M3003

EDITION 3 | NOVEMBER 2012

The Expression of Uncertainty and Confidence in Measurement

CONTENTS

United Kingdom Accreditation Service, 21-47 High Street, Feltham, Middlesex, TW13 4UN
Website: www.ukas.com Publication requests Tel: +44 (0) 20 8917 8400 Fax: +44 (0) 20 8917 8500

1 INTRODUCTION

1.1 The general requirements that testing and calibration laboratories have to 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 17025:2005. This international standard forms the basis for international laboratory accreditation and in cases of differences in interpretation remains the authoritative document at all times. M3003 is not intended as a prescriptive document, and does not set out to introduce additional requirements to those in ISO/IEC 17025:2005 but to provide amplification and guidance on the current requirements within the international standard.

1.2 The purpose of these guidelines is to provide policy on the evaluation and reporting of measurement uncertainty for testing and calibration laboratories. Related topics, such as evaluation of compliance with specifications, are also included. A number of worked examples are included in order to illustrate how practical implementation of the principles involved can be achieved.

1.3 The guidance in this document is based on information in the *Guide to the Expression of Uncertainty in Measurement* [1], hereinafter referred to as the GUM. M3003 is consistent with the GUM both in methodology and terminology. It does not, however, preclude the use of other methods of uncertainty evaluation that may be more appropriate to a specific discipline. For example, the use of Bayesian statistics is becoming recognised as being particularly useful in certain areas of testing.

1.4 M3003 is aimed both at the beginner and at those more experienced in the subject of measurement uncertainty. In order to address the needs of an audience with a wide spectrum of experience, the subject is introduced in relatively straightforward terms and gives details of the basic concepts involved. Cross-references are made to a number of Appendices, where more detailed information is presented for those who wish to obtain a deeper understanding of the subject.

1.5 The following changes have been made since the publication of M3003 Edition 2:

1.5.1 Appendix A has been amended to address the replacement of the term “Best Measurement Capability” (BMC) with “Calibration and Measurement Capability” (CMC).

1.5.2 A definition of “uncertainty budget” has been added to Appendix K, Note B. Minor amendments have been made to the subsequent examples to fully address this definition.

1.5.3 Minor corrections were made to the pressure calibration example in Appendix K8.

1.5.4 The value for 99.9 % confidence in paragraph M2.16 has been corrected from 3.29 to 3.09.
The value for 84 % confidence in paragraph M2.16 has been corrected from 1.00 to 0.99.
The value for 75 % confidence in paragraph M2.16 has been corrected from 0.68 to 0.67.
Mention has also been made in paragraph M2.17 of an Excel function that can be used to evaluate these data.

1.5.5 Minor graphical, text and grammatical corrections have been made throughout the document.

1.5.6 JCGM 101:2008 has been added to the list of references in Appendix R.

1.5.7 In Appendix R, references 1, 5 and 8 have been updated to the latest editions.

United Kingdom Accreditation Service, 21-47 High Street, Feltham, Middlesex, TW13 4UN
Website: www.ukas.com Publication requests Tel: +44 (0) 20 8917 8400 Fax: +44 (0) 20 8917 8500

1.6 No further changes of any significance will be made to M3003 Edition 3 during its lifetime. However, minor text modifications of an editorial nature may be made if the need is identified. Any such changes will be listed below.

Date	Details of amendment
November 2012	First issue of M3003 Edition 3

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 sure that our estimate is within, say, 5 degrees of the “true” room temperature than we are that it 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, in order 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 where 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 Is the thermometer accurate?

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 leads to the concept of *traceability of measurements*, whereby measurements at all levels can be traced back to agreed references. In most cases, ISO/IEC17025:2005 requires that measurements are traceable to SI units. This is usually achieved by an unbroken chain of comparisons to a national metrology institute, which maintains measurement standards that are directly related to SI units.

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 the *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. They "drift". This, of course, is why 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 reading 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, the past data may not indicate a reliable trend, and a limit value may have to be assigned for the likely change since the last calibration. This 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 is also 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 error, perhaps using interpolation between points where calibration data is available. This is not always possible as it depends on the measured data being such that accurate interpolation is practical. It may then be necessary to use other information, such as the manufacturer's specification, to evaluate the additional uncertainty that arises when the reading is not directly at a point that has been calibrated.

2.8 How well can I read it?

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, the finite number of digits in the display will define the limit.

2.8.2 For example, suppose the last digit of a digital thermometer can change in steps of 0.1 °C. The reading happens to be 23.4 °C. What does this mean in terms of uncertainty?

The reading is a rounded representation of an infinite continuum of underlying values that the thermometer would indicate if it had more digits available. In the case of a reading of 23.4 °C, this means that the underlying value cannot be less than 23.35 °C, otherwise the rounded reading would be 23.3 °C. Similarly, the underlying value cannot be more than 23.45 °C, otherwise the rounded reading would be 23.5 °C.

A reading of 23.4 °C therefore means that the underlying 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. As we have no way of knowing where in this range the underlying value is, we have to assume the rounding error is zero with limits of ± 0.05 °C.

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 displayed digit. This 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 uncertainty of 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 in order to obtain a more realistic reading. In this way, we can “smooth out” the effect of short-term variations in the thermometer indication. The average, or arithmetic mean, of a number of 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 – represented by the difference from 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. 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.

2.9.5 The only information we have is a series of readings and a calculated average, or mean, value. We therefore have to use a statistical approach to determine how far our calculated mean could be away from the “true” mean. These statistics are quite straightforward and give us the uncertainty associated with the repeatability (or, more correctly, non-repeatability) of our measurements. This uncertainty is referred to as the *experimental standard deviation of the mean*. For the sake of brevity, this is often referred to as simply the *standard deviation of the mean*.

NOTE

In earlier textbooks on the subjects of uncertainty or statistics, this may be referred to as the *standard error of the mean*.

2.9.6 It is often convenient to regard the calculation of the standard deviation of the mean as a two-stage process, and it can be performed easily on most scientific calculators.

2.9.7 First we calculate the estimated standard deviation 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*.

2.9.8 The standard deviation of the mean is then obtained by dividing the value obtained in 2.9.7 by the square root of the number of measurements that contributed to the mean value.

2.9.9 Let us try an 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.

2.9.10 Using the calculator function $x\sigma_{n-1}$, we obtain an estimated standard deviation of 0.2915 °C.

2.9.11 Five measurements contributed to the mean value, so we divide 0.2915 °C by the square root of 5.

$$\frac{0.2915}{\sqrt{5}} = \frac{0.2915}{2.236} = 0.1304\ °\mathrm{C}.$$

2.9.12 Further information on the statistical analysis processes used for evaluation of non-repeatability can be found in Section 4.

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

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 in order to determine the significance of the effect.

2.10.2 How could we do this? Some fairly basic and obvious 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 this turns 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 This reveals a number of important issues. First, that the measurement may not be independent of the operator and that special consideration may have to be given to operator effects. We may have to train the operator to use the equipment in a particular 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. This is a positive benefit of uncertainty evaluation.

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

2.11.1 Maybe. 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 quite 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 constant but the relative humidity can be varied. This will reveal how sensitive the thermometer is to the quantity we are concerned about.

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

2.11.4 The significance of a particular aspect of the environment has to be considered in the light of the specific measurement being made. It is difficult to see how, for example, 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 can be seen 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. This leads on to the next question…

2.12 Does it matter where in the room I make the measurement?

2.12.1 It depends what we are trying to measure! Are we interested in the temperature at a specific location? Or 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 particular time of day, or the average over a specific period of time?

2.12.3 Such questions have to 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 in a position to evaluate the uncertainties that will arise from that method.

2.12.4 This leads to what is perhaps the most important question of all, one that should be asked before we even start with our evaluation of uncertainty:

2.13 "What exactly is it that I am trying to measure?"

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

2.13.2 The implication of this is that there has to be a defined relationship between the influence quantities and the measurand. This relationship is known as the *mathematical model*. This is an equation that describes how each influence quantity affects the value assigned to the measurand. In effect, it is a description of the measurement process. Further details about the derivation of the mathematical model can be found 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 Some measurement systems are such that the result would be only an approximation to the “true” value, even if no other uncertainties were present, because of assumptions and approximations inherent in the method. The model should include any such assumptions and therefore uncertainties that arise from them will 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 have to know our measurement system and the way in which the 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 provided an introduction to the subject of uncertainty evaluation and has explored a number of the issues involved. This section provides a more formal description of these processes, using terminology consistent with that in the GUM.

3.2 A quantity (Q) is a property of a phenomenon, body or substance to which a magnitude can be assigned. 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.

3.3 The uncertainty evaluation process will encompass a number of influence quantities that affect the result obtained for the measurand. These influence, or input, quantities are 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, … 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 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 value of the measurand, Y. As with the input uncertainties, the value of the measurand is represented by the lower-case letter, i.e. y. The uncertainty associated with y will comprise a combination of the input, or x_i, uncertainties. One of the first steps is to establish the mathematical relationship 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 all have an associated uncertainty. This is referred to as $u(x_i)$, i.e. "the uncertainty of x_i". These values of $u(x_i)$ are known as *standard uncertainties* – but more on this shortly.

3.8 Some uncertainties, particularly those associated with the determination of repeatability, have to be evaluated by statistical methods. Others have been 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 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 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 quantified by variances or standard deviations.

3.13 Therefore any convention as to how they are classified does not affect the estimation of the total uncertainty. But it should always be remembered that, in this publication, 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 character and significant in terms of uncertainty. It may then only be possible to estimate 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 uncertainties.

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 within 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 displayed digit. This concept is illustrated in Figure 1.

3.16

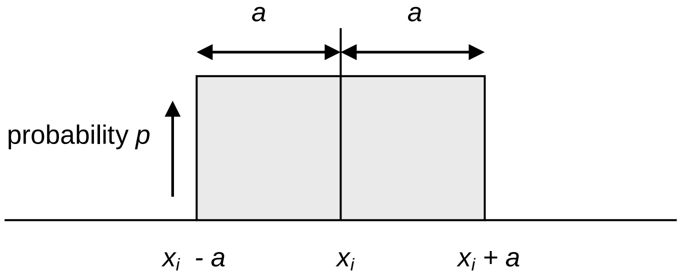


Figure 1

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 of a least significant digit.

3.18 It can be seen from this 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, a contribution of uncertainty from the influence quantity can be characterised as a *probability distribution*, i.e. a range of possible values with information about the most likely value of the input quantity x_i. In this example, it is not possible to say that any particular position of x_i within the range is more or less likely than any other. This is because there is no information available upon which to make such a judgement.

3.20 The probability distributions associated with the input uncertainties are therefore a reflection of the *available knowledge* about that particular quantity. In many cases, there will be insufficient information available to make a reasoned judgement and therefore a uniform, or rectangular, probability distribution has to be assumed. Figure 1 is an example of such a distribution.

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 scale is linear, both of these readings will have an associated rectangular distribution of identical size. If two identical rectangular distributions, each of magnitude ± a, are combined then the resulting distribution will be triangular with a semi-range of ± 2a.

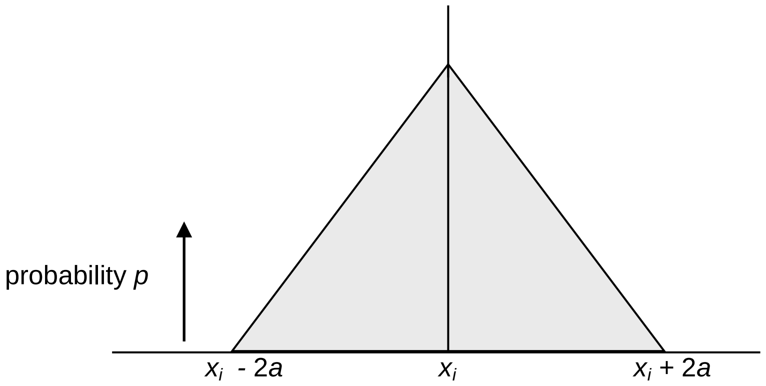


Figure 2

Combination of two identical rectangular distributions, each with semi-range limits of ± a, yields a triangular distribution with a semi-range of ± 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. This means that a cosine function characterises the probability distribution for the uncertainty. Harris and Warner[3] have shown that a symmetrical U-shaped probability distribution arises from this effect. In this example, the distribution has been evaluated from a theoretical analysis of the principles involved.

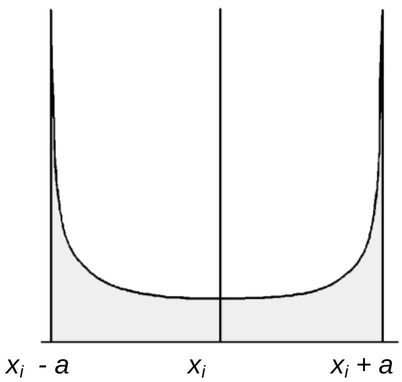


Figure 3

U-shaped distribution, associated with RF mismatch uncertainty. For this situation, x_i is 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[4]. The importance of this is that it makes it possible to assign a confidence level in terms of probability to the combined uncertainty. The exceptional case arises when one contribution to the total uncertainty dominates; in this circumstance the resulting distribution departs little from that of the dominant contribution.

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

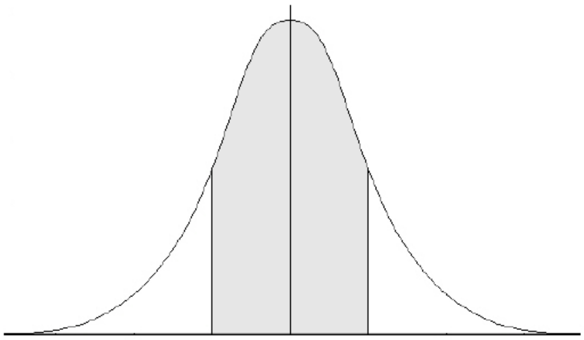


Figure 4

The normal, or Gaussian, probability distribution. This is 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 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 represents 1 standard deviation from the centre of the distribution. This corresponds to approximately 68% of the area under the curve.

3.25 When the input uncertainties are combined, a normal distribution will usually be obtained. The normal distribution is described in terms of a standard deviation. It will be therefore be necessary to express the input uncertainties in terms that, when combined, will cause the resulting normal distribution to be expressed at the one standard deviation level, like the example in Figure 4.

3.26 As some of the input uncertainties are expressed as limit values (e.g., the rectangular distribution), some processing is needed to convert them into this form, which is known as a standard uncertainty and is referred to as $u(x_i)$.

3.27 When it is possible to assess only the upper and lower bounds of an error, 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 this and for other situations.

Table 1

Assumed probability distribution	Expression used to obtain the standard uncertainty	Comments or examples
Rectangular	$u(x_i) = \frac{a_i}{\sqrt{3}}$	A digital thermometer has a least significant digit of 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 $u(x_i) = \frac{a_i}{\sqrt{3}} = \frac{0.05}{1.732} = 0.029$°C.
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 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 $u(x_i) = \frac{a_i}{\sqrt{6}} = \frac{3}{2.449} = 1.2$°C
Normal *(from repeatability evaluation)*	$u(x_i) = s(\bar{q})$	A statistical evaluation of repeatability gives the result in terms of one standard deviation; therefore no further processing is required.
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.
Normal *(from a manufacturer's specification)*	$u(x_i) = \frac{\text{Tolerance limit}}{k}$	Some manufacturers' specifications are quoted at a given coverage probability (sometimes referred to as *confidence level*), e.g. 95% or 99%. In such cases, a normal distribution can be assumed and the tolerance limit is divided by the coverage factor *k* for the stated coverage probability. 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. Indeed, they may be entirely different units altogether. 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 units.

3.29 In order 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.

3.30 The sensitivity coefficient simply describes how sensitive the result is to a particular influence quantity. In this example, the steel used in the manufacture of gauge blocks has a temperature coefficient of expansion of approximately +11.5 *x* 10^{-6} per °C. So, in this case, this figure can be used as the sensitivity coefficient.

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

$$c_i = \frac{\partial f}{\partial x_i} = \left.\frac{\partial f}{\partial X_i}\right|_{X_1 = x_1 \,..\, X_N = x_N}$$

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

3.32 The calculations required to obtain sensitivity coefficients by partial differentiation can be a lengthy process, particularly when there are many input contributions and uncertainty estimates are needed for a range of values. If the functional relationship is not known for a particular measurement system the sensitivity coefficients can sometimes be obtained by the practical approach of changing one of the input variables by a known amount, while keeping all other inputs constant, and noting the change in the output estimate. This approach can also be used if *f* is known but the determination of the partial derivatives is likely to be difficult.

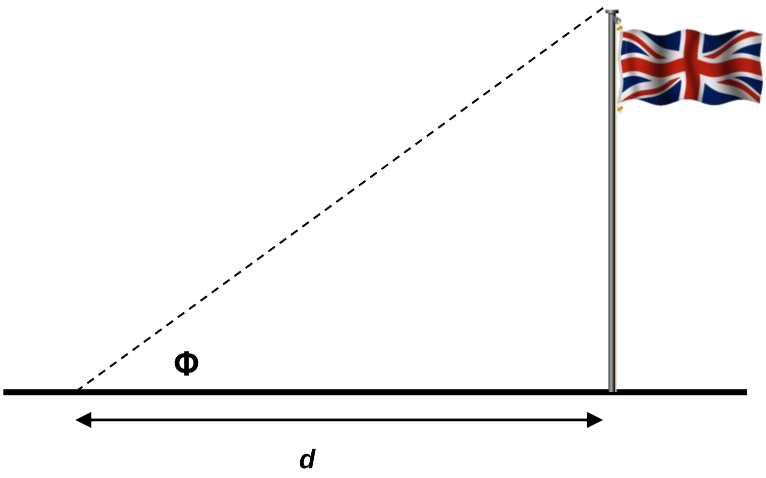
3.33 A more straightforward approach is to replace the partial derivative $\partial f / \partial x_i$ by the quotient $\Delta f / \Delta x_i$, where Δf is the change in f resulting from a change Δx_i in x_i. It is important to choose the magnitude of the change Δx_i around x_i carefully. It should be balanced between being sufficiently large to obtain adequate numerical accuracy in Δf and sufficiently small to provide a mathematically sound approximation to the partial derivative. The following example illustrates this, and why it is necessary to know the functional relationship between the influence quantities and the measurand.

<u>Example</u>

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 Φ.

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

If the measured distance is 7.0 m and the measured angle is 37°, the estimated height is 7.0 *x* tan (37) = 5.275 m.



3.34 If the uncertainty in d is, say, ± 0.1 m then the estimate of h could be anywhere between (7.0 – 0.1) tan (37) and (7.0 + 0.1) tan (37), i.e. between 5.200 m and 5.350 m. 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 $\frac{0.075}{0.1} = 0.75$.

3.35 Similar reasoning can be applied to the uncertainty in the angle Φ. If the uncertainty in Φ is ± 0.5°, then the estimate of h could be anywhere between 7.0 tan (36.5) and 7.0 tan (37.5), 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 $\frac{0.096}{0.5} = 0.192$ metre per degree.

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

3.37 The combined standard uncertainty is calculated as follows:

3.38

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

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

3.40 An example of this process is presented below, using the data from the measurement of the flagpole height described on page 15. For the purposes of the example, it is assumed that the repeatability of the process has been evaluated by making repeat 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.

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

NOTE

The values in the right-hand column of the table are now scaled in accordance with the effect of the corresponding input quantity on the measurand and are expressed as a standard uncertainty.

3.41 In accordance with the Central Limit Theorem, this combined standard uncertainty takes the form of a normal distribution. As the input uncertainties had been expressed in terms of a standard uncertainty, the resulting normal distribution is expressed as one standard deviation, as illustrated in Figure 5.

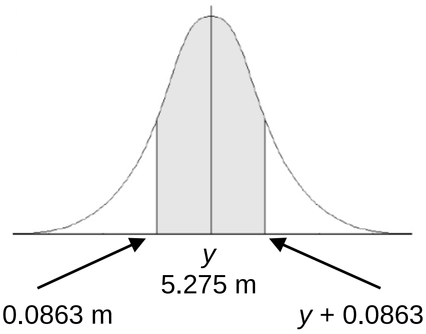


Figure 5

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

3.42 For a normal distribution, 1 standard deviation encompasses 68.27% of the area under the curve. This means that there is about 68% confidence that the measured value y lies within the stated limits.

3.43 The GUM recognises the need for providing a high level of confidence – referred to herein as *coverage probability* - associated with an uncertainty and uses the term *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

$$U = k\, u_c(y). \qquad (2)$$

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 $k = 2$ actually provides a coverage probability of 95.45% for a normal distribution. For convenience this is approximated to 95% which would relate to a coverage factor of $k = 1.96$. However, the difference is not generally significant since, in practice, the coverage probability is usually based on conservative assumptions and approximations to the *true* probability distributions.

3.45 Example: The measurement of the height of the flagpole had a combined standard uncertainty $u_c(y)$ of 0.0863 m. Hence the expanded uncertainty $U = k\, u_c(y) = 2 \times 0.0863 = 0.173$ m.

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

3.47 There may also be situations where a normal distribution can be assumed, but 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 confidence for a normal distribution.

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

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 or randomness of a measurement process. For some measurements, the 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 significant spread in a sample of measurement results, the arithmetic mean or average of the results should be calculated. If there are *n* independent repeated values for a quantity *Q* then the mean value $\bar{q}$ is given by

$$\bar{q} = \frac{1}{n}\sum_{j=1}^{n} q_j = \frac{q_1 + q_2 + q_3 \cdots q_n}{n} \tag{3}$$

4.3 The spread in the results gives an indication of the repeatability of the measurement process, which depends on various factors, including the apparatus used, the method, and sometimes on the person making the measurement. A good description of this spread of values is the standard deviation σ of the *n* values that comprise the sample, which is given by

$$\sigma = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(q_j - \bar{q}\right)^2} \tag{4}$$

4.4 This expression yields the standard deviation σ of the particular set of values sampled. However, these are not the only values that could have been sampled. If the process is repeated, another set of values, with different values of $\bar{q}$ and σ, will be obtained.

4.5 For large values of *n*, these mean values approach the central limit of a distribution of all possible values. This probability distribution can often be assumed to have the normal form.

4.6 As it is impractical to capture all values that are available, it is necessary to make an estimate of the value of σ that would be obtained were this possible. Similarly, the mean value obtained is less likely to be the same as that which would be obtained if a very large number of measurements could be taken, therefore an estimate has to be made of the possible error from the “true” mean.

4.7 Equation (4) gives the standard deviation for the samples actually selected, rather than of the whole population of possible samples. However, from the results of a single sample of measurements, an estimate, $s(q_j)$, can be made of the standard deviation σ of the whole population of possible values of the measurand from the relation

$$s\left(q_j\right) = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\left(q_j - \bar{q}\right)^2} \tag{5}$$

4.8 The mean value $\bar{q}$ will have been derived from a finite number *n* of samples and therefore its value will not be the exact mean that would have been obtained if an infinite number of samples could have been taken. The mean value itself therefore has uncertainty. This uncertainty is referred to as the *experimental standard deviation of the mean*. It is obtained from the estimated standard deviation of the population by the expression:

$$s\left(\bar{q}\right) = \frac{s\left(q_j\right)}{\sqrt{n}} \tag{6}$$

4.9 Example: Four measurements were made to determine the repeatability of a measurement system. The results obtained were 3.42, 3.88, 2.99 and 3.17.

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

The estimated standard deviation $s(q_j) = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\left(q_j - \bar{q}\right)^2} = 0.386$

The experimental standard deviation of the mean $s(\bar{q}) = \frac{s(q_j)}{\sqrt{n}} = \frac{0.386}{\sqrt{4}} = 0.193$

For information on the use of calculators to calculate $s(q_j)$, see Paragraph P4.

4.10 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 of a measurement system may be obtained from data obtained previously, based on a larger number of readings.

4.11 Whenever possible at least two measurements should be made as part of the procedure; however, it is acceptable for a single measurement to be made even though it is known that the system has imperfect repeatability, and to rely on a previous assessment of the repeatability of similar devices. This procedure must be treated with caution because the reliability of a previous assessment will depend on the number of devices sampled and how well this sample represents all devices. It is also recommended that data obtained from prior assessment should be regularly reviewed. Of course, when only one measurement is made on the device being calibrated a value of $s(q_j)$ must have been obtained from prior measurements, and n in Equation (6) is then 1.

In such cases, the estimated standard deviation $s(q_j)$ is given by

$$s(q_j) = \sqrt{\frac{1}{m-1}\sum_{j=1}^{m}\left(q_j - \bar{q}\right)^2}, \qquad (7)$$

4.12 where m is the number of readings considered in the previous evaluation. The standard deviation of the mean $s(\bar{q}) = \frac{s(q_j)}{\sqrt{n}}$, n being the number of measurements contributing to the reported mean value.

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.

4.13 The standard uncertainty is then the standard deviation of the mean, i.e. $u_i(x) = s(\bar{q})$. (8)

4.14 A previous estimate of standard deviation can only be used if there has been no subsequent change in the measurement system or procedure that could have an effect on the repeatability. If an apparently excessive spread in measurement values is found, which is not typical of the measurement system, the cause should be investigated and resolved before proceeding further.

5 TYPE B EVALUATION OF STANDARD UNCERTAINTY

5.1 It is probable that systematic components of uncertainty, i.e. those that account for errors that remain constant while the measurement is made, will be estimated from Type B evaluations. The most important of these systematic components, for a reference instrument, will often be the imported uncertainties associated with its own calibration. However, there can be, and usually are, other important contributions to systematic errors in measurement that arise in the equipment user's own laboratory.

5.2 The successful identification and evaluation of these contributions depends on a detailed knowledge of the measurement process and the experience of the person making the measurements. The need for the utmost vigilance in preventing mistakes cannot be overemphasised. Common examples are errors in the corrections applied to values, 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.

5.3 In evaluating the components of uncertainty it is necessary to consider and include at least the following possible sources:

(a) The reported calibration uncertainty assigned to reference standards and any drift or instability in their values or readings.

(b) The calibration of measuring equipment, including ancillaries such as connecting leads etc., and any drift or instability in their values or readings.

(c) 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 the item being calibrated is not normally included in the uncertainty evaluation for that calibration.

(d) The operational procedure.

(e) Variability between different staff carrying out the same type of measurement.

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

5.4 Whenever possible, corrections should be made for errors revealed by calibration or other sources; the convention is that an error is given a positive sign if the measured value is greater than the conventional true value. The correction for error 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, as if it were a systematic uncertainty equal to (±) the uncorrected error magnitude.

5.5 Having identified all the possible systematic 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 the assessed 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 well-defined geometric shapes for which the standard uncertainty can be obtained from a simple calculation. These distributions and sample calculations are presented in detail in paragraphs 3.15 to 3.22.

United Kingdom Accreditation Service, 21-47 High Street, Feltham, Middlesex, TW13 4UN
Website: www.ukas.com Publication requests Tel: +44 (0) 20 8917 8400 Fax: +44 (0) 20 8917 8500

6 REPORTING OF RESULTS

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

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 ν_{eff} = YY effective degrees of freedom corresponds to a coverage probability of approximately 95%. The uncertainty evaluation has been carried out in accordance with UKAS requirements".

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

6.6 In the special circumstances where a dominant non-Gaussian Type B contribution occurs refer to Appendix C. If uncertainty is being reported as an analytical expression, refer to Appendix L .

6.7 Uncertainties are usually expressed in bilateral terms (±) either in units of the measurand or as relative values, for example as a percentage (%), parts per million (ppm), 1 in 10^X, etc. However there may be situations where the upper and lower uncertainty values are different; for example if cosine errors are involved. If such differences are small then the most practical approach is to report the expanded uncertainty as the larger of the two. However if there is a significant difference between the upper and lower values then they should be evaluated and reported separately.

6.8 The number of figures in a reported uncertainty should always reflect practical measurement capability. In view of the process for estimating uncertainties it is seldom justified to report more than two significant figures. It is therefore recommended that the expanded uncertainty be rounded to two significant figures, using the normal rules of rounding. The numerical value of the measurement result should in the final statement normally be rounded to the least significant figure in the value of the expanded uncertainty assigned to the measurement result.

6.9 Rounding should always be carried out at the end of the process in order to avoid the effects of cumulative rounding errors.

7 STEP BY STEP PROCEDURE FOR EVALUATION OF MEASUREMENT UNCERTAINTY

The following is a guide to the use of this code of practice for the treatment of uncertainties. The left hand column gives the general case while the right hand column indicates how this relates to example K4 in Appendix K. Although this example relates to a calibration activity, the process for testing activities is unchanged.

	General case	**Example K4: Calibration of a weight of nominal value 10 kg of OIML Class M1**
7.1	If possible determine the mathematical relationship between values of the input quantities and that of the measurand: $y = f(x_1, x_2, \dots x_N)$ See Appendix D for details.	It will be assumed that the unknown weight, W_X, can be obtained from the following relationship: $W_X = W_S + D_S + \delta I_d + \delta C + Ab$
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, however, uncertainties for these contributions have been determined, therefore: Drift of standard mass since last calibration: 0 Correction for air buoyancy: 0 Linearity correction: 0 Effect of least significant digit resolution: 0

	General case	Example K4: Calibration of a weight of nominal value 10 kg of OIML Class M1
7.3	List systematic components of uncertainty associated with 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 other documents if the assumed probability distribution is not covered in this publication.	Source of uncertainty — Limit (mg) — Distribution W_S Calibration of std. mass — 30 — Normal (k = 2) D_S Drift of standard mass — 30 — Rectangular δC Comparator linearity — 3 — Rectangular δAb Air buoyancy — 10 — Rectangular δI_d Resolution effects — 10 — Triangular Then: $u(x_1) = u(W_S) = \frac{30}{2}$ = 15 mg $u(x_2) = u(D_S) = \frac{30}{\sqrt{3}}$ = 17.32 mg $u(x_3) = u(\delta C) = \frac{3}{\sqrt{3}}$ = 1.73 mg $u(x_4) = u(Ab) = \frac{10}{\sqrt{3}}$ = 5.77 mg $u(x_5) = u(\delta I_d) = \frac{10}{\sqrt{6}}$ = 4.08 mg
7.4	Use prior knowledge or make trial measurements and calculations to determine if there is going to be a random component of uncertainty that is significant compared with the effect of the listed systematic components of uncertainty. Random components of uncertainty also have to be considered as input quantities.	From previous knowledge of the measurement system 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): $\bar{q} = \frac{1}{n}\sum_{j=1}^{n} q_j$	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 g

	General case	Example K4: Calibration of a weight of nominal value 10 kg of OIML Class M1
7.6	Either calculate the standard deviation of the mean value from Equations (5) and (6): $$s(q_j)=\sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(q_j-\bar{q})^2}$$ $$s(\bar{q})=\frac{s(q_j)}{\sqrt{n}}$$ or refer to the results of previous repeatability evaluations for an estimate of $s(q_j)$ based on a larger number of readings, using Equation (7): $$s(q_j)=\sqrt{\frac{1}{m-1}\sum_{j=1}^{m}(q_j-\bar{q})^2}$$ $$s(\bar{q})=\frac{s(q_j)}{\sqrt{n}}$$ where m is the number of readings used in the prior evaluation and n is the number of readings that contribute to 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 10 measurements using the conventional bracketing technique and was calculated, using Equation (5), to be 8.7 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 (6): $$s(\overline{W_X})=\frac{s(W)}{\sqrt{n}}=\frac{8.7}{\sqrt{3}}=5.0\text{ mg}$$
7.7	Even when a random component of uncertainty is not significant, where possible check the instrument indication at least once to minimise the possibility of unexpected errors.	
7.8	Derive the standard uncertainty for the above Type A evaluation from Equation (8): $$u(x_i)=s(\bar{q})$$	This is then the standard uncertainty for the Type A evaluation: $$u(x_6)=u(\overline{W_R})=s(\overline{W_R})=5.0\text{ mg}$$

	General case	Example K4: Calibration of a weight of nominal value 10 kg of OIML Class M1
7.9	Calculate the combined standard uncertainty for uncorrelated input quantities using Equation (1) if absolute values are used: $u_c(y) = \sqrt{\sum_{i=1}^{N} c_i^2 u^2(x_i)} \equiv \sqrt{\sum_{i=1}^{N} u_i^2(y)}$, where c_i is the partial derivative $\partial f / \partial x_i$, or a known sensitivity coefficient. Alternatively use Equation (11) if the standard uncertainties are relative values: $\frac{u_c(y)}{\lvert y \rvert} = \sqrt{\sum_{i=1}^{N} \left[\frac{p_i u(x_i)}{\lvert x_i \rvert} \right]^2}$, 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 ($c_i = 1$).
7.10	If correlation is suspected use the guidance in paragraph D3 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: $u(W_X) = \sqrt{15^2 + 17.32^2 + 4.08^2 + 1.73^2 + 5.77^2 + 5.0^2}$ = 24.55 mg.
7.11	Either calculate an expanded uncertainty from Equation (2): $U = k \cdot u_c(y)$ or, if there is a significant random contribution evaluated from a small number of readings, use Appendix B to calculate a value for k_p and use this value to calculate the expanded uncertainty.	U = 2 x 24.55 mg = 49.10 mg It was not necessary to use Appendix B to determine a value for k_p. In fact the effective degrees of freedom of $u(W_X)$ are greater than 5000 which gives a value for k_{95} = 2.00.
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 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.

APPENDIX A
CALIBRATION AND MEASUREMENT CAPABILITY

A1 A CMC is a Calibration and Measurement Capability available to customers under normal conditions. Calibration and Measurement Capability is a term normally used to describe the uncertainty that appears in an accredited calibration laboratory's schedule of accreditation and is the uncertainty for which the laboratory has been accredited using the procedure that was the subject of assessment.

NOTES

[1] 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.

[2] The meanings of the terms Calibration and Measurement Capability, CMC, and Best Measurement Capability, BMC, (as used historically in connection with the uncertainties stated in the scope of an accredited laboratory) are identical. The terms BMC and CMC should be interpreted similarly and consistently in the current areas of application.

[3] This Appendix is based on the parts of ILAC-P14:12/2010[9] that address the subject of Calibration and Measurement Capability.

A2 The CMC should be calculated according to the procedures given in this document and should normally be quoted as an expanded uncertainty at a coverage probability of 95%, which usually requires the use of a coverage factor of $k = 2$.

A3 An accredited laboratory is not permitted to report an uncertainty smaller than its accredited CMC but may report an equal or larger uncertainty. The magnitude of the uncertainty reported on a certificate of calibration will often depend on properties of the device being calibrated. No device is perfect and so the concept of a “best existing device” is used in association with the evaluation of a CMC. CMC uncertainty statements therefore incorporate agreed values for the best existing devices. Where necessary the laboratory's schedule of accreditation will include remarks that describe the conditions under which the CMC can be achieved.

A4 A best existing device is one that is available but does not necessarily represent the majority of devices that the laboratory may be asked to calibrate. The properties of devices that are considered to be the “best existing” will depend on the field of calibration but may include an instrument with very low random fluctuations, negligible temperature coefficient, very low voltage reflection coefficient etc. The uncertainty budget that is intended to demonstrate the CMC should still include contributions from the properties of the best existing device but the value of the uncertainty may be entered as zero or a negligible value, if this is the case. Where applicable the uncertainty shall include the same contributions to uncertainty that were included in evaluation of the CMC uncertainty component, except that uncertainty components evaluated for the best existing device shall be replaced with those of the customer’s device.

A5 It may also be the case that a laboratory only intends to calibrate devices that do not represent the state of the art, or best existing, devices. In such cases, for the purpose of CMC evaluation, the type of device they intend to calibrate may be considered to be the “best existing” device for that laboratory and the characteristics of that device should be incorporated into the CMC.

A6 Effects on performance that are caused by the customer’s device itself before or after its calibration (such as those due to transport) should normally be excluded from the uncertainty statement. If, however, a laboratory anticipates that such contributions will be significant compared to the uncertainties assigned by the laboratory, the customer should be notified according to the general clauses regarding tenders and reviews of contracts in ISO/IEC 17025:2005.

A7 It is sometimes the case that a laboratory may wish to be accredited for a measurement uncertainty that is larger than it can actually achieve. If the principles of this document are followed when constructing the uncertainty budget the resulting expanded uncertainty should be a realistic representation of the laboratory's measurement capability. If this is smaller than the uncertainty the laboratory wishes to be accredited for and report on their certificates of calibration, the implication is that the laboratory is uncomfortable in some way about the magnitude of the expanded uncertainty. If this is the case then the contributions to the uncertainty budget should be reviewed and consideration given to making more conservative allowances as necessary.

A8 It can be the case that some calibration laboratories offer a CMC with uncertainties than are smaller than are routinely offered for everyday calibrations. This is because they will maintain their own reference standards, upon which the CMC is based, but use subsidiary equipment – often automated – for routine work. The contract review arrangements between the laboratory and its customer should define the level of service being offered.

A9 In some cases the CMC quoted in a laboratory's schedule has to cover a two (or more) dimensional range of measured values, such as different levels and frequencies, and it may not be practical to give the actual uncertainty for all possible values of the quantities. In these cases the CMC may be given as a range of uncertainties appropriate to the upper and lower values of the uncertainty that has been calculated for the range of the quantity, or may be described as an analytical expression. Guidance about the expression of uncertainty over a range of values is presented in Appendix L.

APPENDIX B
DERIVING A COVERAGE FACTOR FOR UNRELIABLE INPUT QUANTITIES

B1 In the majority of measurement situations it will be possible to evaluate Type B uncertainties with high reliability. Furthermore, if the procedure followed for making the measurements is well established and if the Type A evaluations are obtained from a sufficient number of observations then the use of a coverage factor of $k = 2$ will mean that the expanded uncertainty, U, will provide an interval with a coverage probability close to 95%. This is because the distribution tends to normality as the number of observations increases and $k = 2$ corresponds to 95% confidence for a normal distribution.

B2 However, in some cases it may not be practical to base the Type A evaluation on a large number of readings, which 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 correctly 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.

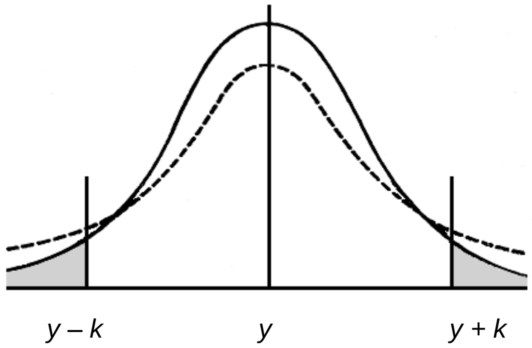


Figure 6

In Figure 6, the solid line indicates the normal distribution. A specified proportion p of the values under the curve are encompassed between $y - k$ and $y + k$. An example of the t-distribution is superimposed, using dashed lines. For the t-distribution, a greater proportion of the values lie outside the region $y - k$ to $y + k$, and a smaller proportion lie 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 of t_p, and hence k_p, from the t-distribution table.

B3 In order to obtain a value for k_p it is necessary to obtain an estimate of the effective degrees of freedom, ν_{eff}, of the combined standard uncertainty $u_c(y)$. The GUM recommends that the Welch-Satterthwaite equation is used to calculate a value for ν_{eff} based on the degrees of freedom, ν_i, of the individual standard uncertainties $u_i(y)$; therefore

$$\nu_{eff} = \frac{u_c^4(y)}{\sum_{i=1}^{N} \frac{u_i^4(y)}{\nu_i}} \qquad (9)$$

B4 The degrees of freedom, ν_i, for contributions obtained from Type A evaluations are $n - 1$, where n is the number of readings used to evaluate $s(q_j)$.

B5 It is often possible to take the degrees of freedom, v_i , of Type B uncertainty contributions as infinite, that is, their value is known with a very high degree of reliability. If this is the case, and there is only one contribution obtained from a Type A evaluation, then the process using the Welch-Satterthwaite formula simplifies, as all the terms relating to the type B uncertainties become zero. This is illustrated in the example in paragraph B10.

B6 However, a Type B contribution may have come from a calibration certificate as an expanded uncertainty based on a *t*-distribution rather than a normal distribution, as described in this Appendix. This is an example of a Type B uncertainty that does not have infinite degrees of freedom. For this eventuality the degrees of freedom will be as quoted on the calibration certificate, or it can be obtained from the *t*-distribution table for the appropriate value of k_{95} .

B7 Having obtained a value for v_{eff} , the *t*-distribution table is used to find a value of k_p. This table, reproduced below, gives values for k_p, for various levels of confidence *p*. Unless otherwise specified, the values corresponding to p = 95.45% should be used.

Degrees of freedom v	Values of $t_p(v)$ from the *t*-distribution for degrees of freedom v that define an interval that encompasses specified fractions p of the corresponding distribution					
	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	235.80
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

B8 Normally v_{eff} will not be an integer and it will be necessary to interpolate between the values given in the table. Linear interpolation will suffice for v_{eff} >3; higher-order interpolation should be used otherwise. Alternatively, the next lower value of v_{eff} may be used.

B9 The value of $t_p(v)$ obtained from the table is the coverage factor k_p that is required to calculate the expanded uncertainty, U_p, from $U_p = k_p\, u_c(y)$. Unless otherwise specified, the coverage probability p will usually be 95%.

B10 Example

B10.1 In a measurement system a Type A evaluation, based on 4 observations, gave a value of $u_i(y)$ of 3.5 units using Equations (4) and (5). 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, $u_c(y)$, had a value of 5.7 units. Then using the Welch-Satterthwaite formula (paragraph B3, Equation 9):

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

B10.2 The value of v_{eff} given in the *t*-distribution table, for a coverage probability p of 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 x 2.13 = 12.14 units.

APPENDIX C
DOMINANT NON-GAUSSIAN TYPE B UNCERTAINTY

C1.1 In some measurement processes there can be one component of uncertainty derived from a Type B evaluation that is dominant in magnitude compared with the other components. When the dominant component is characterised by limits within which there is a high probability of occurrence, a calculated expanded uncertainty, U, using the coverage factor of $k = 2$, may be greater than the arithmetic sum of the semi-range of all the individual limiting values. As it is reasonable to assume that the arithmetic sum of these contributions would be for a coverage probability approaching 100%, there is a degree of pessimism in following the normal recommended procedure for combination of uncertainties.

C1.2 Consequently special consideration needs to be given to the situation in which the calculated expanded uncertainty fails to meet the criterion $U \leq$ arithmetic sum of the limit values of all contributions. For practical purposes, the “limit” value of a normal distribution can be taken as three times its standard deviation.

C1.3 In many cases this criterion will be met, but it can be the case that a single rectangular distribution may dominate over other contributions, A commonly encountered example is the resolution of a digital indicating instrument. This will have a standard uncertainty of $\frac{a_i}{\sqrt{3}}$. If this is large compared to other contributions it is less likely that the criterion above will be met. When it is not met then the dominant contribution, a_d, should be extracted and a new value of the expanded uncertainty should be reported as $U = U' + a_d$, where U' is the expanded uncertainty for the remaining components, treated in the normal statistical manner.

C1.4 The use of a two-part expression such as this means that when both components are imported into a subsequent uncertainty budget it is likely that a_d will no longer be a dominant component and a normal distribution can be assumed for the subsequent combined standard uncertainty.

C1.5 Under these circumstances, the statement of coverage probability associated with the expanded uncertainty will have to be modified; an example is given below:

The uncertainty is stated in two parts, the second of which is dominant and is due to the uncertainty due to the resolution of the instrument being calibrated, for which a rectangular probability distribution has been assumed.

C1.6 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.

C1.7 There may be, however, situations where a single value of uncertainty is required even if there is a Type B uncertainty that causes the distribution to be non-normal. This will involve evaluation of a coverage factor for a stated coverage probability for the convolved distributions.

C1.8 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:

$\frac{u_i(y)_{normal}}{u_i(y)_{rect}}$	$k_{95.45}$	$\frac{u_i(y)_{normal}}{u_i(y)_{rect}}$	$k_{95.45}$	$\frac{u_i(y)_{normal}}{u_i(y)_{rect}}$	$k_{95.45}$
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

C1.9 <u>Example</u>

A digital voltmeter is calibrated with an applied voltage of 1.0000 V; the resulting reading is 1.001 V. The expanded uncertainty of the applied voltage is 0.0002 V (*k* = 2). The only other uncertainty of significance is due to the rounding of the indicator display. The indicator can display readings in steps of 0.001 V; therefore there will be a possible rounding error of ± 0.0005 V. A rectangular probability distribution is assumed.

The uncertainty budget will therefore be as follows:

Symbol	Source of uncertainty	Value ± V	Probability distribution	Divisor	c_i	$u_i(V)$ V	ν_i or ν_{eff}
V_s	Uncertainty of applied voltage	0.0002	Normal	2	1	0.000100	∞
δI_d	Digital rounding of indicator	0.0005	Rectangular	√3	1	0.000289	∞
$u_c(V)$	Combined standard uncertainty		Convolved $\frac{0.000100}{0.000289} = 0.346$			0.000306	∞
U	Expanded uncertainty		Convolved $k = 1.77$			0.000541	∞

Reported result

For an applied voltage of 1.000 V the voltmeter reading was 1.001 V ± 0.00054V.

The reported expanded uncertainty is based on a convolution of a dominant rectangular uncertainty with other, smaller, uncertainties. The resulting standard uncertainty has been multiplied by a coverage factor *k* = 1.77, which, for this particular convolution, corresponds to a coverage probability of approximately 95%. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

NOTE

A simple test to determine whether a rectangular uncertainty is a dominant component is to check whether its standard uncertainty is more than 1.4 times the combined standard uncertainty for the remaining components. If it is not, then $\frac{u_i(y)_{normal}}{u_i(y)_{rect}} \geq 0.71$ and the coverage factor *k* will be within 5% of the usual value of 2.00.

C1.10 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:

$\frac{u_i(y)_{normal}}{u_i(y)_{U-shaped}}$	$k_{95.45}$	$\frac{u_i(y)_{normal}}{u_i(y)_{U-shaped}}$	$k_{95.45}$	$\frac{u_i(y)_{normal}}{u_i(y)_{U-shaped}}$	$k_{95.45}$
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

C1.11 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:

$\frac{u_i(y)_{rect}}{u_i(y)_{U-shaped}}$	$k_{95.45}$	$\frac{u_i(y)_{rect}}{u_i(y)_{U-shaped}}$	$k_{95.45}$	$\frac{u_i(y)_{rect}}{u_i(y)_{U-shaped}}$	$k_{95.45}$
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

C1.12 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:

Ratio $\frac{u_i(y)_{smaller}}{u_i(y)_{larger}}$	*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

APPENDIX D
DERIVATION OF THE MATHEMATICAL MODEL

D1 Measurement process

D1.1 In general a measurement process can be regarded as having estimated input quantities X_i whose values, given the symbol *x,* contribute to the estimated value *y* of the measurand or output quantity *Y*. Where, as in most cases, there are several input quantities, the standard uncertainty associated with the estimated value of each one is represented by $u(x_i)$. Standard uncertainty and its evaluation are discussed in Sections 4 and 5.

D1.2 The measurement process can usually be modelled by a functional relationship between the values of the estimated input quantities and that of the output estimate in the form

$$y = f\,(x_1,\, x_2,\, \ldots\, x_N)$$

For example, if electrical resistance *R* is measured in terms of voltage *V* and current *I* then the relationship is $R = f(V,I) = V/I$. The mathematical model of the measurement process is used to identify the input quantities that need to be considered in the uncertainty budget and their relationship to the total uncertainty for the measurement.

D1.3 Example

The following example of how a mathematical model can be derived relates to example K1 in Appendix K.

In this example two resistors, R_S and R_X, are compared by connecting them in series and passing a constant current through them. The voltage across each is measured. As the same current passes through both resistors the ratio of the two voltages V_S and V_X will be the same as the ratio of the two resistance values, i.e.

$$\frac{R_X}{R_S} = \frac{V_X}{V_S}$$

If the resistance of R_S is known the value of R_X can be determined by rearranging the equation as follows:

$$R_X = R_S \cdot \frac{V_X}{V_S}$$

It is known that there will be various uncertainties associated with the measurement. In the expression above, the calibration uncertainty of R_S means that the same relative uncertainty will exist in the value of R_X. There are, however, further uncertainties to be considered. For example, the value of R_S will change with time, therefore a contribution δR_D has to be included – the model now becomes:

$$R_X = \left(R_S + \delta R_D\right) \cdot \frac{V_X}{V_S}$$

Another consideration is the effect of temperature. The temperature coefficient of resistors of this type is usually described by a parabolic curve. Taking the partial derivatives of the expression describing this curve is not a straightforward process; therefore the practical approach of estimated the temperature coefficient at the nominal temperature was taken, using a linear approximation. This value for the temperature coefficient, R_{TC}, is multiplied by the value assigned for temperature variations, Δt, and the product is inserted in the model:

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

Another influence is the ratio V_X/V_S. In this case, the same voltmeter is used to measure both V_S and V_X. Any systematic error, or offset, in the voltmeter reading will cancel because it is the same for both voltages - the ratio will be unchanged. All that have to be considered are secondary influences, such as resolution and linearity, and it can be seen from examination of the uncertainty budget for example K1 that these influences have been considered separately. This is a good example of negative correlation.

The repeatability of the process also has to be included. As the parameters directly observed during the calibration are the voltages V_X and V_S, it is convenient to assess the repeatability in terms of observed changes in the ratio V_X/V_S. This is carried out in accordance with Equations (4) and (5), which yield the experimental standard deviation of the mean, $s\left(\overline{V}\right)$. The model then becomes:

$$R_X = \left(R_S + \delta R_D + \left(R_{TC} \cdot \Delta t\right) + s\left(\overline{V}\right)\right) \cdot \frac{V_X}{V_S}$$

This is how the model in Example K1 was derived.

D2 Model descriptions

D2.1 If the functional relationship is an addition or subtraction of the input quantities, for example

$$W_X = f\,(W_S, D_S, \delta I_d, \delta C, Ab) = W_S + D_S + \delta I_d + \delta C + Ab,$$

then all the input quantities will be directly related to the output quantity and the partial derivatives will all be unity.

D2.2 In paragraph 3.35 it is stated that the combined standard uncertainty is calculated using the expression

$$u_c(y) = \sqrt{\sum_{i=1}^{N} c_i^{\,2} u^2(x_i)} \equiv \sqrt{\sum_{i=1}^{N} u_i^{\,2}(y)} \qquad (10)$$

D2.3 If the functional relationship is a product or quotient, i.e. the output quantity is obtained from only the multiplication or division of the input quantities, this can be transformed to a linear addition by the use of relative values, e.g. those expressed in % terms or in parts per million. The general form is $y = cx_1^{p_1} \cdot x_2^{p_2} \cdots x_N^{p_N}$ where the exponents p_i are known positive or negative numbers. The standard uncertainty will then be given by

$$\frac{u_c(y)}{|y|} = \sqrt{\sum_{i=1}^{N} \left[\frac{p_i\, u\,(x_i)}{|x_i|}\right]^2} \qquad (11)$$

D2.4 Some examples of the use of relative uncertainties are

$$P = f(V, I) = V \cdot I, \text{ so } \frac{u(P)}{P} = \sqrt{\left[\frac{u(V)}{V}\right]^2 + \left[\frac{u(I)}{I}\right]^2}$$

$$P = f(V, R) = V^2/R, \text{ so } \frac{u(P)}{P} = \sqrt{\left[\frac{2u(V)}{V}\right]^2 + \left[\frac{u(R)}{R}\right]^2}$$

$$V = f(P, Z) = (P \cdot Z)^{½}, \text{ so } \frac{u(V)}{V} = \sqrt{\left[\frac{u(P)}{2P}\right]^2 + \left[\frac{u(Z)}{2Z}\right]^2}$$

D2.5 The use of relative uncertainties can often simplify the calculations and is particularly helpful when the input quantities and the uncertainties are already given in relative terms. However, sensitivity coefficients may still be required to account for known relationships, such as a temperature coefficient. Relative uncertainties should not be used when the functional relationship is already an addition or subtraction.

D3 Correlated input quantities

D3.1 The expressions given for the standard uncertainty of the output estimate, Equations (10) and (11), will only apply when there is no correlation between any of the input estimates, that is, the input quantities are independent of each other. It may be the case that some input quantities are affected by the same influence quantity, e.g. temperature, or by the errors in a particular instrument that is used for separate measurements in the same process. In such cases the input quantities are not independent of each other and the equation for obtaining the standard uncertainty of the output estimate must be modified.

D3.2 The effects of correlated input quantities may serve to reduce the combined standard uncertainty, such as when an instrument is used as a comparator between a standard and an unknown, and this is referred to as negative correlation. In other cases measurement errors will always combine in one direction and this has to be accounted for by an increase in the combined standard uncertainty. This is referred to as positive correlation. Knowledge concerning the possibility of correlation can often be obtained from the functional relationship between the input quantities and the output quantity but it may also be necessary to investigate the effects of correlation by making a planned series of measurements.

D3.3 If positive correlation between input quantities is suspected but the degree of correlation cannot easily be determined then the most straightforward solution is to add arithmetically the standard uncertainties for these quantities to give a new standard uncertainty that is then dealt with in the usual manner in Equation (1) or (11). The GUM should be consulted for a more detailed approach to dealing with correlation based on the calculation of correlation coefficients.

D3.4 An example of the treatment of correlated contributions is shown in paragraph K6.4.

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.

E1 Imported uncertainty

E1.1 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.

E2 Secular stability

E2.1 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 has to 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.

E2.2 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 compliance 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.

E3 Environmental conditions

E3.1 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 has to 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.

United Kingdom Accreditation Service, 21-47 High Street, Feltham, Middlesex, TW13 4UN
Website: www.ukas.com Publication requests Tel: +44 (0) 20 8917 8400 Fax: +44 (0) 20 8917 8500

E3.2 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.

E3.3 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.

E4 Interpolation of calibration data

E4.1 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.

E4.2 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.

E5 Resolution

E5.1 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.

E5.2 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 displayed 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 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.

E5.3 This "digital rounding 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 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.

E5.4 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.

E6 Apparatus layout

E6.1 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.

E7 Thermoelectric voltages

E7.1 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.

E7.2 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.

E7.3 Generally an allowance has to be made as a Type B component of uncertainty for the presence of thermal emfs.

E8 Loading and cable impedance

E8.1 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.

E8.2 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.

E8.3 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

E9.1 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.

E9.2 In a power measurement system, the power, P_O, that would be absorbed in a load equal to the characteristic impedance of the transmission line has been shown[3] to be related to the actual power, P_L, absorbed in a wattmeter terminating the line by the equation

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

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

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

E9.3 When φ is unknown, this expression for absorbed power can have limits:

$$P_0\,(limits) = \frac{P_L}{1-|\Gamma_L|^2}\left(1 \pm 2|\Gamma_G||\Gamma_L|\right) \quad \text{E(3)}$$

E9.4 The calculable mismatch error is $1 - |\Gamma_L|^2$ and is accounted for in the calibration factor, while the limits of mismatch uncertainty are $\pm\, 2\,|\Gamma_L|\,|\Gamma_G|$. Because a cosine function characterises the probability distribution for the uncertainty, Harris and Warner[3] show that the distribution is U-shaped with a standard deviation given by

$$u(mismatch) = \frac{2\,|\Gamma_G||\Gamma_L|}{\sqrt{2}} = 1.414\,\Gamma_G\,\Gamma_L \quad \text{E(4)}$$

E9.5 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[3] have shown that the standard deviation of mismatch, M, expressed in dB is approximated by

$$M = \frac{8.686}{\sqrt{2}}\left[|\Gamma_G|^2\left(|s_{11a}|^2+|s_{11b}|^2\right)+|\Gamma_L|^2\left(|s_{22a}|^2+|s_{22b}|^2\right)+|\Gamma_G|^2\cdot|\Gamma_L|^2\left(|s_{21a}|^4+|s_{21b}|^4\right)\right]^{0.5} \quad \text{E(5)}$$

where 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[3] concluded that the distribution for *M* would approximate to that of a normal distribution due to the combination of its component distributions.

E9.6 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.

E10 Directivity

E10.1 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.

E10.1 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 (34.4 dB).

E11 Test port match

E11.1 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.

E12 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. Although connecting and disconnecting the device can evaluate the repeatability of particular connector pairs in use, these connector pairs are only samples from a whole population. To obtain representative data for guidance for various types of connectors in use is beyond the resources of most measurement laboratories. Reference [5] provides advice on the specifications and use of coaxial connectors including guidance on the repeatability of the insertion loss of connector pairs.

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 can be found in reference [6].

F1 Reference weight calibration

F1.1 The uncertainties assigned to the values on a calibration certificate for the calibration of the reference weights are all contributors to the uncertainty budget.

F2 Secular stability of reference weights

F2.1 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.

F3 Weighing machine/weighing process

F3.1 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. readability;

(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.

F4 Air buoyancy effects

F4.1 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 (see reference [6]) 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 millibar to 1050 millibar.

F4.2 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.

F4.3 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.

F4.4 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.

F4.5 Apart from air buoyancy effects, the environment in which the calibration takes place can introduce 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.

G1 Reference thermometer calibration

G1.1 The uncertainties assigned to the values on a calibration certificate for the calibration of the reference thermometer are all contributors to the uncertainty budget.

G2 Measuring instruments

G2.1 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.

G3 Further influences

G3.1 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.

G4 Contributions associated with the thermometer to be calibrated

G4.1 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.

G4.2 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.

G5 Mathematical interpretation

G5.1 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.

H1 Reference standards and Instrumentation

H1.1 The uncertainties assigned to the reference standards and those for the measuring instruments used to make the measurements.

H2 Secular stability of reference standards and instrumentation

H2.1 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.

H3 Temperature effects

H3.1 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.

H4 Elastic compression

H4.1 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.

H5 Cosine errors

H5.1 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.

H6 Geometric errors

H6.1 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.

United Kingdom Accreditation Service, 21-47 High Street, Feltham, Middlesex, TW13 4UN
Website: www.ukas.com Publication requests Tel: +44 (0) 20 8917 8400 Fax: +44 (0) 20 8917 8500

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.

J1 Reference DWT

J1.1 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.

J2 Secular stability of the reference dead weight tester

J2.1 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. 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.

J2.2 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.

J3 Reference DWT mass set uncertainty

J3.1 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.

J4 Secular stability of the reference DWT mass set

J4.1 It is necessary to account for likely changes the mass set of the reference DWT since the last calibration. Paragraph F2.1 addresses the subject of secular stability of reference weights.

J5 Uncertainty of local gravity determination

J5.1 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.

J6 Air buoyancy effect

J6.1 Air buoyancy affects the mass set of a DWT in the same way as described in paragraph F4.

J7 Temperature effect on DWT area

J7.1 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 has to 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.

J8 Uncertainty due to head correction

J8.1 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 and 9 Pa/mm.

J8.2 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.

J9 Effects of fluid properties

J9.1 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.

J10 Non-verticality of the DWT piston

An uncertainty arises due to the fact that the piston may not be perfectly vertical. If it were, then all of 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.

J11 Uncertainties arising from the calibration process

J11.1 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 calculation or data fitting of the calibration results may also have to be considered.

APPENDIX K
EXAMPLES OF APPLICATION FOR CALIBRATION

NOTES

(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 particular measurements described.

(b) An uncertainty budget is a statement of a measurement uncertainty, of the components of that measurement uncertainty, and of their calculation and combination. It should include the measurement model, estimates, and measurement uncertainties associated with the quantities in the measurement model, covariances, type of applied probability density functions, degrees of freedom, type of evaluation of measurement uncertainty, and any coverage factor. An uncertainty budget is not simply a summary table; it has to 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.

K1 Calibration of a 10 kΩ resistor by voltage intercomparison

K1.1 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

$$R_X = \left(R_S + \delta R_D + \left(R_{TC} \cdot \Delta t\right) + s\left(\overline{V}\right)\right) \cdot \frac{V_X}{V_S}, \text{ where}$$

R_S	=	Calibration value for the standard resistor,
δR_D	=	Relative drift in R_S since the previous calibration,
R_{TC}	=	Relative 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,
$s\left(\overline{V}\right)$	=	Repeatability of ratio V_X/V_S.

K1.2 The calibration certificate for the standard resistor reported an uncertainty of ± 0.5 ppm at a coverage probability of approximately 95% (k = 2).

K1.3 A correction was made for the estimated drift in the value of R_S. The uncertainty in this correction, R_D, was estimated to have limits of ± 0.5 ppm.

K1.4 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.

K1.5 Records of evaluation of the oil bath characteristics showed that the maximum temperature deviation from the set point did not exceed ± 0.1ºC at any point within the bath.

K1.6 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.

K1.7 Type A evaluation: Five measurements were made to record the departure from unity in the ratio V_X / V_S in ppm. The readings were as follows:

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

From Equation (3), the mean value $\overline{V}$ = +10.5 ppm.

From Equations (5) and (6), $u(\overline{V}) = s(\overline{V}) = \frac{0.158}{\sqrt{5}}$ = 0.0707 ppm.

K1.8 Summary table

Symbol	Source of uncertainty	Value ±	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	∞
$u(\overline{V})$	Repeatability of indication	0.071 ppm	Normal	1	1	0.071	4
$u_c(y)$	Combined standard uncertainty		Normal			0.445	>500
U	Expanded uncertainty		Normal (k = 2)			0.891	>500

K1.9 Reported result

The measured value of the 10 kΩ resistor at 20 ºC ± 0.1º C was **10 000.105 Ω ± 0.89 ppm**

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 normally not included, as it is an unknown quantity. The relevant temperature conditions are therefore included in the reporting of the result.

K2 Calibration of a coaxial power sensor at a frequency of 18 GHz

K2.1 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 $\frac{\textit{Incident power at reference frequency}}{\textit{Incident power at calibration frequency}}$, for the same power sensor response and is determined from the following:

Calibration Factor, $K_X = (K_S + D_S) \times \delta DC \times \delta M \times \delta REF$, where

K_S = Calibration Factor of the standard sensor,
D_S = Drift in standard sensor since the previous calibration,
δDC = Ratio of DC voltage outputs,
δM = Ratio of Mismatch Losses,
δREF = Ratio of reference power source (short-term stability of 50 MHz reference).

K2.2 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.

K2.3 None of the uncertainty contributions are considered to be correlated to any significant extent.

K2.4 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 E9.6.

K2.5 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.

K2.6 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.

K2.7 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.

K2.8 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%.

From Equations (5) and (6), $u(K_R) = s(\overline{K}_X) = \frac{0.7415}{\sqrt{4}}$ = 0.3707%.

K2.9 Summary table

Symbol	Source of uncertainty	Value ± %	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	∞
D_S	Drift since last calibration	0.2	Rectangular	√3	1.0	0.116	∞
δDC	Instrumentation linearity	0.1	Normal	2.0	1.0	0.05	∞
δM	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_R	Repeatability of indication	0.37	Normal	1.0	1.0	0.37	3
$u(K_X)$	Combined standard uncertainty		Normal			1.69	>500
U	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 random components of uncertainty may arise due to the use of different connector pairs.

K3 Calibration of a 30 dB coaxial attenuator

K3.1 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 the following:

$A_X = A_b - A_a + A_{IF} + D_{IF} + L_M + R_D + M + A_L + A_R$, where

A_b = Indicated attenuation with unknown attenuator set to zero,
A_a = Indicated attenuation with unknown attenuator set to 30 dB,
A_{IF} = Calibration of reference IF attenuator,
D_{IF} = Drift in reference IF attenuator since last calibration,
L_M = Departure from linearity of mixer,
R_D = Error due to resolution of detection system,
M = Mismatch error,
A_L = Effect of signal leakage,
A_R = Repeatability.

K3.2 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).

K3.3 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.

K3.4 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/10dB. An uncertainty of ± 0.006 dB was therefore assigned at 30 dB. The probability distribution is assumed to be rectangular.

K3.5 The resolution of the detection system was estimated to cause possible rounding errors of 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.

K3.6 No correction is made for mismatch error. The mismatch uncertainty is calculated from the scattering coefficients using Equation E(5). The values used were as follows:

$\Gamma_L = 0.03$ $\Gamma_G = 0.03$ $s_{11a} = 0.05$ $s_{11b} = 0.09$ $s_{22a} = 0.05$ $s_{22b} = 0.01$ $s_{21a} = 1$ $s_{21b} = 0.031$

K3.7 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.

K3.8 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.

From Equations (5) and (6), $u(A_R) = s(\overline{A}_X) = \dfrac{0.018}{\sqrt{4}}$ = 0.009 dB.

K3.9 Summary table

Symbol	Source of uncertainty	Value ± dB	Probability distribution	Divisor	c_i	$u_i(A_X)$ dB	ν_i or ν_{eff}
A_{IF}	Calibration of reference attenuator	0.005	Normal	2.0	1.0	0.0025	∞
D_{IF}	Drift since last calibration	0.002	Rectangular	√3	1.0	0.0016	∞
L_M	Mixer non-linearity	0.006	Rectangular	√3	1.0	0.0035	∞
R_D	Resolution of indication	0.010	Triangular	√6	1.0	0.0041	∞
M	Mismatch	0.022	Normal	1	1.0	0.022	∞
A_L	Signal leakage effects	0.001	Rectangular	√3	1.0	0.0006	∞
A_R	Repeatability of indication	0.009	Normal	1.0	1.0	0.009	3
$u(A_X)$	Combined standard uncertainty		Normal			0.0246	>500
U	Expanded uncertainty		Normal (k = 2)			0.0491	>500

K3.10 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 E9.5.

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

K4 Calibration of a weight of nominal value 10 kg of OIML Class M1

K4.1 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:

Measured value of unknown weight, $W_X = W_S + D_S + \delta I_d + \delta C + W_R + A_b$, where

W_S = Weight of the standard,
D_S = Drift of standard since last calibration,
δI_d = The rounding of the value of the least significant digit of the indication,
δC = Difference in comparator readings,
W_R = Repeatability,
A_b = Correction for air buoyancy.

K4.2 The calibration certificate for the standard mass gives an uncertainty of 30 mg at a coverage probability of approximately 95% ($k = 2$).

K4.3 The 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.

K4.4 The least significant digit I_d for the mass comparator represents 10 mg. Digital rounding δI_d 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.

K4.5 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.

K4.6 A previous Type A evaluation of the repeatability of the measurement process, comprising 10 comparisons between standard and unknown, gave a standard deviation, $s(W_R)$, of 8.7 mg. This evaluation replicates the normal variation in positioning single weights on the comparator, and therefore includes effects due to eccentricity errors.

K4.7 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.

K4.8 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:	+ 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 10 000.005 g + 0.020 g = 10 000.025 g.

K4.9 Since three comparisons between standard and unknown were made (using 3 readings on the unknown weight), this is the value of *n* that is used to calculate the standard deviation of the mean:

From Equations (5) and (6), $u(W_R) = s(\overline{W_X}) = \frac{8.7}{\sqrt{3}}$ = 5.0 mg.

K4.10 Summary table

Symbol	Source of uncertainty	Value ± 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	∞
A_b	Air buoyancy (1 ppm of nominal value)	10.0	Rectangular	√3	1.0	5.77	∞
W_R	Repeatability of indication	5.0	Normal	1.0	1.0	5.0	9
$u(W_X)$	Combined standard uncertainty		Normal			24.55	>500
U	Expanded uncertainty		Normal (k = 2)			49.11	>500

K4.11 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).

K5 Calibration of a weighing machine of 205 g capacity by 0.1 mg digit

K5.1 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 weighings near full load. The span of the weighing machine has been adjusted using its internal weight before calibration. The following uncertainty calculation is carried out for a near full loading of 200 g. The machine indications are obtained from

Unknown indication $I_X = W_S + D_S + \delta I_{d0} + \delta I_d + A_b + I_R$, where

W_S = Weight of the standard,
D_S = Drift of standard since last calibration,
δI_{d0} = The rounding of the value of one digit at the zero reading,
δI_d = The rounding of the value of one digit of the indicated value,
A_b = Correction for air buoyancy,
I_R = Repeatability of the indication.

K5.2 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).

K5.3 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.

K5.4 It is often the case that when a weighing machine is zeroed, or tared, it does so to a greater resolution than that provided by the digital readout. This can be thought of as an “internal digit” that is not presented to the user. In this example, it is assumed that the “internal digit” size is one-tenth of that provided on the external display. The resulting possible rounding error is therefore ± 0.005 mg. The probability distribution is assumed to be rectangular.

NOTE

This is a good example of how a detailed knowledge of the principles of operation of a measuring instrument may be required in order to identify and quantify the associated uncertainties.

K5.5 No correction can be made for the rounding due to the resolution of the digital display of the machine. 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.

K5.6 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.

K5.7 The repeatability of the machine was established from a series of 10 readings (Type A evaluation), which gave a standard deviation, $s(W_R)$, of 0.05 mg. The degrees of freedom for this evaluation are 9, i.e. n - 1.

K5.8 Only one reading was taken to establish the weighing machine indication for each linearity and eccentricity point. For this calibration point the weighing machine indication, I_X , was 199.9999 g when the 200 g standard mass was applied. The value of n that is used to calculate the standard deviation of the mean of the indication, using the previously obtained repeatability standard deviation, $s(W_R)$, is therefore one:

From Equations (5) and (6), $u(I_R) = s(\overline{W_X}) = \frac{0.05}{\sqrt{1}}$ = 0.05 mg.

K5.9 Summary table

Symbol	Source of uncertainty	Value ± mg	Probability distribution	Divisor	c_i	$u_i(I_X)$ mg	ν_i or ν_{eff}
W_S	Calibration of standard weight	0.1	Normal	2.0	1.0	0.05	∞
D_S	Drift since last calibration	0.1	Rectangular	√3	1.0	0.058	∞
δI_{d0}	Digital rounding error (at zero)	0.005	Rectangular	√3	1.0	0.003	∞
δI_d	Digital rounding error (for indicated value)	0.05	Rectangular	√3	1.0	0.029	∞
A_b	Air buoyancy (1 ppm of nominal value)	0.2	Rectangular	√3	1.0	0.115	∞
I_R	Repeatability of indication	0.05	Normal	1.0	1.0	0.05	9
$u(I_X)$	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 indication of the weighing machine was **199.9999 g ± 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.

K6 Calibration of a Grade 2 gauge block of nominal length 10 mm

K6.1 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

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

L_S = Certified length of the standard gauge block at 20ºC
L_D = Drift with time of certified length of standard gauge block
δL = Measured difference in length
L = Nominal length of gauge block
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 20ºC when δL is determined
D_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
L_r = Repeatability of measurement

K6.2 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).

K6.3 The change in value L_D of the standard gauge block with time 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.

K6.4 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, δα, is ± 2 μm m^{-1} $ºC^{-1}$ with a triangular distribution. For L = 10 mm this corresponds to ± 20 nm/ºC. This difference will have two influences:

(a) The difference in temperature, δ*t*, 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 δ*T* between the mean temperature of the two gauge blocks and the reference temperature of 20 ºC was measured and was assigned limits of ± 0.2 ºC, giving rise to a length uncertainty of ± 4 nm.

As the influence of δα appears directly in both these uncertainty contributions they are considered to be correlated and, in accordance with paragraph D3.3, the corresponding uncertainties have been added before being combined with the remaining contributions. This is included in the uncertainty budget as $\delta_{Ts,x}$.

K6.5 The error due to the discrimination and non-linearity of the comparator, D_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.

K6.6 The variation in length of the unknown gauge block, $L_{V(x)}$, was considered to comprise two components: (i) 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 an uncertainty of ± 17 nm. (ii) 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. Quadrature combination of these contributions gives an uncertainty due to surface irregularities of ± 30 nm. As this is the combination of two rectangular distributions, of similar magnitude, the resulting distribution was assumed to be triangular.

K6.7 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 11 measurements and using Equation (7), yielded a standard deviation $s(L_R)$ of 16 nm.

K6.8 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 this previous value of standard deviation, with n = 1, because only one reading is made for the actual calibration.

From Equation (6), $u(L_R) = s(\bar{L}_X) = \dfrac{16}{\sqrt{1}}$ = 16 nm.

The measured length of the unknown gauge block was 9.999 94 mm.

K6.9 Summary table

Symbol	Source of uncertainty	Value ± nm	Probability distribution	Divisor	c_i	$u_i(L_X)$ nm	ν_i or ν_{eff}
L_S	Calibration of the standard gauge block	30	Normal	2.0	1.0	15.0	∞
L_D	Drift since last calibration	15	Triangular	√6	1.0	6.1	∞
D_C	Comparator	50	Rectangular	√3	1.0	28.9	∞
δC	Difference in elastic compression	5.0	Rectangular	√3	1.0	2.9	∞
$\delta T_{s,x}$	Temperature effects	5.6	Triangular	√6	1.0	2.3	∞
$L_{V(x)}$	Length variation of unknown gauge block	30	Triangular	√6	1.0	12.3	∞
L_r	Repeatability	16	Normal	1.0	1.0	16.0	10
$u(L_X)$	Combined standard uncertainty		Normal			39.0	>400
U	Expanded uncertainty		Normal (k = 2)			77.9	>400

K6.10 Reported result

The measured length of the gauge block was **9.999 94 mm ± 0.078 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.

K7 Calibration of a Type N thermocouple at 1000ºC

K7.1 A Type N thermocouple is calibrated against two reference standard Type R thermocouples in a horizontal furnace at a temperature of 1000ºC. The emfs generated by the thermocouples are measured using a digital microvoltmeter via a selector/reversing switch. All the thermocouples have their reference junctions at 0ºC. The unknown thermocouple is connected to the reference point using compensating cables.

K7.2 The temperature t_x of the hot junction of the unknown thermocouple is given by

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

$$\approx t_s \cdot V_{iS} + C_s \cdot \delta V_{iS1} + C_s \cdot \delta V_{iS2} + C_s \cdot \delta V_R - \frac{C_s}{C_{S0}}\delta t_{0S} + \delta t_D + \delta t_F$$

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

$$V_x(t) \approx V_x(t_x) + \frac{\Delta t}{C_x} - \frac{\delta t_{0X}}{C_{X0}} = V_{iX} + \delta V_{iX1} + \delta V_{iX2} + h_{TH} + \delta V_R + \frac{\Delta t}{C_X} - \frac{\delta t_{0X}}{C_{X0}}$$

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.
h_{TH}	=	Error due to inhomogeneity of the unknown thermocouple.

K7.3 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.

K7.4 The Type R 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 U is 0.3 ºC with a coverage factor k = 2.

K7.5 No correction is made for drift of the reference thermocouples since the last calibration but an uncertainty of ± 0.3ºC has been estimated from previous calibrations. A rectangular probability distribution has been assumed.

K7.6 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	C_S = 0.189 ºC/μV	C_{S0} = 0.077 ºC/μV
unknown	C_X = 0.038 ºC/μV	C_{X0} = 0.026 ºC/μV

K7.7 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.

K7.8 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.

K7.9 Residual parasitic offset voltages due to the switch contacts were estimated to be zero within ± 2.0 μV.

K7.10 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. - 0.407.

K7.11 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

K7.12 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.

K7.13 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.

K7.14 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:	1000.6 ºC		

K7.15 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

$$V_x = 36248 \times \frac{1000}{1000.5} \mu V = 36230 \mu V.$$

K7.16 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.

K7.17 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.

K7.18 In order to determine the random uncertainty associated with these measurements a Type A evaluation had been carried out on a previous occasion. A series of ten measurements had been undertaken at the same temperature of operation. Using Equation (7), this gave estimates of the standard deviations for the temperature of the furnace, $s_p(t_S)$ of 0.10 ºC and the voltage of the thermocouple to be calibrated, $s_p(V_{iX})$, of 1.6 μV.

The resulting standard uncertainties were as follows:

From Equation (6): $u(t_S) = s_p(\bar{t}_S) = \frac{s_p(t_S)}{\sqrt{n}} = \frac{0.1}{\sqrt{1}}$ = 0.10 ºC,

and $u(V_{iX}) = s_p(\bar{V}_{iX}) = \frac{s_p(V_{iX})}{\sqrt{n}} = \frac{1.6}{\sqrt{1}}$ = 1.6 μV.

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.

K7.19 Summary table - temperature of the furnace

Symbol	Source of uncertainty	Value ±	Probability distribution	Divisor	c_i	$u_i(T)$ ºC	ν_i or ν_{eff}
δt_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	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	∞
t_S	Repeatability	0.1 ºC	Normal	1.0	1.0	0.10	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	∞
$u_c(T)$	Combined standard uncertainty		Normal			0.641	>500

K7.20 Summary table - emf of unknown thermocouple

Symbol	Source of uncertainty	Value ±	Probability distribution	Divisor	c_i	$u_i(V)$ μV	ν_i or ν_{eff}
Δt_X	Furnace temperature	0.641 ºC	Normal	1.0	38.6	24.7	>500
δV_{LX}	Compensating leads	5.0 μV	Rectangular	√3	1.0	2.89	∞
δV_{iX1}	Voltmeter calibration	2.0 μV	Normal	2.0	1.0	1.0	∞
δV_R	Switch contacts	2.0 μV	Rectangular	√3	1.0	1.155	∞
δt_{0X}	Determination of reference point	0.1 ºC	Rectangular	√3	26.1	1.48	∞
$u(V_{iX})$	Repeatability	1.6 μV	Normal	1.0	1.0	1.6	9
δV_{iS2}	Voltmeter resolution	0.5 μV	Rectangular	√3	1.0	0.29	∞
h_{TH}	Inhomogeneity of thermocouple	0.3 ºC	Rectangular	√3	38.6	6.686	∞
$u_c(V)$	Combined standard uncertainty		Normal			25.9	>500
U	Expanded uncertainty		Normal (k = 2)			51.9	>500

K7.21 Reported result

The type N thermocouple shows, at the temperature of 1000.0 °C and with its cold junction at a temperature of 0 ºC, an emf of **36 230 μV ± 52 μV**.

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.

K8 Calibration of a Digital Pressure Indicator (DPI) at a nominal pressure of 2 MPa using a reference hydraulic dead weight tester

K8.1 The 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 indication of the unknown DPI is obtained from:

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

where

- I_X = Indication of unknown 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 level of standard and reference level of generated pressure
- B_V = Buoyancy volume of the reference piston – from calibration certificate.
- p_f = Density of hydraulic fluid
- σ = 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
- δI_d = The rounding of the value of one changing digit of the indication
- I_R = Repeatability of indication
- e_V = Error due to the piston not being perfectly vertical.

K8.2 The calibration certificate for the reference DWT gives the piston area and its uncertainty as:

A_0 = 80.6516 mm^2 ± 0.0026 mm^2. This results in a relative uncertainty in A_0 of 32.2 ppm.

K8.3 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.

K8.4 The drift limit in the effective area of the DWT, Ap_D, based on results from previous calibrations, has been set to ± 30 ppm.

K8.5 Mass uncertainties

K8.5.1 The mass of the piston is shown on the calibration certificate as 0.567 227 kg ± 0.000 010 kg.

K8.5.2 The drift of the piston mass, based on previous calibrations, has been set to 0.000 015 kg.

K8.5.3 The uncertainty of the mass set is shown on the calibration certificate for the three weights “A”, “B” and “C” used to generate 2 MPa as:

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

K8.5.4 The 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.

K8.5.5 To obtain the combined mass uncertainty, $u(m)$, in relative terms:

$$\text{Relative uncertainty in mass } u(m) = \frac{\text{Uncertainty of Piston} + \text{A} + \text{B} + \text{C (mg)}}{\text{Mass of Piston} + \text{A} + \text{B} + \text{C (kg)}} \text{ ppm}$$

$$= \frac{10 + 10 + 50 + 50}{0.567\,227 + 0.255\,242 + 7.402\,137 + 8.224\,784} = 7.30 \text{ ppm}$$

Similarly, for drift this gives m_D = 7.60 ppm.

K8.6 The uncertainty in the temperature of the piston, u_t, coming from the analysis of the temperature indictor used has been set as 0.5ºC. 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 figure was obtained from the calibration certificate for the DWT.

K8.7 A correction had been made for the value of the local acceleration due to gravity. This value had been estimated from knowledge of the measurement location and took the Bouguer anomalies into account. The expanded uncertainty associated with this estimate, U_g, was 3 ppm (k = 2).

K8.8 No correction has been made for air buoyancy therefore, as all masses have an assumed density of 7800 kgm^{-3}, the uncertainty, a_b, is estimated to be 13 ppm.

K8.9 The uncertainty relating to fluid head effects, f_h, arises from the height different between the reference level of the reference DWT and the generated pressure datum point. It is estimated as + 2 mm and the uncertainty in this estimate is 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 ms^{-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.

K8.10 The uncertainty contribution, f_e, from buoyancy volume, surface tension and fluid density effects has been estimated as 2 ppm based on a relative uncertainty in each of ± 10%.

K8.11 The repeatability of the calibration process had been established from previous measurements using DPIs of similar type and nominal range. This Type A evaluation, based upon 10 measurements and using Equation (7), yielded a relative standard deviation $s(P_R)$ of 16 ppm.

K8.12 An uncertainty arises due to the fact that the piston may not be perfectly vertical. If it were, then all of 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.

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.

K8.13 The calibration of the unknown DPI was established from a single measurement. As the conditions were the same as for the previous evaluation of repeatability the standard uncertainty due to repeatability can be obtained from this previous value of standard deviation, with n = 1, because only one reading is made for the actual calibration. Therefore $I_R = \dfrac{s(P_R)}{\sqrt{1}}$ = 16 ppm.

K8.14 No correction can be made for the rounding, δI_d, due to the resolution of the digital display of the DPI. The least significant digit on the range being calibrated of 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.

K8.15 Summary table

Symbol	Source of uncertainty	Value ±	Probability distribution	Divisor	c_i	$u_i(P_X)$ ppm	v_i or v_{eff}
A_0	Calibration of DWT (area)	32.2 ppm	Normal	2.0	1.0	16.25	∞
a_p	Calibration of DWT (distortion)	0.5 ppm/MPa	Normal	2.0	2.0	0.500	∞
Ap_D	Drift in area	30 ppm	Rectangular	√3	1.0	17.32	∞
$(u)m$	Calibration of total load	7.3 ppm	Normal	2	1.0	3.65	∞
m_D	Drift of total load	7.6 ppm	Rectangular	√3	1.0	4.38	∞
u_t	Temperature of the piston	0.5 ºC	Rectangular	√3	23	6.64	∞
a_b	Air buoyancy	13 ppm	Rectangular	√3	1.0	7.51	∞
U_g	Local gravity determination	3.0 ppm	Normal	2	1.0	1.50	∞
f_h	Fluid head effects	13.5 ppm	Rectangular	√3	1.0	7.79	∞
f_e	Other fluid effects	2.0 ppm	Rectangular	√3	1.0	1.15	∞
u_L	Levelling effects	3.4 ppm	Rectangular	√3	1.0	1.96	∞
δI_d	Digital rounding error	50 ppm	Rectangular	√3	1.0	28.9	∞
I_R	Repeatability of indication	16 ppm	Normal	1	1.0	16.0	9
$u(I_X)$	Combined standard uncertainty		Normal			43.0	>350
U	Expanded uncertainty		Normal (k = 2)			86.0	>350

K8.16 The generated pressure was calculated from the mass of the piston (K8.5.1), that of the mass set (K8.5.3) and the following quantities:

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

This results in an applied pressure of 2.000 806 MPa.

K8.17 Reported result

The pressure was applied in an increasing direction until it reached a final value of 2.000 806 MPa. The indication of the digital pressure indicator was **2.000 84 MPa ± 86 ppm**.

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 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 Gaussian, due to the presence of the other uncertainties, even though they are of smaller magnitude. This has been verified by Monte Carlo Simulation techniques.

2 The resolution uncertainty is based on the least significant digit of the DPI. However, this changes in steps of 2 digits for this particular DPI, therefore the digital rounding error is 1 digit.

3 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 particular 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 0.000 17 MPa.

APPENDIX L
EXPRESSION OF UNCERTAINTY FOR A RANGE OF VALUES

L1 Introduction

L1.1 On occasions it is convenient to provide a statement of uncertainty that describes a range of values rather than a single result.

L1.2 The GUM deals with expression of uncertainties for the reporting of a single value of a measurand, or more than one parameter derived from the same set of data. In practice many measuring instruments are calibrated at several points on a range and the use of an expression describing the uncertainty at any of these points can be desirable.

L1.3 This Appendix therefore describes the situations when this can occur, explains how it can be dealt with using the principles of this code of practice and provides an illustration of the process using a worked example.

L2 Principles

L2.1 When measurements are made over a range of values and the corresponding sources of uncertainty are examined it may be found that some are absolute in nature (i.e. they arise in a manner that is independent of the value of the measurand) and some are relative in nature (i.e. they arise in a manner that makes them proportional to the value of the measurand).

L2.2 It is possible, of course, to calculate a value for the expanded uncertainty for each reported value over the range. This can give problems when reporting values near zero as a relative term may not be appropriate and an absolute term has to be used. Conversely, when reporting values higher up the scale it may be desirable to express the uncertainty in relative terms, as this is often how instrument specifications are expressed.

L2.3 If the instrument being calibrated is subsequently to be used in a situation where a further analysis of uncertainty is required, the user may also require to express these uncertainties in both absolute and relative terms. However if the user has only been provided with a single value of uncertainty for each reported value, it would not be possible to extract the absolute and relative parts from these single values. Reporting uncertainties in both absolute and relative terms therefore provides more information to the user than if a series of single values are quoted. Additionally, it is often more representative of the way instrument specifications are expressed.

L2.4 The process of calculating a value of expanded uncertainty describing a range of values is identical to that for single values except that the absolute and relative terms are identified as such and, in effect, a separate uncertainty evaluation is carried out for each. These evaluations are carried out in the manner already described in this publication.

L2.5 The results of these evaluations are then expressed as separate absolute and relative terms. Traditionally this has been expressed in the form

$\pm (U_{REL} + U_{ABS})$

This linear addition of quantities is not in accordance with the principles embodied in the GUM, unless there happens to be a high degree of correlation between the absolute and relative terms - which is not usually the case. The two values should normally be reported separately with an appropriate statement describing how they should be combined. A suggested statement is given in the example in L3.13.

L3 Example of uncertainty evaluation for a range of values

L3.1 In this example a 6½-digit electronic multimeter is calibrated on its 1 V dc range using a multi-function calibrator.

L3.2 The calibrations were carried out in both polarities at 0.1 V increments from zero to 1 V and additionally at 1.5 V and 1.9 V. Only one measurement was carried out at each point and therefore reliance was placed on a previous evaluation of repeatability using similar multimeters.

L3.3 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.

L3.4 The indication of the multimeter under test, I_{DVM}, can be described as follows:

$I_{DVM} = V_{CAL} + V_D + V_{UE} + V_{TC} + V_{LIN} + V_T + V_{CM} + \delta V_{RES} + I_R$, where

V_{CAL}	=	Calibrated voltage setting of multifunction calibrator.
V_D	=	Drift in voltage of multifunction calibrator since last calibration.
V_{UE}	=	Uncorrected errors of multifunction calibrator.
V_{TC}	=	Temperature coefficient of multifunction calibrator.
V_{LIN}	=	Linearity and zero offset of multifunction calibrator.
V_T	=	Thermoelectric voltages generated at junctions of connecting leads, calibrator and multimeter.
V_{CM}	=	Effects on voltmeter reading due to imperfect common-mode rejection characteristics of the measurement system.
δV_{RES}	=	Rounding errors due to the resolution of multimeter being calibrated.
I_R	=	Repeatability of indication

L3.5 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 but there was an additional 0.5 μV in absolute units (k = 2).

L3.6 The manufacturer's 1-year performance specification for the calibrator included the following effects:

V_D , V_{UE} , V_{TC} These contributions were assumed to be relative in nature.

V_{LIN} This contribution was assumed to be absolute in nature.

L3.7 The specification for the calibrator on the 1 V dc range was ± 8 ppm of reading ± 1 ppm of full-scale. On this particular multifunction calibrator the full-scale value is twice the range value; therefore the absolute term is $\pm 2\ \text{V} \times 10^{-6} = \pm 2\ \mu\text{V}$. The performance of the calibrator had 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 was assumed.

L3.8 The effects of thermoelectric voltages, V_T, for the particular connecting leads used had 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 was assigned, based on previous experiments with the leads. The probability distribution was assumed to be rectangular.

L3.9 Effects due to common-mode signals, V_{CM}, had also been the subject of a previous evaluation and an uncertainty of 1 μV, with a rectangular distribution, was assigned. This contribution is absolute in nature, as the common-mode voltage is unrelated to the measured voltage.

L3.10 No correction is made for the rounding due to the resolution V_{RES} 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 δV_{RES} of ± 0.5 µV. The probability distribution is assumed to be rectangular and this term is absolute in nature.

L3.11 A previous evaluation had been carried out on the repeatability of the system using a similar voltmeter. Ten measurements were carried out at zero voltage, 1 V and 1.9 V. Repeatability at the zero scale point was found not to be significant compared with other absolute contributions. From Equation (7), this was found to give a standard deviation $s(V_i)$ of 2.5 ppm. Then, from Equation (6);

$$I_R = s_p(\bar{V}_i) = \frac{s(V_i)}{\sqrt{n}} = \frac{2.5}{\sqrt{1}} = 2.5 \text{ ppm}.$$

L3.12 Summary table

Symbol	Source of uncertainty	value (relative) ± ppm	value (absolute) ± µV	Probability distribution	Divisor	c_i	$u_i(V)$ (relative) ppm	$u_i(V)$ (absolute) µV	v_i or v_{eff}
V_{CAL}	Calibration uncertainty	2.8	0.5	Normal	2.0	1	1.40	0.25	∞
V_{SPEC}	Specification of calibrator	8.0	2.0	Rectangular	√3	1	4.62	1.15	∞
V_T	Thermoelectric voltages		1.0	Rectangular	√3	1		0.58	∞
V_{CM}	Common mode effects		1.0	Rectangular	√3	1		0.58	∞
δV_{RES}	Voltmeter resolution		0.5	Rectangular	√3	1		0.29	∞
I_R	Repeatability	2.5		Normal	1.0	1	2.5		9
$u_c(V)$	Combined standard uncertainty			Normal			5.42	1.46	>100
U	Expanded uncertainty			Normal (k = 2)			10.8	2.92	>100

L3.13 It is assumed that the results of this calibration will be presented in tabular form. After the results the following statements regarding uncertainty can be given:

The expanded uncertainty for the above measurements is stated in two parts:

Relative uncertainty: 11 ppm

Absolute uncertainty: 2.9 µV

The reported two-part expanded uncertainty is in each case 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. For each stated result the user may, if required, combine the uncertainties shown by quadrature summation in either relative or absolute terms as appropriate.

APPENDIX M
ASSESSMENT OF COMPLIANCE WITH SPECIFICATION

M1 Introduction

M1.1 In many situations it will be necessary to make a statement in calibration certificates or test reports about whether or not the reported result complies with a given specification.

M1.2 For calibration activities, this will often be the case for general purpose test and measurement equipment. For measurement standards it is more likely that the measured value and expanded uncertainty will be of more interest to the user, and specification compliance is less relevant.

M1.3 For testing activities it is very likely that the result will have to be compared with specified limits in order to arrive at a conclusion relating to conformance or fitness for purpose.

M2 Theory

M2.1 It has been shown elsewhere in this document that an uncertainty associated with a measurement result will usually take the form of a normal or Gaussian distribution centred around the reported value y of the measurand Y.

M2.2 If the result y and the entirety of the distribution lie within the specified limits then it is clear that compliance with the specification has been demonstrated. Conversely, If the result y and the entirety of the distribution lie outside the specified limits then non-compliance with the specification has been demonstrated.

M2.3 A “perfect” normal distribution has tails that extend to infinity in each direction, implying that there will always be some doubt about whether or not compliance has or has not been demonstrated. In practice, an uncertainty budget is always based on a finite number of contributions, thereby limiting the amount by which the tails of the distribution extend away from the reported value.

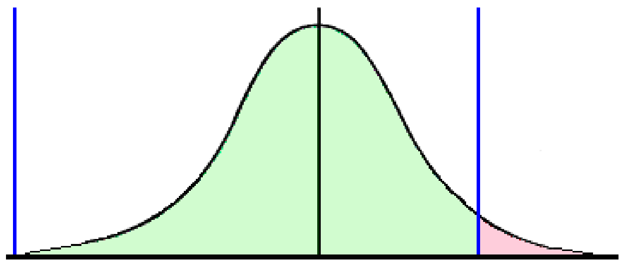
M2.4 Nevertheless, it can be the case that one, or both, of the tails of the distribution significantly overlap one, or both, of the specification limits. This means that consideration has to be given to the area of the distribution that is contained within the limits when assessing compliance with the specification.

M2.5 Furthermore, the implication of this is that compliance – or non-compliance – can only be stated in conjunction with an associated confidence level. This is because there will always be a possibility of one, or both, of the tails of the distribution overlapping the limits.

M2.6 This concept is illustrated in Figure 7 overleaf.

M2.7

Figure 7



In this example, it can be seen that the majority of the distribution lies within the specification limits, but a significant proportion lies above the upper limit. If, for example, 95% of the area of the distribution is within the specification limits, and 5% of it is outside them, then compliance has been demonstrated at a confidence level of 95%. Similar reasoning applies regarding non-compliance with the specification if the result were to be outside the limits.

M2.8 As an expanded uncertainty is normally expressed for a coverage probability of approximately 95%, it is generally accepted practice that statements regarding compliance will relate to the same level of confidence.

M2.9 If this is the case for a given situation, then comparison of the expanded uncertainty using a coverage factor $k = 2$ with the specification limit is unduly pessimistic. It will yield a confidence level of 97.5% or greater. This is because only one tail of the distribution will usually be the subject of comparison with the limit. If, as should be the case, the uncertainty is small compared with the specification, the probability contained within the other tail will already be within the specification limits.

M2.10 A new coverage factor, k_s, should therefore be used for the purpose of comparison with a specification limit. Assuming a normal distribution, the value of k_s required to achieve at least 95% confidence is 1.64.

M2.11 NOTE

It is common practice for a compliance decision to be made by simply comparing the expanded uncertainty, with a coverage probability of approximately 95%, with the specification. This will always yield a “safe” decision, however the method described above is preferred as it takes into account the actual amount of probability that breaches the specification limit.

M2.12 It may be the case that compliance, or non-compliance, with a specification cannot be demonstrated at 95% confidence, as described above. One solution is to reduce the uncertainty, possibly by applying corrections to the result, using more accurate equipment or by taking the mean of a larger amount of readings. If this is not practical or desirable, then it may be possible to evaluate compliance, or non-compliance, at a different level of confidence. Such an approach should, of course, be taken with the agreement and understanding of the customer.

M2.13 This procedure may also be used in cases where a customer has requested a compliance statement for a confidence level other than 95%.

M2.14 Two pieces of information are needed to deduce the confidence level at which compliance can be stated:

The combined standard uncertainty $u_c(y)$
The difference between the specification limit and the result, $L_S - y$

M2.15 The probability of compliance when the result lies within the specification, or that of non-compliance when it lies outside the specification, can be obtained from the following table:

M2.16

Probability of compliance or non-compliance (%)	$\left\lvert \frac{L_S - y}{u_c(y)} \right\rvert$	Probability of compliance or non-compliance (%)	$\left\lvert \frac{L_S - y}{u_c(y)} \right\rvert$	Probability of compliance or non-compliance (%)	$\left\lvert \frac{L_S - y}{u_c(y)} \right\rvert$
99.9	3.09	91	1.34	80	0.84
99.73	2.78	90	1.28	79	0.81
99	2.32	89	1.23	78	0.77
98	2.05	88	1.17	77	0.74
97	1.88	87	1.13	76	0.71
96	1.75	86	1.08	75	0.67
95.45	1.69	85	1.04	74	0.64
95	1.64	84	0.99	73	0.61
94	1.56	83	0.95	72	0.58
93	1.48	82	0.92	71	0.55
92	1.41	81	0.88	70	0.52

M2.17 Normally $\left\lvert \frac{L_S - y}{u_c(y)} \right\rvert$ will not be an integer and it will be necessary to interpolate between the values given in the table. Linear interpolation will suffice for $\left\lvert \frac{L_S - y}{u_c(y)} \right\rvert < 2$; higher-order interpolation should be used otherwise. Alternatively, the Excel function NORM.S.INV() will yield the correct value. If a simple approach is desired then the next lower value may be used.

M2.18 It should be noted that this procedure is only valid when the uncertainty breaches *one* of the specification limits and for this reason the uncertainty should be sufficiently small that an insignificant portion of the distribution approaches the other limit.

M2.19 Furthermore, as the result approaches either limit there will come a point at which no reasonable decision can be made regarding compliance or non-compliance with the specification. In the extreme, if the result coincided exactly with one of the limits, there would always be 50% confidence in the decision, regardless of the magnitude of the uncertainty. For this reason, and by general convention, the table above is limited to a confidence probability of 70% and above.

M2.20

Example

A measurement yields a result y of 0.80 units with a combined standard uncertainty $u_c(y)$ of 0.15 units. The specification is ± 1.00 unit. At what confidence level can compliance with the specification be made?

$\left\lvert \frac{L_S - y}{u_c(y)} \right\rvert = \left\lvert \frac{1.00 - 0.80}{0.15} \right\rvert = 1.33$. The next lower value in the table is 1.28, therefore it has been demonstrated that the specification is met for at least 90% confidence.

M2.21 It may be the case that a result cannot be demonstrated to comply with a specification for a given confidence level, but only a statement of the likelihood of compliance is required. An example of how this may be reported is given below:

M2.22 The measured result is below (above) the specification limit by a margin less than the measurement uncertainty; it is therefore not possible to state compliance (non-compliance) based on the stated coverage probability. However the result indicates that compliance (non-compliance) is more probable than non-compliance (compliance) with the specification limit.

M2.23 In cases such as those described above it is essential that the client is made aware of the situation because the end-user is taking some of the risk that the item may not meet the specification. It should also be noted that this "shared risk" approach may be superseded by legal requirements; for example, it is not permitted in some areas of legal metrology.

M3 Specification limits – further considerations

In Section M2 the specification limits have been treated as absolute, analogous to a rectangular probability distribution. This may not always be the case.

M3.1 There are situations where the specification is characterised by a normal distribution. It is stated in the GUM that, when a specification is quoted for a given coverage probability, then a normal distribution can be assumed. Some manufacturers state confidence levels for their specifications.

M3.2 If both the uncertainty U and the specification L are stated at the same coverage probability, then the specification is met at that coverage probability when $y < \sqrt{L_S^2 - U^2}$ and is failed when $y > \sqrt{L_S^2 + U^2}$, where

y = reported result
U = expanded uncertainty
L_S = specification limit

A worked example is shown below.

M3.3 <u>Example</u>

A digital multimeter is calibrated with an applied voltage of 10 .000 000 V dc. The expanded uncertainty, U, is 3 ppm (approximately 95% confidence, k = 2) and the reading is + 7 ppm from the nominal value.

The manufacturer's specification for this reading is stated as 10 ppm at 99% confidence. Is the specification met?

The comparison with specification has to be carried out with both the specification and the uncertainty at the same coverage probability. The specification at the 95.45% coverage probability, $L_{95.45}$, can be obtained from $L_{95.45} = L_{99} \cdot \frac{k_{95.45}}{k_{99}}$

From Paragraph 3.47, $k_{95.45}$ = 2 and k_{99} = 2.58, therefore $L_{95.45} = 10 \cdot \frac{2.0}{2.58}$ = 7.75 ppm.

$\sqrt{L_{95.45}^2 - U_{95.45}^2} = \sqrt{7.75^2 - 3^2}$ = 7.15 ppm. As 7 < 7.15, then compliance with the specification has been demonstrated, taking the measurement uncertainty into account.

M4 Reporting compliance with specification

M4.1 If compliance or non-compliance with a specification is clearly demonstrated for a given confidence level then a statement to this effect can be made in calibration certificates and test reports. However care must be taken to ensure that there is no implication that parameters that have not been measured also comply with a specification. For this reason a broad statement such as "the equipment complies with its specification" should not be made. A suggested statement of compliance is as follows:

The equipment complies with the stated specification at the measured points for the stated confidence level, due allowance having been made for the uncertainty of the measurements.

This statement can be modified as necessary where non-compliance with a specification is to be reported.

M4.2 When making compliance or non-compliance statements, the specification and any relevant clauses within it should be unambiguously identified in the calibration certificate or test report.

M4.3 There may be cases where the uncertainty attainable for a given test or calibration is larger than the specification for the item under consideration. Such situations should be subject to the contract arrangements between the laboratory and its customer. A statement regarding compliance should not be given under these circumstances but a note should be included indicating those uncertainties associated with the measured values that are greater than the required specifications.

United Kingdom Accreditation Service, 21-47 High Street, Feltham, Middlesex, TW13 4UN
Website: www.ukas.com Publication requests Tel: +44 (0) 20 8917 8400 Fax: +44 (0) 20 8917 8500

APPENDIX N
UNCERTAINTIES FOR TEST RESULTS

N1 Introduction

N1.1 ISO/IEC 17025:2005 requires that “testing laboratories shall have and apply procedures for estimating uncertainty of measurement”.

N1.2 It is recognised that the present state of development and application of uncertainties in testing activities is not as comprehensive as in the calibration fields, to which much of this document is addressed. It is therefore accepted that the implementation of ISO/IEC 17025:2005 criteria on this subject will take place at an appropriate pace, which may differ from one field to another. However laboratories should be able to satisfy requests from clients, or requirements of specifications, to provide statements of uncertainty.

N1.3 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.

N1.4 Some tests are qualitative in nature, i.e., they do not yield a numeric result. Therefore there can be no meaning in reporting uncertainties directly associated with the test result. Nevertheless, there will be uncertainties associated with the underlying test conditions and these should be subject to the same type of evaluation as is required for quantitative test results.

N1.5 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.

N2 Objectives

N2.1 The objective of a measurement is to determine the value of the measurand, ie the specific quantity subject to measurement. When applied to testing, the general term measurand may cover many different quantities, for example:

the electrical breakdown characteristics of an insulating material;

the strength of a material;

the concentration of an analyte;

the level of emissions of electromagnetic radiation from an appliance;

the quantity of micro-organisms in a food sample;

the susceptibility of an appliance to electric or magnetic fields;

the quantity of asbestos particles in a sample of air.

N2.2 A measurement begins with an appropriate specification of the measurand, the generic method of measurement and the specific detailed measurement procedure. Knowledge of the influence quantities involved for a given procedure is important so that the sources of uncertainty can be identified.

N3 Sources of uncertainty

N3.1 There are many possible sources of uncertainty. As these will depend on the nature of the tests involved, it is not possible to give detailed guidance here. However the following general points 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.

N3.2 These sources are not necessarily independent and, in addition, unrecognised systematic effects may exist that cannot be taken into account but contribute to error. This is one reason that participation in inter-laboratory comparisons, participation in proficiency testing schemes and internal cross-checking of results by different means are encouraged.

N3.3 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.

N3.4 Definitions are given in paragraph 3.10 for Type A evaluations and in paragraph 3.11 for Type B evaluations. Further detail on the means of evaluation is given in Sections 4 and 5.

N3.5 It is 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 17025:2005 requires that a reasonable estimation be made and that the form of the reporting does not give an incorrect impression of the uncertainty.

N3.6 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.

N3.7 Some analysis processes appear at first sight to be quite complex, for example there may be various stages of weighing, dilutions and processing before results are obtained. However 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 uncertainty of measurement can be evaluated from the resolution and repeatability of the process; matrix effects may also have to be considered.

N4 Process

N4.1 The process of assigning a value of uncertainty to a measurement result is summarised below:

(a) Identify all sources of error that are likely to have a significant effect, and their relationship with the measurand.

(b) Assign values to these using information such as described in N3.3, or in the case of Type A evaluations, calculate the standard deviation using Equations (5) and (6).

(c) Consider each uncertainty component and decide whether any are interrelated and whether a dominant component exists (see D3 and Appendix C respectively).

(d) Add any interdependent components algebraically (i.e., account for whether they act together or cancel each other) and derive a net value.

(e) Express each uncertainty value as the equivalent of a standard deviation (see paragraph 3.24), taking into account any sensitivity coefficients (see paragraphs 3.28 to 3.35).

(f) Take the independent components and the values of any derived net components and, in the absence of a dominant component, combine them by taking the square root of the sum of the squares. This gives the combined standard uncertainty (Equation (1)).

(g) Multiply the combined standard uncertainty by a coverage factor *k*, selected on the basis of the coverage probability required, to provide the expanded uncertainty *U* (see paragraphs 3.42 to 3.44).

(h) Report the result and, if required, the expanded uncertainty, coverage factor and coverage probability in accordance with Section 6.

N4.2 If one uncertainty contribution is significantly larger than the others then modifications may be required to this procedure. In the case of a dominant component derived from Type B evaluation, see Appendix C. If the non-repeatability of the system is significant, and its effects are evaluated by using a Type A analysis, it may be necessary to use the procedure in Appendix B.

N4.3 Further information regarding uncertainty evaluation for testing activities can be obtained from specialist publications that address particular fields of testing, such as are described in References [7] and [8].

APPENDIX P
ELECTRONIC DATA PROCESSING

P1 Introduction

P1.1 Due to the nature and quantity of the calculations involved it is inevitable that some form of electronic processing will be involved in these calculations. This Appendix gives brief details of precautions that may be necessary under these circumstances. Mention is also made of other techniques of uncertainty evaluation that electronic data processing has made practical.

P2 Use of calculators

P2.1 Most scientific calculators are easily capable of all the calculations required for the evaluation of measurement uncertainty. It is recommended that readers of this document gain familiarity with the functions involved by practicing the calculations presented in Appendix K - this can give rise to a better understanding of the process as well as giving the users confidence in their own abilities.

P2.2 It is also recommended that, where possible, intermediate results are stored in the calculator memory for use later, or are written down to a reasonable amount of significant figures, in order to prevent the cumulative effects of rounding errors having a significant effect on the result. Most scientific calculators work with sufficient accuracy so that they do not in themselves introduce any significant errors - with one notable exception (see Section P4).

P3 Use of spreadsheets

P3.1 The widespread use of personal computers has made repetitive calculations a much easier process than in the past. It is possible to construct a spreadsheet using the various equations presented in this document in order to perform the uncertainty calculations. The time spent constructing the spreadsheet can easily be recovered by the subsequent ease of producing, and amending, uncertainty budgets.

P3.2 It is recommended that a sample of results produced by a spreadsheet program is checked manually using a scientific calculator to ensure that correct results are being generated.

P4 Calculation of standard deviations

P4.1 Many scientific calculators include statistical functions for calculation of standard deviations in accordance with Equation 5. The calculator key associated with this function will usually be marked σ_{n-1} or, sometimes, *s*.

P4.2 It will often be the case that this particular function is not capable of evaluating small values of standard deviation correctly. The following data set is presented as an example:

1000.025 1000.015 1000.019 1000.021

P4.3 The value of *s* that is obtained from this data, using Equation 5, should be 0.004163. Most calculators will either:

(a) Display an error message, or

(b) Display a value of zero, or

(c) Display an incorrect non-zero value.

P4.4 A solution to this problem is to use only the few least significant digits in the set of data. So, for the data above, the numbers 0.025, 0.015, 0.019 and 0.021 can be entered, yielding the correct value of $s(q_j)$: 0.00416.

NOTE

This could equally have been evaluated using just the last two digits, i.e. 25, 15, 19 and 21. However it is useful to include a decimal point, as this will then appear in the correct place in the results thereby minimising the likelihood of errors being made.

P4.5 It is possible that electronic spreadsheets will also suffer from similar errors therefore suitable checks should be devised to evaluate any such effects.

P5 Other techniques

P5.1 The use of high speed data processing has opened new avenues for the analysis of uncertainty. One such technique is known as Monte Carlo simulation.

P5.2 In a Monte Carlo simulation the mathematical model of the underlying system is run over and over again, each time using a different set of random numbers representing the input variables. Each of these sets of random numbers combines via the model to represent a different outcome. Each run of the model is called a simulation, or trial, and at the end of each trial the outcome of the process is recorded. Each different outcome arises, through the measurement model, corresponding to a particular set of random numbers being applied to it. If the model is a good representation of the real-world system, then, by running a large enough number of trials (each with a different set of random numbers), the whole range of possible outputs can be produced; these form the distribution of the output.

P5.3 Sampling techniques, such as Monte Carlo simulation, provide an alternative approach to uncertainty evaluation in which the propagation of uncertainties is undertaken numerically rather than analytically. Such techniques are useful for validating the results returned by the application of the GUM, as well as in circumstances where the assumptions made by the GUM do not apply. In fact, these techniques are able to provide much richer information, by propagating the distributions (rather than just the uncertainties) of the inputs x_i through the measurement model f to provide the distribution of the output y. From the output distribution confidence intervals can be produced, as can other statistical information.

P5.4 Further information regarding Monte Carlo simulation techniques can be found in reference [10].

P5.5 Another method of uncertainty analysis is by the use of Bayesian statistics; this is based on the concept of *degree of belief*. Bayes' Theorem gives the 'posterior' odds on the correctness of a belief (given the new evidence that has just been observed), making use of the 'prior' odds that the belief is correct, i.e. an estimate of the plausibility of the belief before one has the new data. A concept of *likelihood ratio* allows the 'prior' odds to be adjusted according to the each piece of new evidence. Another iteration is then performed using the next piece of evidence, and so on until all the evidence is used and a probability associated with the belief is obtained.

P5.6 Bayes' Theorem is particularly useful for analysis of qualitative data, where the conventional GUM methodology is not appropriate. It can also be applied to the expression of professional opinions that may appear in test reports.

APPENDIX Q
SYMBOLS

The symbols used are taken mainly from the GUM. The meanings have also been described further in the text, usually where they first occur, but are summarised here for convenience of reference.

a_i	Estimated semi-range of uncorrelated systematic component of uncertainty, probability distributions unknown, where $i = 1$ ..... N.
a_d	A systematic component of uncertainty that so dominates other contributions to uncertainty in magnitude that special consideration has to be given to its presence in calculating the expanded uncertainty.
c_i	Sensitivity coefficient used to multiply the value x_i of an input quantity X_i to express it in terms of the measurand Y.
f	Functional relationship between the measurand Y and the input quantities X_i on which Y depends, and between output estimate y and input estimates x_i upon which y depends.
$\partial f / \partial x_i$	Partial derivative with respect to input quantity X_i of the functional relationship f between the measurand and the input quantities.
k	Coverage factor (general).
k_p	Coverage factor used to calculate an expanded uncertainty U_p for a specified coverage probability p.
k_s	Coverage factor chosen for the purpose of comparison with a specification limit.
m	Number of readings or observations that are used for the evaluation of $s(q_j)$, if different from n.
n	Number of readings or observations that contribute to a mean value.
N	Number of input estimates x_i on which the value of the measurand depends.
q_j	jth repeated observation of randomly varying quantity q. *See Note 1 overleaf.*
$\overline{q}$	Arithmetic mean or average of n repeated observations of randomly varying quantity q.
p	Coverage probability or level of confidence expressed in percentage terms or in the range zero to one.
σ	The standard deviation of a population of data using all the samples in that population.
$s(q_j)$	Estimate of the standard deviation σ of the population of values of a random variable q based on a limited sample of results from that population. *See Note 1 overleaf.*
$s(\overline{q})$	Experimental standard deviation of the mean value $\overline{q}$.
$t_p(\nu_{eff})$	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 .

$u_c(y)$ Combined standard uncertainty of output estimate y.

U Expanded uncertainty of output estimate y that describes the measurand as an interval $Y = y \pm U$, with a high coverage probability.

U_p Expanded uncertainty of output estimate y that describes the measurand as an interval $Y = y \pm U_p$, with a specified coverage probability p.

v Degrees of freedom; in general, the number of terms in a sum minus the number of constraints on the terms of the sum.

v_i Degrees of freedom of standard uncertainty $u(x_i)$ of input estimate x_i.

v_{eff} Effective degrees of freedom of $u_c(y)$ used to obtain $t_p(v_{eff})$

NOTE 1 The GUM uses the symbols q_k and $s(q_k)$ where q_j and $s(q_j)$ are used here. M3003 uses the subscript j instead of k in order to avoid any possible confusion with the coverage factor k.

APPENDIX R
REFERENCES

1 BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, JCGM 100:2008, *Evaluation of measurement data - Guide to the Expression of Uncertainty in Measurement.* Joint Committee for Guides in Metrology, First Edition, September 2008.

2 European cooperation for Accreditation, EA-4/02 *Expression of the Uncertainty of Measurement in Calibration*, December 1999.

3 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.

4 DIETRICH, C.F. *Uncertainty, Calibration and Probability: The Statistics of Scientific and Industrial Measurement Uncertainty,* Institute of Physics, January 1991.

5 National Physical Laboratory, *Guidance on using Precision Coaxial Connectors in Measurement*, Edition 3, August 2007.

6 United Kingdom Accreditation Service, *Calibration of Weighing Machines*, LAB14 Edition 4, November 2006.

7 United Kingdom Accreditation Service, *The Expression of Uncertainty in EMC Testing*, LAB34 Edition1, August 2002.

8 EURACHEM/CITAC Guide CG 4: *Quantifying Uncertainty in Analytical Measurement*, QUAM:2012.P1, Third Edition, 2012. ISBN 978-0-948926-30-3.

9 International Laboratory Accreditation Cooperation, *ILAC Policy for Uncertainty in Calibration*, ILAC-P14:12/2010

10 BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, 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.* Joint Committee for Guides in Metrology, First Edition, 2008.

United Kingdom Accreditation Service, 21-47 High Street, Feltham, Middlesex, TW13 4UN
Website: www.ukas.com Publication requests Tel: +44 (0) 20 8917 8400 Fax: +44 (0) 20 8917 8500