

Quantity	Source of uncertainty	Uncertainty	Probability distribution	Divisor	c_i	$u_i(t_X)$ / ºC	ν_i or ν_{eff}
t_S	Calibration of standard thermocouples	0.3 ºC	Normal	2.0	1.0	0.150	∞
δt_D	Drift in standard thermocouples	0.3 ºC	Rectangular	√3	1.0	0.173	∞
δV_{iS1}	Voltmeter calibration and drift	2.0 µV	Normal	2.0	0.077	0.077	∞
δV_R	Switch contacts	2.0 µV	Rectangular	√3	0.077	0.089	∞
δt_{0S}	Determination of reference point	0.1 ºC	Rectangular	√3	-0.407	-0.024	∞
δt_r	Repeatability	0.1 ºC	Normal	1.0	1.0	0.100	9
δV_{iS2}	Voltmeter resolution	0.5 µV	Rectangular	√3	0.077	0.022	∞
δt_F	Furnace non-uniformity	1.0 ºC	Rectangular	√3	1.0	0.577	∞
$u_c(t_X)$	Combined standard uncertainty		Normal			0.641	>500

K.7.20 Summary table - EMF of unknown thermocouple

$$V_X(t) = V_{iX} + \delta V_{iX1} + \delta V_{iX2} + \delta V_{LX} + \delta V_{TH} + \delta V_R + \frac{\delta t}{c_X} - \frac{\delta t_{0X}}{c_{X0}} + \delta V_r$$

Quantity	Source of uncertainty	Uncertainty	Probability distribution	Divisor	c_i	$u_i(V)$ / µV	ν_i or ν_{eff}
t_X	Furnace temperature	0.641 ºC	Normal	1.0	38.5	24.65	>500
δV_{LX}	Compensating leads	5.0 µV	Rectangular	√3	1.0	2.89	∞
δV_{iX1}	Voltmeter calibration and drift	2.0 µV	Normal	2.0	1.0	1.00	∞
δV_R	Switch contacts	2.0 µV	Rectangular	√3	1.0	1.15	∞
δt_{0X}	Determination of reference point	0.1 ºC	Rectangular	√3	26.3	1.52	∞
δV_r	Repeatability	1.6 µV	Normal	1.0	1.0	1.60	9
δV_{iX2}	Voltmeter resolution	0.5 µV	Rectangular	√3	1.0	0.29	∞
δV_{TH}	Inhomogeneity of thermocouple	0.3 ºC	Rectangular	√3	38.5	6.66	∞
$u_c(V)$	Combined standard uncertainty		Normal			25.9	>500
U	Expanded uncertainty		Normal ($k = 2$)			51.9	>500

K.7.21 Reported result

The N-type thermocouple generates an EMF of **36 230 µV ± 52 µV** at a temperature of 1000 °C with its reference (cold) junction at a temperature of 0 ºC.

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

K.8 Calibration of a digital pressure indicator (DPI) at a nominal pressure of 2 MPa using a reference hydraulic dead weight tester

K.8.1 The calibration pressure was generated using a dead weight tester (DWT) the performance characteristics of which had previously been determined. The indication was approached with increasing pressure to account for the existence of possible hysteresis in the DPI.
The measurement error for the unknown DPI is obtained from:

$$\Delta I_X = I_X - P_X + \delta I_d + \delta\Delta_r$$

$$\Delta I_X = I_X - \left\{ \frac{\left[\sum m_L g\left(1 - \frac{\rho_a}{\rho_m}\right)\right] + \left(-B_V\left(\rho_f - \rho_a\right)g\right) + \phi \times C}{A_0\left(1 + a_p\right)\left(1 + \lambda(t - 20)\right)} + \left(\rho_f - \rho_a\right)gh + \delta e_V \right\} + \delta I_d + \delta\Delta_r$$

where

ΔI_X	=	Error in the indication of the unknown DPI,
I_X	=	Indication of the DPI,
P_X	=	Generated pressure at the reference level of the DPI,
m_L	=	Mass of the component parts of the load, including the piston,
ρ_a	=	Density of ambient air,
ρ_m	=	Density of the mass, m, and can be significantly different for each load component,
g	=	The value of the local acceleration due to gravity,
h	=	Height different between reference levels for standard and for generated pressure,
B_V	=	Buoyancy volume of the reference piston – from calibration certificate,
ρ_f	=	Density of hydraulic fluid,
ϕ	=	Surface tension coefficient of hydraulic fluid,
C	=	Circumference of reference piston,
A_0	=	Effective area at zero pressure of reference piston,
a_p	=	Distortion coefficient of reference piston (pressure dependant term),
λ	=	Temperature coefficient of piston and cylinder,
δe_V	=	Error due to the piston not being perfectly vertical,
δI_d	=	The rounding error corresponding to the least significant digit of the indication,
$\delta\Delta_r$	=	Repeatability error when measuring ΔI_X.

K.8.2 The calibration certificate for the reference DWT gives the piston area and its expanded uncertainty as:
A_0 = 80.6516 mm^2 ± 0.0026 mm^2.
This results in a relative expanded uncertainty ($k = 2$) in A_0 of 32.2 ppm.

K.8.3 The calibration certificate for the reference DWT gives the distortion coefficient of the reference piston as a_p = 6.0 x 10^{-6}/MPa ± 0.5 x 10^{-6}/MPa, where the coverage interval is defined in terms of (±) the expanded uncertainty, 0.5 ppm/MPa ($k = 2$). The input uncertainty is the product of this uncertainty and the generated pressure, which in the numerical example below is 2 MPa. The corresponding input uncertainty is therefore $0.5\,\frac{\mathrm{ppm}}{\mathrm{MPa}} \times 2\ \mathrm{MPa} = 1.0\ \mathrm{ppm}\ (k = 2)$

K.8.4 The allowed drift limit in the effective area of the DWT, based on results from previous calibrations, has been set to ±30 ppm, to be treated as the limits of a rectangular distribution.

K.8.5 Mass uncertainties

K.8.5.1 The mass of the piston is shown on the calibration certificate as 0.567 227 kg with an expanded uncertainty ($k = 2$) of 0.000 010 kg.

The values of the mass set, as shown on the calibration certificate for the three weights “A”, “B” and “C” used to generate 2 MPa are:

A = 0.255 242 kg ± 0.000 010 kg
B = 7.402 137 kg ± 0.000 050 kg
C = 8.224 784 kg ± 0.000 050 kg

where the coverage intervals are defined in terms of (±) the expanded uncertainty ($k = 2$)

As all calibrated mass values are likely to be strongly correlated the uncertainty for the combined load is found by summing the individual uncertainties. Expressed as a relative value for the 2 MPa load this is:

$$\text{Relative expanded uncertainty in mass} = \frac{U_{\text{piston}} + U_{\text{A}} + U_{\text{B}} + U_{\text{C}}}{\text{Total mass of piston} + \text{A} + \text{B} + \text{C}}$$

$$= \frac{(10 + 10 + 50 + 50)\ \text{mg}}{(0.567\ 227 + 0.255\ 242 + 7.402\ 137 + 8.224\ 784)\ \text{kg}}$$

$$= 7.3\ \text{ppm}$$

to be treated as the semi-range of a rectangular distribution

K.8.5.2 The allowed drift limit of the piston mass, based on previous calibrations, has been set to ±0.000 015 kg.

The allowed limits to the drift of the mass set have been set to be equal to the expanded uncertainty of its calibration, i.e., 10 mg, 50 mg and 50 mg respectively.

Again, as all calibrated mass values are likely to be strongly correlated, the uncertainty for the combined load is found by summing the individual uncertainties. Expressed in relative terms the value is:

$$\text{Relative mass drift limit} = \pm\frac{(15 + 10 + 50 + 50)\text{mg}}{(0.567\ 227 + 0.255\ 242 + 7.402\ 137 + 8.224\ 784)\text{kg}} = 7.6\ \text{ppm}$$

to be treated as the semi-range of a rectangular distribution.

K.8.6 The uncertainty in the temperature of the piston is estimated to be no more than 0.5 ºC. This will affect the pressure generated in proportion to the temperature coefficient of the piston and cylinder combination. In this case a steel piston and cylinder has a temperature coefficient of 23 ppm/ºC. This sensitivity coefficient was obtained from the calibration certificate for the DWT.

K.8.7 A correction has already been made for the value of the local acceleration due to gravity. The correction was estimated from knowledge of the measurement location and known (Bouguer) anomalies. The expanded uncertainty associated with this estimate, is 3 ppm ($k = 2$).

K.8.8 A standard air buoyancy correction has been made assuming an air density of 1.2 kg.m^{-3} and a mass density of 8000 kg.m^{-3}. As all masses have an assumed density of 7800 kg.m^{-3}, and assuming normal laboratory ambient conditions, the additional uncertainty for this approximation is estimated to be 13 ppm, to be treated as the semi-range of a rectangular distribution.

K.8.9 The uncertainty relating to fluid head effects arises from the height different between the reference level of the reference DWT and the datum level for generated pressure, estimated as +2 mm with an uncertainty of 1 mm. No correction is made; therefore, a limit value of 3 mm has been assigned for the uncertainty associated with fluid head effects. Assuming that the density of the oil used is 917 kg/m^3 and the local value of g is 9.81 m.s^{-2}, then the uncertainty associated with the fluid head effect is 917 x 9.81 x 0.003 = 27.0 Pa. In relative terms this corresponds to an uncertainty of 13.5 ppm at 2 MPa, to be treated as the semi-range of a rectangular distribution.

K.8.10 The uncertainty contribution arising from the combined effects of the uncertainty in the knowledge of buoyancy volume, surface tension and fluid density has been estimated as 2 ppm based on a relative uncertainty in each of ±10 %, to be treated as the semi-range of a rectangular distribution.

K.8.11 An uncertainty arises because the piston may not be perfectly vertical. If it were, then all the force would act on the area. Any departure from vertical will reduce the force, and therefore the pressure, by the cosine of the angle. In this example, it is assumed that, after levelling, the piston is vertical to within 0.15°. The effect in terms of generated pressure is proportional to the cosine of the angle from true vertical. The cosine of 0.15° is 0.999 996 6. The maximum error is therefore -3.4 ppm of the generated pressure. This is to be represented by a rectangular distribution with limits of ±3.4 ppm.

NOTE: This effect always acts in one direction, i.e., the generated pressure will always be smaller than that obtained if the piston were truly vertical. As this uncertainty is small compared with others in this particular calibration, it is convenient to treat it as bilateral.

K.8.12 No correction can be made for the rounding error, δI_d, due to the resolution of the digital display of the DPI. The least significant digit on the range being calibrated for this particular DPI changes in steps of 200 Pa and there is therefore a possible rounding error of ±100 Pa or, in relative terms, ±50 ppm. The probability distribution is assumed to be rectangular

K.8.13 The repeatability standard deviation of the calibration process had been established from previous measurements using DPIs of similar type and range. This Type A evaluation, based upon $m = 10$ measurements, yields a relative standard deviation of 16 ppm, with $\upsilon = (m-1) = 9$ degrees of freedom.

This calibration of the unknown DPI was established from a single measurement. As the conditions were the same as previously, the repeatability uncertainty can be obtained using the previously obtained repeatability standard deviation, and $n = 1$ (as only one measurement is made for the calibration).

So, from equation (4), the repeatability uncertainty is

$$u_{\text{rep}}(\delta\Delta_r) = \frac{s(\delta\Delta_r)}{\sqrt{n}} = \frac{16}{\sqrt{1}} = 16 \text{ ppm}$$

K.8.14 All uncertainties associated with the indicated value I_X are accounted for by the error terms δI_d and $\delta\Delta_r$, so no further contributions need be considered.

K.8.15 The generated pressure was calculated from the mass of the piston, that of the mass set, and the following quantities:

Temperature of Piston:	21 ºC
Buoyancy volume of piston:	3.0 x 10^{-7} m^3
Air density:	1.2 $kg.m^{-3}$
Local gravity:	9.811812 $m.s^{-2}$
Oil surface tension coefficient:	0.0315 N/m

This results in an applied pressure of $P_X = 2.000\,806\ \mathrm{MPa}$

K.8.16 Summary table for

$$\Delta I_X = I_X - \left\{ \frac{\left[\sum m_L g \left(1 - \frac{\rho_a}{\rho_m}\right)\right] + \left(-B_V\left(\rho_f - \rho_a\right)g\right) + \phi \times C}{A_0\left(1 + a_p\right)\left(1 + \lambda(t - 20)\right)} + \left(\rho_f - \rho_a\right)gh + \delta e_V \right\} + \delta I_d + \delta \Delta_r$$

Quantity	Source of uncertainty	Uncertainty	Probability distribution	Divisor	c_i	$u_i(\Delta I_X)$ / ppm	ν_i or ν_{eff}
A_0	Calibration of DWT (area)	32.2 ppm	Normal	2.0	1.0	16.1	∞
a_p	Calibration of DWT (distortion)	1.0 ppm	Normal	2.0	1.0	0.500	∞
A_0	Drift in area	30 ppm	Rectangular	√3	1.0	17.32	∞
m_L	Calibration of total load	7.3 ppm	Normal	2	1.0	3.65	∞
m_L	Drift of total load	7.6 ppm	Rectangular	√3	1.0	4.39	∞
t	Temperature of the piston	0.5 ºC	Rectangular	√3	23 ppm/ºC	6.64	∞
g	Local gravity determination	3.0 ppm	Normal	2	1.0	1.50	∞
B_V	Air buoyancy	13 ppm	Rectangular	√3	1.0	7.51	∞
h	Fluid head effects	13.5 ppm	Rectangular	√3	1.0	7.79	∞
ϕ	Other fluid effects	2.0 ppm	Rectangular	√3	1.0	1.15	∞
δe_V	Levelling effects	3.4 ppm	Rectangular	√3	1.0	1.96	∞
δI_d	Digital rounding error	50 ppm	Rectangular	√3	1.0	28.9	∞
$\delta \Delta_r$	Repeatability	16 ppm	Normal	1	1.0	16.0	9
$u_c(\Delta I_X)$	Combined standard uncertainty		Normal			43.0	>450
U	Expanded uncertainty		Normal ($k = 2$)			86.0	>450

:

NOTES

1 Unlike in most of the other examples in this Appendix, each individual source of uncertainty is not represented here by a unique term in the measurement model. The quantity m_L is associated with two sources of uncertainty. This approach is commonplace but is arguably not as good practice as it is to identify each source with an individual quantity in the model.

2 In this example, the uncertainty due to resolution, δI_d, is larger than any other contribution and is assigned a rectangular distribution. Nevertheless, the combined standard uncertainty is still normally distributed, due to the presence of the other uncertainties, even though they are of smaller magnitude. This has been verified by MCS.

3 The resolution uncertainty is based on the least significant digit of the DPI. For this instrument it changes in steps of 2 digits, therefore the uncertainty is ±1 digit.

4 This uncertainty budget has been constructed in relative terms (ppm), as most of the errors that arise are proportional to the generated pressure and it is the convention in this field of measurement to express uncertainties

in this manner. If it is required that the uncertainty is reported in absolute units, it can be calculated from the reported value and the relative uncertainty. In this case, the expanded uncertainty in absolute terms is 0.00017 MPa. See also Appendix Q

K.8.17 Reported result

The pressure was applied in an increasing direction until it reached a final value of 2.000806 MPa. The indication of the digital pressure indicator was 2.00084 MPa.

The corresponding measurement error ΔI_X is **17 ppm ± 86 ppm** (34 Pa $\pm$ 170 Pa)

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

K.9 Volume flowrate expressed at 'standard' pressure and temperature

K.9.1 A laboratory uses a flowmeter to measure the volumetric flowrate at the prevailing pressure and temperature of the transported fluid. The flowrate is to be reported with reference to 'standard' conditions.

K.9.2 The 'standard' flowrate is given by,

$$Q_\mathrm{s} = Q_\mathrm{m}\left(\frac{P_\mathrm{m}}{P_\mathrm{s}}\right)\left(\frac{T_\mathrm{s}}{T_\mathrm{m}}\right) + \delta Q_\mathrm{S}$$

where

Q_s	=	measured flowrate, referred to standard conditions $(P_\mathrm{s}, T_\mathrm{s})$,
Q_m	=	measured flowrate at prevailing fluid conditions $(P_\mathrm{m}, T_\mathrm{m})$,
P_s	=	standard pressure, $1013.25\ \mathrm{hPa}$
T_s	=	standard temperature, $293.15\ \mathrm{K}$
P_m	=	fluid pressure at the meter during flow measurement,
T_m	=	fluid temperature at the meter during flow measurement,
δQ_s	=	repeatability error for replicated measurements of Q_s.

K.9.3 Sensitivity coefficients for the input quantities are:

$$c_{Q_\mathrm{m}} = \frac{\partial Q_\mathrm{s}}{\partial Q_\mathrm{m}} = \left(\frac{P_\mathrm{m}}{P_\mathrm{s}}\right)\left(\frac{T_\mathrm{s}}{T_\mathrm{m}}\right)$$

$$c_{P_\mathrm{m}} = \frac{\partial Q_\mathrm{s}}{\partial P_\mathrm{m}} = Q_\mathrm{m}\left(\frac{1}{P_\mathrm{s}}\right)\left(\frac{T_\mathrm{s}}{T_\mathrm{m}}\right) = \frac{Q_\mathrm{s}}{P_\mathrm{m}}$$

$$c_{T_\mathrm{m}} = \frac{\partial Q_\mathrm{s}}{\partial T_\mathrm{m}} = -Q_\mathrm{m}\left(\frac{P_\mathrm{m}}{P_\mathrm{s}}\right)\left(\frac{T_\mathrm{s}}{{T_\mathrm{m}}^2}\right) = \frac{-Q_\mathrm{s}}{T_\mathrm{m}}$$

$$c_{\delta Q_\mathrm{S}} = \frac{\partial Q_\mathrm{s}}{\partial \delta Q_\mathrm{S}} = 1$$

Sensitivity coefficients are not required for P_s and T_s which are both constant, having zero uncertainty

K.9.4 The flowmeter output signal is in the form of a 4 mA to 20 mA current.

The calibration certificate provides data for reference flowrate and measured current. A calibration function is also given

$$q = A \times\ i + B$$

where i is the measured current,
$A = 6.25\ \mathrm{Ls^{-1}/mA}$, and
$B = -24.95\ \mathrm{Ls^{-1}}$.

K.9.5 In later use, the measured flowrate is given by

$$Q_\mathrm{m} = q_\mathrm{m} - \delta q_\mathrm{d} - \delta q_\mathrm{T} - \delta q_\mathrm{r}$$

where

q_{m} = flowrate indicated by the flowmeter (calculated from measured output current)
δq_{d} = error due to drift of the flowmeter calibration,
δq_{T} = error due to 'environmental' effects (installation, temperature…),
δq_{r} = error due to measurement of the flowmeter output (current).

Following convention, the best estimate of the value for all δ-terms is zero (however the associated uncertainty is not zero).

K.9.5.1 The corresponding sources of uncertainty are:

Flowmeter calibration. The expanded uncertainty ($k = 2$) associated with the calibration function is given as $U(q) = 0.26\ \mathrm{Ls}^{-1}$ for measured currents with negligibly small measurement uncertainty.

Calibration drift. The maximum difference between corresponding points in successive calibrations is $0.23\ \mathrm{Ls}^{-1}$ for the last four recalibrations. This is treated as the limit of a rectangular PDF.

Environmental effects. Contributions for environmental effects such as ambient temperature, position and orientation of the flowmeter are estimated to be no larger than $0.11\ \mathrm{Ls}^{-1}$, which is treated as the limit of a rectangular PDF.

Measurement of current. A separate uncertainty budget for the measurement of current estimates the expanded uncertainty ($k = 2$) to be $0.021\ \mathrm{mA}$, with a normal distribution.

K.9.6 Summary table for evaluation of the measurement uncertainty associated with $Q_{\mathrm{m}} = q_{\mathrm{m}} - \delta q_{\mathrm{d}} - \delta q_{\mathrm{T}} - \delta q_{\mathrm{r}}$

Quantity	Source of uncertainty	Uncertainty	Probability distribution	Divisor	c_i	$u_i(Q_{\mathrm{m}})$ / Ls^{-1}	ν_i or ν_{eff}
q_{m}	Flowmeter calibration	$0.26\ \mathrm{Ls}^{-1}$	Normal	2	1	0.13	∞
δq_{d}	Calibration drift	$0.23\ \mathrm{Ls}^{-1}$	Rectangular	√3	1	0.133	∞
δq_{T}	Environmental effects	$0.11\ \mathrm{Ls}^{-1}$	Rectangular	√3	1	0.064	∞
δq_{r}	Measurement of current	$0.021\ \mathrm{mA}$	Normal	2	$6.25\ \mathrm{Ls}^{-1}/\mathrm{mA}$	0.066	∞
$u_{\mathrm{c}}(Q_{\mathrm{m}})$	Standard uncertainty		Normal			0.207	>500

Repeatability uncertainty is *not* included in this part of the budget as it will be encompassed in the repeatability term δQ_{S} which, in this model, accounts for *all* repeatability errors.

Similar (partial) uncertainty budgets have also been created for measurement of fluid pressure and measurement of fluid temperature at the location of the flowmeter. Again, individual repeatability uncertainty is not included as it will be encompassed in the repeatability term δQ_{S}.

K.9.7.1 $P_\mathrm{m} = p_\mathrm{m} - \Delta p_\mathrm{m} - \delta p_\mathrm{f} - \delta p_\mathrm{d} - \delta p_\mathrm{T}$

where

p_m = pressure indicated by the gauge,

Δp_m = measurement error found during gauge calibration,

δp_f = fitting error (e.g., difference between calibration data (Δp_m) and values obtained from the data for use),

δp_d = error due to drift of the gauge calibration,

δp_T = error due to 'environmental' effects (installation, temperature, etc …),

Following convention, the best estimate of the value for all δ-terms is zero (however the associated uncertainty is not zero).

K.9.7.2 The corresponding sources of uncertainty are:

Resolution of indication: The pressure indication has a resolution of $1\ \mathrm{hPa}$, which is treated as the (full) width of a rectangular PDF.

Calibration certificate. The expanded uncertainty ($k = 2$) associated with the calibration results is given as $U = 1.8\ \mathrm{hPa}$.

Calibration data fitting. Across the range of use of the pressure gauge the measurement error varies between $-0.056\ \mathrm{hPa}$ and $+0.047\ \mathrm{hPa}$. These errors are small in relation to other potential errors, so no correction is to be applied. This 'uncorrected error' is incorporated into the uncertainty evaluation as a rectangular PDF with a half-width of $0.056\ \mathrm{hPa}$

Calibration drift. The maximum difference between corresponding points in successive calibrations is $1.25\ \mathrm{hPa}$ for the last five recalibrations. This is treated as the limit of a rectangular PDF.

Environmental effects. Contributions for environmental effects such as ambient temperature, position and orientation of the pressure gauge are estimated to be no larger than $1\ \mathrm{hPa}$, which is treated as the limit of a rectangular PDF.

K.9.7.3 Summary table for evaluation of the measurement uncertainty associated with $P_\mathrm{m} = p_\mathrm{m} - \Delta p_\mathrm{m} - \delta p_\mathrm{f} - \delta p_\mathrm{d} - \delta p_\mathrm{T}$

Quantity	Source of uncertainty	Uncertainty / hPa	Probability distribution	Divisor	c_i	$u_i(P_\mathrm{m})$ / hPa	ν_i or ν_eff
p_m	Resolution of indication	0.5	Rectangular	√3	1	0.289	∞
Δp_m	Calibration certificate	1.8	Normal	2	1	0.9	∞
δp_f	Calibration data fitting	0.056	Rectangular	√3	1	0.032	∞
δp_d	Calibration drift	1.25	Rectangular	√3	1	0.722	∞
δp_T	Environmental effects	1	Rectangular	√3	1	0.577	∞
$u_\mathrm{c}(P_\mathrm{m})$	Standard uncertainty		Normal			1.322	>500

K.9.8.1 $T_\mathrm{m} = t_\mathrm{m} - \Delta t_\mathrm{m} - \delta t_\mathrm{f} - \delta t_\mathrm{d} - \delta t_\mathrm{T}$

where

t_m = temperature indicated by the thermometer,

Δt_m = measurement error found during thermometer calibration,

δt_f = fitting error (e.g., difference between calibration data (Δt_m) and values obtained from the data for use),

δt_d = error due to drift of the thermometer calibration,

δt_T = error due to 'environmental' effects (installation, temperature, etc …),

Following convention, the best estimate of the value for all δ-terms is zero (however the associated uncertainty is not zero).

K.9.8.2 The corresponding sources of uncertainty are:

Resolution of indication: The thermometer indication has a resolution of $0.1\ \mathrm{K}$, which is treated as the (full) width of a rectangular PDF.

Calibration certificate. The expanded uncertainty ($k = 2$) associated with the calibration results is given as $U = 0.055\ \mathrm{K}$.

Calibration data fitting. Across the range of use of the thermometer the measurement error varies between $-0.43\ \mathrm{K}$ and -$0.47\ \mathrm{K}$. These errors are too large and too biased to remain uncorrected. To avoid using a more complicated fit to the results, the average error is used to correct all thermometer indications. The residual uncorrected error is then no larger than $0.02\ \mathrm{K}$. This is incorporated into the uncertainty evaluation as the limit of a rectangular PDF

Calibration drift. The maximum difference between corresponding points in successive calibrations is $0.13\ \mathrm{K}$ for the last four recalibrations. This is treated as the limit of a rectangular PDF.

Environmental effects. Contributions for environmental effects, such as installation of the thermometer, are estimated to be no larger than $0.2\ \mathrm{K}$, which is treated as the limit of a rectangular PDF.

K.9.8.3 Summary table for evaluation of the measurement uncertainty associated with $T_\mathrm{m} = t_\mathrm{m} - \Delta t_\mathrm{m} - \delta t_\mathrm{f} - \delta t_\mathrm{d} - \delta t_\mathrm{T}$

Quantity	Source of uncertainty	Uncertainty / K	Probability distribution	Divisor	c_i	$u_i(T_\mathrm{m})$ / K	ν_i or ν_eff
t_m	Resolution of indication	0.05	Rectangular	√3	1	0.029	∞
Δt_m	Calibration certificate	0.055	Normal	2	1	0.028	
δt_f	Calibration data fitting	0.02	Rectangular	√3	1	0.012	
δt_d	Calibration drift	0.13	Rectangular	√3	1	0.075	∞
δt_T	Environmental effects	0.2	Rectangular	√3	1	0.115	∞
$u_\mathrm{c}(T_\mathrm{m})$	Standard uncertainty		Normal			0.144	>500

K.9.9 The laboratory has previously established that under normal conditions of use the repeatability standard deviation for replicate measurements of Q_s is $s = 0.18\ \text{Ls}^{-1}$, based upon a set of $m = 25$ measurements, with $\upsilon = m - 1 = 24$ degrees of freedom.

Furthermore, during normal use only a single ($n = 1$) reading is obtained.

The repeatability uncertainty is therefore estimated to be $u_{\text{rep}} = s/\sqrt{n} = 0.18\ \text{Ls}^{-1}$.

K.9.10 Summary table for evaluation of the measurement uncertainty associated with $Q_s = Q_m\left(\frac{P_m}{P_s}\right)\left(\frac{T_s}{T_m}\right) + \delta Q_s$

Quantity	Source of uncertainty	Uncertainty	Probability distribution	Divisor	c_i	$u_i(Q_s)$ / Ls^{-1}	υ_i or υ_{eff}
Q_m	Indicated flowrate	$0.207\ \text{Ls}^{-1}$	Normal	1	$c_{Q_m} = 0.998$	0.207	∞
P_m	Indicated pressure	$1.322\ \text{hPa}$	Normal	1	$c_{P_m} = 0.012\ \text{Ls}^{-1}\text{hPa}^{-1}$	0.016	∞
T_m	Indicated temperature	$0.144\ \text{K}$	Normal	1	$c_{T_m} = -0.043\ \text{Ls}^{-1}\text{K}^{-1}$	0.006	∞
δQ_s	Repeatability, u_{rep}	$0.18\ \text{Ls}^{-1}$	Normal	1	$c_{\delta Q_s} = 1$	0.18	24
$u_c(Q_s)$	Standard uncertainty		Normal			0.275	130
$U_{95\%}(Q_s)$	Expanded uncertainty		Normal ($k = 2$)			0.55	130

K.9.11 Suppose that during a test a flowmeter output of $i = 6.0\ \text{mA}$ is measured at prevailing conditions of $P_m = 1018\ \text{hPa}$ and $T_m = 295.1\ \text{K}$.

The corresponding flowrate is $Q_m = q_m = 12.55\ \text{Ls}^{-1}$ at the prevailing conditions (where q_m is obtained using the linear fit described at K.9.4).

The standardised flowrate is therefore $Q_s = 12.53\ \text{Ls}^{-1}$ at $P_s = 1013.25\ \text{hPa}$ and $T_s = 293.15\ \text{K}$.

K.9.12 Reported result

The measured flowrate is $(\mathbf{12.53 \pm 0.55})\ \mathbf{Ls^{-1}}$, for standard conditions of $P_s = 1013.25\ \text{hPa}$ and $T_s = 293\ \text{K}$.

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

K.10 Thickness of a test sample

K.10.1 A laboratory's procedure requires that the mean thickness w at the centre of a sample taken from a flat sheet of aluminium is to be measured before it is used in a test.
Only the centre thickness is needed for this test.

The procedure specifies the measuring equipment to be used and its method of operation so that the 'centre' is reliably located and any mechanical or geometrical effects are negligible.

The assigned thickness is given by the model

$$w = w_{\mathrm{ave}} + \Delta w_{\mathrm{cal}} + \delta w_{\mathrm{T}}$$

where

w_{ave} = mean indication, for n measurements made at the centre of the sample.
Uncertainty in this quantity arises from errors of precision (repeatability) and resolution,

Δw_{cal} = calibration correction corresponding to w_{ave}.
There is uncertainty in this quantity that is 'brought-in' from the calibration process; uncertainty due to drift in the calibration value over the course of time; and uncertainty that arises because in practice calibration corrections are not made,

δw_{T} = measurement error due thermal effects.
Uncertainty in this quantity arises from the wide difference permitted between ambient operational temperature and the temperature at which the measuring equipment was calibrated.

K.10.2 The procedure requires that $n = 5$ measurements are made.
The indicated values are:

measurement	Indicated thickness, w /mm
1	1.51
2	1.51
3	1.52
4	1.50
5	1.53

The mean thickness is therefore $w_{\mathrm{ave}} = 1.514\ \mathrm{mm}$

The repeatability standard deviation is estimated from the data to be $s(w) = 0.0114\ \mathrm{mm}$, therefore the repeatability uncertainty is

$$u_{\mathrm{rep}}(w) = \frac{s(w)}{\sqrt{n}} = \frac{0.0114\ \mathrm{mm}}{\sqrt{5}} = 0.0051\ \mathrm{mm}$$

K.10.3 The resolution of the measuring equipment is estimated to cause possible rounding errors of one-half of one least significant indicated digit i.e., $\pm 0.005\ \mathrm{mm}$ which describe the limits for a rectangular PDF.

K.10.4 The measuring system is routinely calibrated. The most recent results have a reported expanded uncertainty $(k = 2)$ of $U_{95\%} = 0.015\ \mathrm{mm}$. It is assumed that the calibration results are described by a normal PDF. (For an accredited calibration certificate this is a reasonable assumption, unless otherwise stated).

K.10.5 Drift in the calibration correction is evaluated in terms of the difference between successive calibrations.

The most recent drift value is $0.002\ \mathrm{mm}$, and previous values are $-0.004\ \mathrm{mm}$ and $-0.001\ \mathrm{mm}$.

Since there is very little information available, the lab chooses to evaluate the uncertainty contribution due to drift in terms of a limit value. The numerical value of the limit is established as the larger of a) the most recent absolute drift value, and b) the average of the most recent absolute drift value and the previous two absolute values. For the current year this corresponds to
$\max\big(0.002, \mathrm{ave}(0.002, 0.004, 0.001)\big)\ \mathrm{mm} = \max(0.002, 0.00233)\ \mathrm{mm} = 0.00233\ \mathrm{mm}.$
This is treated as the half width of a rectangular PDF.

This algorithm has the advantages of giving weight to the latest drift value if it proves to be higher than average, whilst not giving long-lasting weight to a previous high value (as would be the case if a simple maximum of current and recent values was used).

NOTE: Caution is needed when applying an algorithm such as this, particularly when only a small amount of data is available. The credibility of the estimate should always be questioned, particularly in cases where it seems very large (perhaps indicating a fault with the instrument), or when it seems very small (since a very low or 'zero' estimate of calibration drift is often not physically realistic, unless it is hidden by the resolution of the measurement).

K.10.6 Calibration errors are small in comparison with their uncertainty and with several other contributions to the combined standard uncertainty. Under these circumstances the lab chooses not to correct the indicated values and instead includes the 'uncorrected errors' in its uncertainty evaluation, using the maximum uncorrected error to define the half-width of a rectangular PDF.

For the most recent calibration this value is $0.001\ \mathrm{mm}$

NOTE: Although this approach is relatively common in practice, for large corrections it is inconsistent with the principles underpinning the GUM framework. The GUM (at 3.2.4) states that within the GUM framework "It is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects."
The GUM, at F.2.4.5, does however describe a consistent approach which involves establishing a mean correction together with an associated uncertainty contribution. (The approach taken by the lab in this example in effect assumes that the mean correction is given by $\Delta w_{\mathrm{cal}} = 0$.)

K.10.7 Thickness measurements are to be performed in an uncontrolled environment in which the measuring equipment is located. The ambient temperature has been found to be as much as $\pm 10\ \mathrm{K}$ away from the temperature at which it was calibrated. No other information is routinely available; therefore, the associated uncertainty is assumed to be characterised by a rectangular PDF with half-width of $10\ \mathrm{K}$.

Based upon a difference in linear thermal expansion coefficient between the steel measuring instrument and the aluminium sheet of $11.9 \times 10^{-6}\ \mathrm{K}^{-1}$, and a maximum measurement thickness of $3\ \mathrm{mm}$, the maximum temperature sensitivity is calculated to be
$c_{\mathrm{T}} = 11.9 \times 10^{-6}\ \mathrm{K}^{-1} \times 3\ \mathrm{mm} = 3.6 \times 10^{-5}\ \mathrm{mm\ K}^{-1}.$

The estimated uncertainty of about 10% in linear expansion coefficients has previously been shown to be negligible for this measurement and is not included in the evaluation.

K.10.8 Summary table for evaluation of the measurement uncertainty associated with $w = w_{\mathrm{ave}} + \Delta w_{\mathrm{cal}} + \delta w_{\mathrm{T}}$

Quantity	Source of uncertainty	Uncertainty	Probability distribution	Divisor	c_i	$u_i(w)$ /mm	υ_i or υ_{eff}
w_{ave}	Repeatability, u_{rep}	0.0051 mm	Normal	1	1	0.00510	4
w_{ave}	Resolution	0.005 mm	Rectangular	√3	1	0.00289	∞
Δw_{cal}	Calibration uncertainty	0.015 mm	Normal	2	1	0.00750	∞
Δw_{cal}	Calibration drift	0.0023 mm	Rectangular	√3	1	0.00135	∞
Δw_{cal}	Uncorrected error	0.001 mm	Rectangular	√3	1	0.00058	∞
δw_{T}	Environmental effects	10 K	Rectangular	√3	3.6×10^{-5} mm/K	0.00021	∞
$u_{\mathrm{c}}(w)$	Combined standard uncertainty		Normal			0.00963	50
$U_{95\%}(w)$	Expanded uncertainty		Normal ($k = 2$)			0.019	50

NOTE: Unlike in most of the other examples in this Appendix, each individual source of uncertainty is not represented here by a unique term in the measurement model.
This approach is commonplace but is arguably not as good practice as it is to identify each source with an individual quantity in the model.

K.10.9 Reported result

The measured thickness was **1.514 mm ± 0.019 mm**

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

K.11 Temperature of gas in a container

K.11.1 A laboratory wishes to measure the temperature of gas in a storage vessel. To make the measurement a PT100 probe, L, will be attached to the outside of the vessel. The probe will periodically be calibrated against a reference standard probe, S, using a dry block calibrator.

K.11.2 When calibrating probe L,

The temperature measured by reference probe S is

$$T_{\mathrm{S}} = I_{\mathrm{S}} - \Delta T_{\mathrm{S}},$$

where

T_{S}	=	measured temperate according to the reference standard S,
I_{S}	=	temperate indicated by the reference standard S,
ΔT_{S}	=	measurement error corresponding to I_{S}.

The associated measurement error determined for the lab probe L is,

$$\Delta T_{\mathrm{L}} = I_{\mathrm{L}} - T_{\mathrm{L}} - \delta\Delta T_{\mathrm{L}},$$

where

ΔT_{L}	=	measurement error corresponding to I_{L},
I_{L}	=	temperature indicated by the lab standard L during its calibration against S,
T_{L}	=	temperature at probe L, corresponding to $T_{\mathrm{L}} = T_{\mathrm{S}} - \Delta T_{\mathrm{SL}} = I_{\mathrm{S}} - \Delta T_{\mathrm{S}} - \Delta T_{\mathrm{SL}}$,
ΔT_{SL}	=	temperature difference between the positions of S and L,
$\delta\Delta T_{\mathrm{L}}$	=	error in ΔT_{L} associated with *all* repeatability effects.

Combining all expressions gives

$$\Delta T_{\mathrm{L}} = I_{\mathrm{L}} - (I_{\mathrm{S}} - \Delta T_{\mathrm{S}} - \Delta T_{\mathrm{SL}}) - \delta\Delta T_{\mathrm{L}}.$$

Self-heating effects for probes S and L are found to be negligible during the calibration of L within the dry block calibrator.

K.11.3 The digital resolution of the indication from probe L is 0.1 ℃, whereas the resolution for probe S is 0.01 ℃. In each case the resolution is treated as the full-width of a rectangular PDF.

K.11.4 From the calibration certificate for probe S, the following results are found.

I_{S} /℃	ΔT_{S} /℃	$U_{95\%}(\Delta T_{\mathrm{S}})$ /℃
-20.01	0.011	0.024
0.01	0.008	0.024
20.01	0.016	0.024
40.02	0.008	0.024
60.02	0.012	0.024

A linear fit to the calibration data provides the calibration function

$$\Delta T_{\mathrm{S}} = I_{\mathrm{S}} \times 0.00001 + 0.0108 \text{ ℃}$$

K.11.5 The error associated with the use of the linear fit to the calibration data for S is estimated from the maximum residual difference between the data points and the fitted function. This is found to be $0.005\ ℃$, and is treated as the half width of a rectangular PDF.

K.11.6 Calibration drift for S is evaluated in terms of the difference between successive calibrations.

The most recent drift value is $-0.008\ ℃$, and previous values are $0.006\ ℃$ and $-0.002\ ℃$.

Since there is very little information available, the lab chooses to evaluate the uncertainty contribution due to drift in terms of a limit value. The numerical value of the limit is established as the larger of a) the most recent absolute drift value, and b) the average of this and the previous two absolute values. For the current year this corresponds to
$\max\left(0.008, \mathrm{ave}(0.008, 0.006, 0.002)\right)\ ℃ = \max(0.008, 0.00533)℃ = 0.008\ ℃$.
This is treated as the half width of a rectangular PDF.

This algorithm has the advantages of giving weight to the latest drift value if it proves to be higher than average, whilst not giving long-lasting weight to a previous high value (as would be the case if a simple maximum of current and recent values was used).

K.11.7 The laboratory has previously established that the repeatability standard deviation for measurement of ΔT_{L} is $s = 0.00123\ ℃$, based upon a set of $m = 25$ measurements (with $\upsilon = m - 1 = 24$ degrees of freedom).

During calibration of probe L only $n = 1$ measurement is performed at each nominal temperature.

The repeatability uncertainty is therefore estimated to be $u_{\mathrm{rep}} = s/\sqrt{n} = 0.00123\ ℃$.

K.11.8 During characterisation of the temperature block calibrator, the maximum temperature difference between any two different locations has been determined to be $0.048\ ℃$, which is treated as the half-width of a rectangular PDF.

K.11.9 Summary table for evaluation of the measurement uncertainty associated with $\Delta T_{\mathrm{L}} = I_{\mathrm{L}} - (I_{\mathrm{S}} - \Delta T_{\mathrm{S}} - \Delta T_{\mathrm{SL}}) - \delta\Delta T_{\mathrm{L}}$

Quantity	Source of uncertainty	Uncertainty / °C	Probability distribution	Divisor	c_i	$u_i(\Delta T_{\mathrm{L}})$ / °C	υ_i or υ_{eff}
I_{L}	Resolution of indicated value of probe L	0.05	Rectangular	√3	1	0.02887	∞
I_{S}	Resolution of indicated value of probe S	0.005	Rectangular	√3	1	0.00289	∞
ΔT_{S}	Calibration uncertainty for probe S	0.024	Normal	2	1	0.012	∞
ΔT_{S}	Fitting error using calibration data for S	0.005	Rectangular	√3	1	0.00289	∞
ΔT_{S}	Drift since calibration of S	0.0080	Rectangular	√3	1	0.00462	∞
ΔT_{SL}	Gradients, contact effects, installation effects…	0.048	Rectangular	√3	1	0.02771	∞
$\delta\Delta T_{\mathrm{L}}$	Repeatability (characterised by repeatability uncertainty u_{rep})	0.00123	Normal	1	1	0.00123	24
$u_{\mathrm{c}}(\Delta T_{\mathrm{L}})$	Combined standard uncertainty		Normal			0.04225	>500
$U_{95\%}(\Delta T_L)$	Expanded uncertainty		Normal ($k = 2$)			0.08449	>500

K.11.10 During the calibration of probe L, the following results were found:

I_{L} /°C	ΔT_{L} /°C
0.4	0.428
10.4	0.412
20.4	0.432
30.4	0.431
40.5	0.446

A linear fit to the calibration data provides the calibration function

$$\Delta T_{\mathrm{L}} = I_{\mathrm{L}} \times 0.00055 + 0.419\ ℃$$

The error associated with the use of the linear fit to the calibration data for L is estimated from the maximum residual difference between the data points and the calibration function. This is found to be 0.012 ℃, and is treated as the half width of a rectangular PDF.

NOTE: The expanded uncertainty for $U_{95\%}(\Delta T_{\mathrm{L}})$ established in K10.9 is 0.084 ℃ when expressed to two significant figures, as is appropriate for a coverage interval of 95 %. Consequently ΔT_{L} is also expressed to the same decimal precision (3 d.p.), even though the probe indication is resolution-limited to fewer decimal places. The key point is that the measurement uncertainty is evaluated from *all* of the information available, *not just* the resolution of the probe indication.

K.11.11 In later use (post calibration),

The temperature measured by lab probe L is

$$T_{\mathrm{u}} = I_{\mathrm{u}} - \Delta T_{\mathrm{u}}$$

where

T_{u} = measured temperate according to the probe L,

I_{u} = temperature indicated by the lab standard L when in use,

ΔT_{u} = is the measurement error corresponding to I_{u}.
In use, it is assumed that the functional relationship between ΔT_{u} and I_{u} is that which was earlier established by the calibration data ΔT_{L} and I_{L}, i.e., $\Delta T_{\mathrm{u}}(I) = \Delta T_{\mathrm{L}}(I)$, so that

$$\Delta T_{\mathrm{u}} = I_{\mathrm{u}} \times 0.00055 + 0.419\ ℃$$

The temperature assigned to the gas when probe L temperature is T_u, is

$$T_{\mathrm{g}} = T_{\mathrm{u}} - \Delta T_{\mathrm{ug}} - \delta T_{\mathrm{h}} - \delta T_{\mathrm{r}}$$

where

T_{g} = temperature of the gas,

ΔT_{ug} = temperature difference between the probe and the gas,

δT_{h} = error in T_{g} associated with self-heating of probe L when attached to the outside of the vessel containing the gas.

δT_{r} = error in T_{g} associated with repeatability effects.

Combining expressions gives

$$T_{\mathrm{g}} = (I_{\mathrm{u}} - \Delta T_{\mathrm{u}}) - \Delta T_{\mathrm{ug}} - \delta T_{\mathrm{h}} - \delta T_{\mathrm{r}}$$

K.11.12 The digital resolution of the indication from probe L is 0.1 ℃, which is treated as the full-width of a rectangular PDF.

K.11.13 Calibration drift for L is evaluated in terms of the largest difference between successive calibrations.

The most recent drift value is $0.003\ ^\circ\mathrm{C}$, and previous values are $-0.007\ ^\circ\mathrm{C}$ and $-0.004\ ^\circ\mathrm{C}$.

Since there is very little information available, the lab chooses to evaluate the uncertainty contribution due to drift in terms of a limit value. The numerical value of the limit is established as the larger of a) the most recent absolute drift value, and b) the average of this and the previous two absolute values. For the current year this corresponds to
$\max\left(0.003, \mathrm{ave}(0.003, 0.007, 0.004)\right)\ ^\circ\mathrm{C} = \max(0.003, 0.00467)^\circ\mathrm{C} = 0.00467\ ^\circ\mathrm{C}$.
This is treated as the half width of a rectangular PDF.

K.11.14 The mean temperature of the gas inside the vessel is believed to be closely related to the mean temperature of the vessel. In turn, the mean temperature of the vessel has previously been found to be within $\pm 0.1\ ^\circ\mathrm{C}$ of the temperature of probe L.
The best estimate of the measurement error is therefore assumed to be $\Delta T_{\mathrm{ug}} = 0\ ^\circ\mathrm{C}$ with an uncertainty defined by a rectangular PDF with limits of $\pm 0.1\ ^\circ\mathrm{C}$.

K.11.15 The temperature stability of the vessel has previously been evaluated and the repeatability standard deviation was found to be $s = 0.076\ ^\circ\mathrm{C}$, based upon a set of $m = 50$ measurements with $\upsilon = m - 1 = 49$ degrees of freedom.

During use the average of $n = 4$ measurements is obtained.

The repeatability uncertainty is therefore estimated to be $u_{\mathrm{rep}} = s/\sqrt{n} = 0.076\ ^\circ\mathrm{C}/\sqrt{4} = 0.038\ ^\circ\mathrm{C}$.

K.11.16 Probe L is attached to the outside of the vessel in such a way that errors due to self-heating effects are no greater than 4 mK over the range of operating conditions. This value is treated as the half width of a rectangular PDF.

K.11.17 Summary table for evaluation of the measurement uncertainty associated with
$T_{\mathrm{g}} = (I_{\mathrm{u}} - \Delta T_{\mathrm{u}}) - \Delta T_{\mathrm{ug}} - \delta T_{\mathrm{h}} - \delta T_{\mathrm{r}}$

Quantity	Source of uncertainty	Uncertainty / °C	Probability distribution	Divisor	c_i	$u_i(T_{\mathrm{g}})$ / °C	υ_i or υ_{eff}
I_{u}	Resolution of indicated value of probe L	0.05	Rectangular	√3	1	0.02887	∞
ΔT_{u}	Calibration uncertainty for probe L	0.08449	Normal	2	1	0.04225	∞
ΔT_{u}	Fitting error when using calibration data for probe L	0.012	Rectangular	√3	1	0.00693	∞
ΔT_{u}	Drift since calibration of probe L	0.0047	Rectangular	√3	1	0.00269	∞
ΔT_{ug}	Gradients, contact effects, installation effects…	0.1	Rectangular	√3	1	0.05774	∞
δT_{h}	Self-heating of probe L	0.004	Rectangular	√3	1	0.00231	∞
δT_{r}	Repeatability errors (characterised by repeatability uncertainty u_{rep} for T_{g})	0.038	Normal	1	1	0.038	49
$u_{\mathrm{c}}(T_{\mathrm{g}})$	Combined standard uncertainty		Normal			0.08638	>500
$U_{95\%}(T_g)$	Expanded uncertainty		Normal ($k = 2$)			0.17276	>500

K.11.18 Reported result

For the measurement of interest, the average probe indication is $I_u = 23.4$ ℃, for which it is calculated that $\Delta T_\mathrm{u} = 0.431$ ℃

The gas temperature T_g (corresponding to average probe indication $I_u = 23.4$ ℃),
is therefore $\mathbf{22.97}$ **℃** $\pm$ $\mathbf{0.17}$ **℃.**

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

NOTE: Unlike in most of the other examples in this Appendix, in this example each individual source of uncertainty is not represented by a unique term in the measurement models.
This approach is commonplace but is arguably not as good practice as it is to identify each source with an individual quantity in the model.

Appendix L Expression of uncertainty for a range of values

L.1 Introduction

L.1.1 The GUM [1] and its supplements deal with the evaluation of measurement uncertainty for a single measurement 'point'. But in practice measurements may be performed for several points in a range, as is common when calibrating measuring instruments, and the use of a mathematical expression to describe the uncertainty at any of these points can be desirable.

L.1.2 This Appendix explains how the principles of this guide can be applied in such a situation, and provides an illustration of the process using a worked example.

L.2 Principles

L.2.1 When measurements are made over a range and the corresponding sources of uncertainty are examined it may be found that some are absolute in nature (i.e., they are independent of the measurand) and some are 'relative' in nature (i.e., they are proportional to the value of the corresponding input quantity or to the measurand).

L.2.2 It is possible to follow this guide and to individually calculate the measurement uncertainty at each tabulated point over the range.
This is often the least ambiguous approach and arguably provides the clearest form of reporting. It may be the only option available in situations where the coverage factor varies for different points in the reported range.

L.2.3 However, for many users there are circumstances where it is useful to be able to separate the absolute and relative constituents of the measurement uncertainty. For example, so that these can be imported into a subsequent uncertainty evaluation as separate inputs, or to avoid reporting measurement uncertainty on a point-by-point basis in a results table.

L.2.4 The process for evaluating the measurement uncertainty describing a range of values is identical to that for single values except that the absolute and relative input quantities are separated, and in effect, an uncertainty evaluation is carried out for each type in the manner already described in this guide.

L.2.5 Historically, the measurement uncertainty has sometimes (incorrectly) then been expressed as a 'linear' combination of the relative and absolute components
$U = U_{\mathrm{rel}} + U_{\mathrm{abs}}$
but this linear addition of quantities is not in accordance with the principles embodied in the GUM (unless there happens to be full correlation between the absolute and relative terms - which is not usually the case).

Instead, the measurement uncertainty should be expressed as a quadrature sum (root of the sum of the squares of the components),as demonstrated in the following example.

L.3 Example of uncertainty evaluation for a range of values

L.3.1 In this example a 6½ - digit electronic multimeter (DMM) is calibrated on its 1 V dc range using a multi-function calibrator.

L.3.2 The calibrations were carried out in both polarities at 0.1 V increments from zero to 1 V. Only one measurement was carried out at each point and therefore reliance was placed on a previous evaluation of repeatability using similar multimeters.

L.3.3 No corrections were made for known errors of the calibrator as these were identified as being small relative to other sources of uncertainty. The uncorrected errors are assumed to be zero with an uncertainty obtained by analysis of information obtained from the calibration certificate for the calibrator.

L.3.4 The measurement error ΔI_{DMM} associated with indication I_{DMM} of the multimeter under calibration, can be described as follows:

$$\Delta I_{\mathrm{DMM}} = I_{\mathrm{DMM}} - V_{\mathrm{REF}} + \delta I_{\mathrm{RES}} + \delta\Delta_{\mathrm{r}}$$

with

$$V_{\mathrm{REF}} = V_{\mathrm{CAL}} + \delta V_{\mathrm{D}} + \delta V_{\mathrm{UE}} + \delta V_{\mathrm{TC}} + \delta V_{\mathrm{LIN}} + \delta V_{\mathrm{T}} + \delta V_{\mathrm{CM}}$$

i.e.,

$$\Delta I_{\mathrm{DMM}} = I_{\mathrm{DMM}} - \{V_{\mathrm{CAL}} + \delta V_{\mathrm{D}} + \delta V_{\mathrm{UE}} + \delta V_{\mathrm{TC}} + \delta V_{\mathrm{LIN}} + \delta V_{\mathrm{T}} + \delta V_{\mathrm{CM}}\} + \delta I_{\mathrm{RES}} + \delta\Delta_{\mathrm{r}}$$

where

ΔI_{DMM}	=	error in the indication of the DMM,
I_{DMM}	=	indication of the DMM,
δI_{RES}	=	rounding error due to the resolution of DMM being calibrated,
$\delta\Delta_r$	=	repeatability error of indication when DMM is measuring V_{REF},
V_{REF}	=	reference voltage presented to the DMM,
V_{CAL}	=	calibrated voltage setting of multifunction calibrator,
δV_D	=	error due to drift in voltage of multifunction calibrator since last calibration,
δV_{UE}	=	error due to uncorrected errors of multifunction calibrator,
δV_{TC}	=	error due to temperature coefficient of multifunction calibrator,
δV_{LIN}	=	error due to linearity and zero offset of multifunction calibrator,
δV_T	=	error due to thermoelectric voltages generated at junctions of connecting leads calibrator and multimeter,
δV_{CM}	=	effects due to imperfect common-mode rejection characteristics of the measurement system.

L.3.5 The calibration uncertainty was obtained from the certificate for the multi-function calibrator. This had a value of 2.8 ppm as a relative uncertainty and a further 0.5 μV in absolute units ($k = 2$).

L.3.6 The manufacturer's 1-year performance specification for the calibrator includes information relating to the following effects: V_{D}, V_{UE}, and V_{TC}, that are found to be relative in nature; and V_{LIN}, that is found to be absolute in nature.

The specification for the calibrator on the 1 V dc range is ±8 ppm of reading and ±1 ppm of full-scale. It is assumed that these are independent quantities.
For this multifunction calibrator, the full-scale value is twice the range value; therefore, the absolute term is ±2 × 1 ×10^{-6} V = ±2 μV.
The performance of the calibrator has been verified by examining its calibration data and history, using internal quality control checks, and ensuring that it was used within the temperature range and other conditions as specified by the manufacturer. A rectangular distribution is assumed.

L.3.7 The effects of thermoelectric voltages for the connecting leads used have been evaluated on a previous occasion. Thermoelectric voltages are independent of the voltage setting are therefore an absolute uncertainty contribution. An uncertainty of 1 μV is assigned, based on previous experiments with the leads. The probability distribution is assumed to be rectangular.

L.3.8 Effects due to common-mode signals have also been the subject of a previous evaluation and an uncertainty of 1 μV, with a rectangular distribution, is assigned. This contribution is absolute in nature, as the common-mode voltage is unrelated to the measured voltage.

L.3.9 No correction can be made for the rounding due to the resolution of the digital display of the multimeter. The least significant digit on the range being calibrated corresponds to 1 μV and there is therefore a possible rounding error between ±0.5 μV. The probability distribution is assumed to be rectangular, and this term is absolute in nature.

L.3.10 The repeatability standard deviation of the calibration process had been established from previous measurements using a similar multimeter. This Type A evaluation, based upon $m = 10$ measurements at each value, was performed at 0 V and at 1.0 V (as well as other nearby values). Repeatability at the zero point was found to be insignificant compared with other absolute contributions, whereas for measurements at 1.0 V, the repeatability standard deviation $s(\delta\Delta_r)$ was found to be 2.5 ppm, (consistent with previous evaluations at other voltage levels) with $\upsilon = (m - 1) = 9$ degrees of freedom.

This calibration of the unknown DMM was established from a single measurement. However, assuming that the previous evaluation of repeatability using a similar multimeter is representative of this later measurement, the repeatability uncertainty can be obtained using the previously obtained repeatability standard deviation, and $n = 1$ (as only one reading is made for the actual calibration).

So, from equation (4), the repeatability uncertainty is

$$u_{\text{rep}}(\delta\Delta_r) = \frac{s(\delta\Delta_r)}{\sqrt{n}} = \frac{2.5}{\sqrt{1}} = 2.5 \text{ ppm}$$

L.3.11 All uncertainties associated with the indicated value I_{DMM} are accounted for by other terms, so no further contributions need be considered.

L.3.12 Summary tables for
$\Delta I_{\mathrm{DMM}} = I_{\mathrm{DMM}} - \{V_{\mathrm{CAL}} + \delta V_{\mathrm{D}} + \delta V_{\mathrm{UE}} + \delta V_{\mathrm{TC}} + \delta V_{\mathrm{LIN}} + \delta V_{\mathrm{T}} + \delta V_{\mathrm{CM}}\} + \delta I_{\mathrm{RES}} + \delta\Delta_r$

Quantity	Source of relative uncertainty	Uncertainty / ppm	Probability distribution	Divisor	c_i	$\frac{u_i(\Delta I_{\mathrm{DMM}})}{I_{\mathrm{DMM}}}$ / ppm	v_i or v_{eff}
V_{CAL}	Calibration uncertainty	2.8	Normal	2.0	1	1.40	∞
δV_{SPEC}	Specification of calibrator	8.0	Rectangular	√3	1	4.62	∞
$\delta\Delta_r$	Repeatability	2.5	Normal	1.0	1	2.5	9
$u_c(\Delta I_{\mathrm{DMM}})/I_{\mathrm{DMM}}$	Standard uncertainty		Normal			5.44	>200
$U_{95\%}(\Delta I_{DMM})/I_{DMM}$	Expanded uncertainty		Normal ($k = 2$)			10.9	>200

Symbol	Source of absolute uncertainty	Uncertainty / μV	Probability distribution	Divisor	c_i	$u_i(\Delta I_{\mathrm{DMM}})$ / μV	v_i or v_{eff}
V_{CAL}	Calibration uncertainty	0.5	Normal	2.0	1	0.25	∞
δV_{SPEC}	Specification of calibrator	2.0	Rectangular	√3	1	1.15	∞
δV_{T}	Thermoelectric voltages	1.0	Rectangular	√3	1	0.58	∞
δV_{CM}	Common mode effects	1.0	Rectangular	√3	1	0.58	∞
δI_{RES}	Voltmeter resolution	0.5	Rectangular	√3	1	0.29	∞
$u_c(\Delta I_{\mathrm{DMM}})$	Standard uncertainty		Normal			1.46	>1000
$U_{95\%}(\Delta I_{DMM})$	Expanded uncertainty		Normal ($k = 2$)			2.9	>1000

L.3.13 Relative uncertainty for all input and output quantities should expressed relative to the same quantity. In this example the chosen quantity is I_{DMM}, the indicated voltage and it has been assumed that ΔI_{DMM} is always small in comparison with I_{DMM} and V_{CAL} so that the input uncertainty $u(x_i)/I_{\mathrm{DMM}} \approx u(x_i)/V_{\mathrm{CAL}}$
(See L.4 for further discussion of how to treat relative input quantities.)

L.3.14 The expanded uncertainty for the above measurement of ΔI_{DMM} is

$$U_{95\%}(\Delta I_{DMM}) = \sqrt{(I_{\mathrm{DMM}} \times 11\ \mathrm{ppm})^2 + (2.9\ \mu\mathrm{V})^2}$$

So, for example, if $I_{\mathrm{DMM}} = 500\ \mathrm{mV}$ we find that

$$U_{95\%}(\Delta I_{DMM}) = \sqrt{(500\ \mathrm{mV} \times 11\ \mathrm{ppm})^2 + (2.9\ \mu\mathrm{V})^2} = \sqrt{(5.5\ \mu\mathrm{V})^2 + (2.9\ \mu\mathrm{V})^2} = 6.2\ \mu\mathrm{V}$$

L.3.15 Alternatively, this expression can be written in terms of a defined function, e.g.,

$$U_{95\%}(\Delta I_{DMM}) = Q[I_{\mathrm{DMM}} \times 11\ \mathrm{ppm}, 2.9\ \mu\mathrm{V}]$$

where
$Q[a, b] = [a^2 + a^b]^{1/2}$

In practice, the variable (I_{DMM} in this case) is often omitted and expressions such as
$U = \sqrt{(11\ \mathrm{ppm})^2 + (2.9\ \mu\mathrm{V})^2}$, or
$U = Q[11\ \mathrm{ppm}, 2.9\ \mu\mathrm{V}]$, might be encountered.

Although this simplification doesn't usually create any confusion, it should be recognised that these expressions are not mathematically complete or dimensionally consistent (in this example the relative term is 'dimensionless' whereas the absolute term has dimensions corresponding to voltage).
The simplification should be avoided when there is any possibility of ambiguity in its use.

L.3.16 A non-mathematical representation can also be used to describe the measurement uncertainty. For example:

> The expanded uncertainty for the measurement of ΔI_{DMM} is found by taking the square root of the sum of the squares of two values corresponding to:
> a) Relative uncertainty of 11 ppm of the DMM indication, and
> b) Absolute uncertainty of 2.9 μV

L.3.17 In each case the result should be accompanied by the usual statement:

> The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a coverage probability of approximately 95 %. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

L.3.18 Uncertainties reported in this form can be imported into subsequent evaluations as single enumerated values i.e., as a single input value computed from the expression.
Alternatively, they can be imported as two separate (independent) inputs corresponding to the relative and the absolute quantity.

L.3.19 Tests for dominance of a single term in each of the two budgets indicate that the relative component of the 'specification of calibrator' quantity may be dominant.
Applying the test described in Appendix C to the relative terms we find that $\frac{4.62}{\sqrt{1.4^2+2.5^2}} = 1.61$ which is greater than the usually adopted critical value, 1.42

To discover if this term is indeed dominant for the overall uncertainty, the balance of the contributions to the combined uncertainty (i.e., the absolute terms) also need to be considered. This can only be achieved by evaluating the combined uncertainty for specific results in a single evaluation, as shown below.

Uncertainty budget evaluated at $\Delta I_{\mathrm{DMM}} = 0.95\ \mathrm{V}$

Quantity	Source of uncertainty	Uncertainty	Probability distribution	Divisor	$u(x_i)$	c_i	$u_i(\Delta I_{\mathrm{DMM}})$ / $\mu\mathrm{V}$	ν_i or ν_{eff}
V_{CAL}	Calibration uncertainty	$2.8\ \mathrm{ppm}$	Normal	2.0	$1.33\ \mu\mathrm{V}$	1	$1.33\ \mu\mathrm{V}$	∞
δV_{SPEC}	Specification of calibrator	$8.0\ \mathrm{ppm}$	Rectangular	√3	$4.38\ \mu\mathrm{V}$	1	$4.39\ \mu\mathrm{V}$	∞
$\delta\Delta_r$	Repeatability	$2.5\ \mathrm{ppm}$	Normal	1.0	$2.37\ \mu\mathrm{V}$	1	$2.38\ \mu\mathrm{V}$	9
V_{CAL}	Calibration uncertainty	$0.5\ \mu\mathrm{V}$	Normal	2.0	$0.25\ \mu\mathrm{V}$	1	$0.25\ \mu\mathrm{V}$	∞
δV_{SPEC}	Specification of calibrator	$2.0\ \mu\mathrm{V}$	Rectangular	√3	$1.15\ \mu\mathrm{V}$	1	$1.15\ \mu\mathrm{V}$	∞
δV_{T}	Thermoelectric voltages	$1.0\ \mu\mathrm{V}$	Rectangular	√3	$0.58\ \mu\mathrm{V}$	1	$0.58\ \mu\mathrm{V}$	∞
δV_{CM}	Common mode effects	$1.0\ \mu\mathrm{V}$	Rectangular	√3	$0.58\ \mu\mathrm{V}$	1	$0.58\ \mu\mathrm{V}$	∞
δI_{RES}	Voltmeter resolution	$0.5\ \mu\mathrm{V}$	Rectangular	√3	$0.29\ \mu\mathrm{V}$	1	$0.29\ \mu\mathrm{V}$	∞
$u_c(\Delta I_{\mathrm{DMM}})$	Standard uncertainty		Normal				$5.37\ \mu\mathrm{V}$	>500
$U_{95\%}(\Delta I_{\mathrm{DMM}})$	Expanded uncertainty		Normal $(k = 2)$				$10.73\ \mu\mathrm{V}$	>500

L.3.20 In this case it is found that the specification uncertainty dominates for measured values $|I_{\mathrm{DMM}}|$ greater than about $0.95\ \mathrm{V}$. (applying the test described in Appendix C we find that $\frac{4.39}{3.09} = 1.42$)
In the worst case, for $|I_{\mathrm{DMM}}| = 1$, the value of the test ratio is 1.435 corresponding to a coverage factor of $k = 1.898$ for 95.45 % coverage probability
The value is not significantly different to the value of $k = 1.900$ corresponding to the adopted critical value.

Appendix M Assessment of conformity with specification

M.1.1 In test reports and to some extent in calibration certificates, there will be occasions where it becomes necessary to make a statement about whether the measured result indicates conformity of the measurand with some form of specification.

M.1.2 Guidance on performing and reporting assessment of conformity with specifications is given in UKAS document LAB 48: 'Decision Rules and Statements of Conformity' [14] and the references therein.

Appendix N Uncertainties for test results

N.1 Introduction

N.1.1 ISO/IEC 17025 [5] requires that testing laboratories shall have and apply procedures for estimating measurement uncertainty.

N.1.2 Testing laboratories should therefore have a defined policy covering the evaluation and reporting of the uncertainties associated with the tests performed. The laboratory should use documented procedures for the evaluation, treatment, and reporting of the uncertainty.

N.1.3 The methodology for estimation of uncertainty in testing is no different from that in calibration and therefore the procedures described in this document apply equally to testing results.

N.1.4 It is however recognised that in certain areas of testing it may be known that a significant contribution to uncertainty exists but that the nature of the test precludes a rigorous evaluation of this contribution. In such cases, ISO/IEC 17025 requires that a reasonable estimation be made and that the form of the reporting does not give an incorrect impression of the uncertainty.

NOTES:

1. In some fields of testing, it may be the case that the contribution of measuring instruments to the overall uncertainty can be demonstrated to be insignificant when compared with the repeatability of the process. Nevertheless, such instruments have to be shown to comply with the relevant specifications, normally by calibration.

2. It will sometimes be the case that the procedure requires standard reference materials to be subject to the same process, the result being the difference between the readings for the analyte and the reference material. In such cases, much of the process can be considered to be negatively correlated and the measurement uncertainty can be evaluated from the resolution and repeatability of the process; matrix effects may also have to be considered.

N.1.5 Many tests involve some form of examination (or inspection) where the outcome of the test is a nominal property (e.g., colour, shape, species, sequence of markers…). In the case of these ‘qualitative’ tests the concept of measurement uncertainty does not readily apply.

But that is *not* to say that measurement uncertainty doesn't play a role in such tests… in fact, in most cases, such tests are performed under defined conditions that are themselves subject to measurement.

<u>For example</u>:
A test requires an inspector to examine the colour of a fluid sample after preparation according to a defined procedure and processing in an oven at (40 ± 1) °C for between 60 min and 65 min.

For this test, the examination of the colour involves a subjective judgement from a trained and competent examiner whose reliability can be established by proficiency testing. However, the oven temperature and elapsed time are both measurable quantities for which a value and a measurement uncertainty can be established.

N.2 Common sources of uncertainty in testing

N.2.1 The following general issues will apply to many areas of testing:

(a) Incomplete definition of the test - the requirement may not be clearly described, e.g., the temperature of a test may be given as 'room temperature'.

(b) Imperfect realisation of the test procedure; even when the test conditions are clearly defined it may not be possible to produce the theoretical conditions in practice due to unavoidable imperfections in the materials or systems used.

(c) Sampling - the sample may not be fully representative. In some disciplines, such as microbiological testing, it can be very difficult to obtain a representative sample.

(d) Inadequate knowledge of the effects of environmental conditions on the measurement process, or imperfect measurement of environmental conditions.

(e) Personal bias and human factors; for example:

- Reading of scales on analogue indicating instruments.

- Judgement of colour.

- Reaction time, e.g., when using a stopwatch.

- Instrument resolution or discrimination threshold, or errors in graduation of a scale.

(f) Values assigned to measuring equipment and reference materials.

(g) Changes in the characteristics or performance of measuring equipment or reference materials since the last calibration.

(h) Values of constants and other parameters used in data evaluation.

(i) Approximations and assumptions incorporated in the measurement method and procedure.

(j) Variations in repeated observations made under similar but not identical conditions - such random effects may be caused by, for example, electrical noise in measuring instruments, short-term fluctuations in the local environment, e.g. temperature, humidity and air pressure, variability in the performance of the person carrying out the test and variability in the homogeneity of the sample itself.

These sources are not necessarily independent.

In addition, unrecognised systematic effects may exist that are not considered, but contribute to error. This is one reason that participation in inter-laboratory comparisons, proficiency testing schemes and internal cross-checking of results by different means are encouraged.

N.2.2 Information on some of the sources of these errors can be obtained from:

(a) Data in calibration certificates - this enables corrections to be made and uncertainties to be assigned.

(b) Previous measurement data - for example, history graphs can be constructed and can yield useful information about changes with time.

(c) Experience with or general knowledge about the behaviour and properties of similar materials and equipment.

(d) Accepted values of constants associated with materials and quantities.

(e) Manufacturers' specifications.

(f) All other relevant information.

These are all referred to as Type B evaluations because the values were not obtained by statistical means. However, the influence of random effects is often evaluated by the use of statistics; if this is the case then the evaluation is designated Type A.

N.3 The 'top-down' approach

N.3.1 The organisation of the uncertainty budget around a measurement equation as described in this guidance document is commonly described as a 'bottom-up' approach. However, in some situations, it is not feasible to establish a detailed measurement equation and a different 'top-down' approach is taken.

N.3.2 Guidance documents such as, EURACHEM/CITAC Guide CG 4 "Quantifying Uncertainty in Analytical Measurement" [10] and ISO 21748 "Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty evaluation" [8] describe a process for evaluating measurement uncertainty using data obtained from collaborative studies conducted in accordance with ISO 5725-2 "Accuracy (trueness and precision) of measurement methods and results — Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method" [7]

N.3.3 Measurement methods that have been evaluated by this 'collaborative study' approach are characterised in terms of 'precision' i.e., by a 'repeatability standard deviation' and a 'reproducibility standard deviation'. In some cases, a 'method bias' and its associated uncertainty will also be established.

N.3.4 In its simplest form, a laboratory that demonstrates control of precision and bias, and introduces no additional factors (from operations not conducted during the collaborative study) may use the reproducibility standard deviation for estimating the standard uncertainty.

Appendix P Comparing independent results using E_n ratio

P.1 The equivalence of two independent measurement results is often compared by computing the Normalised Error ratio (E_N ratio), where

$$E_N = \frac{L_1 - L_2}{\sqrt{(U_1)^2 + (U_2)^2}}$$

L_1 = the value of measurement result 1
L_2 = the value of measurement result 2
U_1 = the expanded uncertainty associated with measurement 1
U_2 = the expanded uncertainty associated with measurement 2

It is common to perform the test with uncertainties expanded for 95 % coverage probability. In this case, if the results are equivalent, the E_N ratio should be within the range ±1, on approximately 95 % of occasions.

If the analysis reveals that E_N lies outside this range more frequently than expected, some investigation and corrective action will be required.

NOTES:

1 The commonly used formation of the E_N ratio described above, involving the direct combination of *expanded* uncertainties, is not consistent with the GUM [1]
Instead, it should be evaluated from

$$E_N = \frac{L_1 - L_2}{k_p\sqrt{(u_1)^2 + (u_2)^2}}$$

where u_1 and u_2 are the standard uncertainties associated with L_1 and L_2, and k_p is the coverage factor established for the PDF of the quantity $(L_1 - L_2)$.
In situations where the PDFs for both L_1 and L_2 have normal distributions this gives identical values to the commonly used equation.

2 It is not physically realistic for comparisons to produce consistently small E_N values. This usually indicates that U_1 and/or U_2 are significant overestimates. This can arise from failure to identify correlation between L_1 and L_2 ,or simply from the (poor) practice of basing uncertainty evaluations on ‘safe’ or ‘conservative’ estimates rather than on realistic estimates for input quantities (as required by the GUM).

3 The E_N ratio (as defined above) is not usually appropriate for testing the equivalence of operators or processes in the same laboratory, as the measurement results are not likely to be independent. To use it in such cases an approximation can be obtained by omitting all common inputs (such as those relating to temperature effects and calibration traceability) from each uncertainty evaluation, so that only independent quantities, such as errors of resolution and repeatability remain.

Appendix Q Relative input quantities

Q.1.1 To allow standardised input uncertainties to be correctly combined using Equation (1) they must be 'alike'.

Q.1.2 For input quantities where the information is expressed in absolute terms (value and unit), this is ensured by the process of standardising using the PDF divisor, and scaling using the sensitivity coefficient.

Q.1.3 For relative quantities some additional scaling or reformulation of the input uncertainty information may be required.

In general terms, if uncertainty information about an input quantity x_i is obtained as a relative value $w(x_i)$ expressing the uncertainty relative to some quantity q_i but the combined standard uncertainty is to be expressed relative to a different quantity z, then the input information must be transformed by an additional scaling factor of q_i/z.

This factor is sometimes confused with the sensitivity coefficient, and it is sometimes (incorrectly) represented in the tabular representation of an uncertainty budget under that label. However, it is important to be aware that this is not the sensitivity coefficient as defined by GUM equation 11b as used throughout this guide, which must still be applied in addition to any additional scaling factor.

The factor could be incorporated into a suitably defined, modified sensitivity factor, say

$$c'_i = c_i \frac{q_i}{z}$$

However, to avoid ambiguity it is better practice to apply the conversion *before* incorporating the information into the standard evaluation process (and associated summary table), replacing $w(x_i)$ as input information with $w(x_i) \times \frac{q_i}{z}$

For example, suppose that a calibration measurement aims to establish a measurement error ΔV for a voltmeter. Although the measurand is ΔV, suppose also that the combined standard uncertainty is to be expressed as a proportion of the measured voltage, V_m (which might for example correspond to a meter indication I_DMM, where $V_\mathrm{m} = I_\mathrm{DMM} - \Delta V$)

Finally, suppose that one of the input quantities is a resistance R. The given uncertainty information for R is a relative quantity, $w(R) = 50\ \mathrm{ppm}$, and it is understood that $w(R)$ is implicitly defined as

$$w(R) = \frac{u(R)}{R}$$

where $u(R)$ is the standard uncertainty (the same argument applies for expanded uncertainties).

For the standard process to be followed (in which the PDF and corresponding divisor are established, and the sensitivity coefficient $c_R = \frac{\partial(\Delta V)}{\partial R}$ is applied), the information about the input quantity must first be scaled so that it expresses the information in terms relative to V_m, i.e., in place of $w(R)$ an input quantity corresponding to $w(R) \times \frac{R}{V_\mathrm{m}}$ should be used.

NOTE: Unless there is a reason why the combined standard uncertainty needs to be expressed as a relative quantity, the most straightforward approach to the uncertainty evaluation might be to calculate $u(R) = u(R) \times R$, then to proceed with the evaluation in terms of absolute quantities.

Appendix R Symbols

The symbols used are taken mainly from the GUM [1]. The meanings have also been described further in the text, usually where they first occur, but are summarised here for convenience of reference.

f	Functional relationship between the measurand Y and the input quantities X_i on which Y depends $Y = f(X_i)$, and between output estimate y and input estimates x_i, thus, $y = f(x_i)$.
$c_i = \frac{\partial f}{\partial x_i}$	Sensitivity coefficient i th input quantity. Partial derivative of the functional relationship f, taken with respect to input quantity X_i, evaluated at $X_i = x_i$.
a_i	Estimated semi-range of a probability distribution for an input quantity x_i with width $2a_i$.
$a_\mathrm{R} = c_\mathrm{R} a$	Semi-width of a rectangular input distribution, corresponding to the semi-width a of the input distribution scaled by the corresponding sensitivity factor c_R.
k	Coverage factor (general).
k_p	Coverage factor used to calculate the expanded uncertainty U_p for a defined coverage probability p., e.g., $k_{95\%}$.
m	Number of readings or observations that are used for the evaluation of $s(q_j)$.
n	Number of readings or observations that contribute to a mean value.
N	Number of input quantities x_i on which the value of the measurand depends, $i = 1$ to N.
q_j	jth observation of randomly varying quantity q. *See NOTE.*
$\bar{q}$	Arithmetic mean or average of n repeated observations of randomly varying quantity q.
p	Coverage probability or 'level of confidence' $(0 \leq p \leq 1)$. Often expressed in percentage terms $(0\ \% \leq p \leq 100\ \%)$.
σ	The standard deviation that characterises the variation in a measurement process (The limit value of s that would be found were it possible to repeat a measurement a very large number of times).
s	Best available estimate of σ. Referred to here as the repeatability standard deviation.
$u_\mathrm{rep} = \frac{s}{\sqrt{n}}$	Repeatability uncertainty.
$s(q_j)$	Value of s obtained from a sample of m observations $\{q_1, \dots, q_m\}$. *See NOTE.*
$t_p(\nu_\mathrm{eff})$	Student *t*-factor for ν_eff degrees of freedom corresponding to a given coverage probability p.
$u(x_i)$	Standard uncertainty of input estimate x_i.
$u_i(y)$	Standard uncertainty of input estimate x_i.expressed in terms of the measurand,

$u_i(y) = |c_i|u(x_i)$

$u_c(y)$ Combined standard uncertainty of output estimate y.

U_p, U Expanded uncertainty of output estimate y that defines an interval $y \pm U$, with a defined coverage probability, p of containing the measurand Y.

ν_i Degrees of freedom of standard uncertainty $u(x_i)$ of input estimate x_i.

ν_{eff} Effective degrees of freedom of $u_c(y)$ used to obtain $t_p(\nu_{\text{eff}})$.

δx The prefix δ is used to demote a quantity whose most likely value is zero, although there is an associated (non-zero) uncertainty.

Δx The symbol Δ is used to represent a difference, a measurement error or a correction for which a non-zero value is usually known.

N, R, T, U Labels, respectively used to designate, normal, rectangular, triangular and U-shaped probability distributions.

$\mathrm{N, R, T, U}$ Subscripts, respectively used to designate, normal, rectangular, triangular and U-shaped probability distributions.

NOTE The GUM uses the symbols q_k and $s(q_k)$ whereas q_j and $s(q_j)$ are used here. M3003 uses the subscript j instead of k to avoid any possible confusion with the coverage factor k.

Appendix S References

[1] JCGM 100:2008,
Evaluation of measurement data - Guide to the Expression of Uncertainty in Measurement. Known as 'The GUM'.

[2] JCGM 101:2008,
Evaluation of measurement data - Supplement 1 to the "Guide to the expression of uncertainty in measurement" - Propagation of distributions using a Monte Carlo method.

[3] JCGM GUM-6:2020,
Guide to the expression of uncertainty in measurement — Part 6: Developing and using measurement models.

[4] JCGM 200:2008,
International vocabulary of metrology – Basic and general concepts and associated terms. Known as 'The VIM'.

[5] BS EN ISO/IEC 17025,
General requirements for the competence of testing and calibration laboratories.

[6] BS EN ISO 15189,
Medical laboratories - Requirements for quality and competence.

[7] ISO 5725-2:2019,
Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method.

[8] ISO 21748:2017,
Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty evaluation.

[9] ISO/TS 20914:2019,
Medical laboratories - Practical guidance for the estimation of measurement uncertainty.

[10] EURACHEM/CITAC Guide CG 4
Quantifying Uncertainty in Analytical Measurement,
QUAM:2012. P1, 3rd Ed, 2012. ISBN 978-0-948926-30-3.

[11] ILAC-P14:2020,
ILAC Policy for Uncertainty in Calibration.

[12] EURAMET Calibration Guide No. 18,
Guidelines on the Calibration of Non-Automatic Weighing Instruments

[13] UKAS LAB 14,
Calibration of Weighing Machines.

[14] UKAS LAB 48,
Decision Rules and Statements of Conformity.

[15] UKAS LAB 45,
Schedules of Accreditation for Calibration Laboratories.

[16] DIETRICH, C.F.
Uncertainty, Calibration and Probability: The Statistics of Scientific and Industrial Measurement Uncertainty,
Institute of Physics, January 1991,
ISBN 0-7503-0060-4.

[17] HARRIS, I.A and WARNER, F.L.
Re-examination of mismatch uncertainty when measuring microwave power and attenuation.
IEE Proceedings, Vol 128 Pt H No 1, February 1981.