UKAS

UNITED KINGDOM ACCREDITATION SERVICE

The Expression of Uncertainty and Confidence in Measurement

M 3003

United Kingdom Accreditation Service • 21 – 47 High Street • Feltham • Middlesex TW13 4UN

The Expression of Uncertainty and Confidence in Measurement

Contents

About the United Kingdom Accreditation Service

The United Kingdom Accreditation Service (UKAS) is recognised by the UK Government as the national body responsible for assessing and accrediting the competence of organisations of measurement, testing, inspection and certification of systems, products and personnel.

About UKAS publications relating to laboratory accreditation

UKAS publications are categorised as follows:

▶ **M series**	publications containing general policy, requirements and guidance applicable to UKAS accredited calibration, testing and sampling;
▶ **NIS series**	publications containing specialised policy, requirements and guidance applicable to specific fields of UKAS accredited calibration, testing and sampling;
▶ **Technical Policy Statements**	brief announcements on matters of technical policy;
▶ **Directories**	classified listings of accredited organisations;
▶ **Publicity material**	leaflets descriptive of the UKAS service for accreditation of calibration, testing and sampling.

M 3003 ✱ Edition 1 ✱ December 1997
(Supersedes NIS 3003, Edition 8)
Reprinted with minor text corrections Jul 98, Nov 98, May 99, Feb 00

United Kingdom Accreditation Service
21 – 47 High Street
Feltham
Middlesex TW13 4UN
England

Tel: 020 8917 8555
Fax: 020 8917 8500

0 Introduction

0.1 The *NAMAS Accreditation Standard*, M10, requires accredited laboratories to produce estimates of uncertainty of its measurements for all calibrations using accepted methods of analysis. This requirement may be met by following the procedures described in this publication which supersedes NIS 3003 Edition 8: 1995 and is intended for application to all fields of measurement.

0.2 The requirements for estimation of measurement uncertainty apply to all results provided by calibration laboratories. They also apply to results produced by testing laboratories under the following circumstances:

(a) Where it is required by the client.

(b) Where it is required by the specification to which the test is carried out.

(c) Where the uncertainty is relevant to the application or validity of the result; eg where the uncertainty affects compliance to a specification or stated limit.

0.3 It is also a requirement that any laboratory, testing or calibration, which carries out internal calibrations in support of its accredited activities produces estimates of uncertainty for those internal calibrations. Further details can be found in Section 3 of the M10 Supplement, *Measurement and Calibration Systems*.

0.4 The need for an internationally accepted procedure for expressing measurement uncertainty led, in 1981, to the international authority in metrology, the Comité International des Poids et Mesures (CIPM), approving brief outline recommendations [1] submitted by a working group of representatives from the major national standards laboratories. The International Organisation for Standardisation (ISO) was then given the task of developing a detailed guide applicable to all levels of accuracy from fundamental research to shop floor operations. The responsibility for the preparation of such a comprehensive document for this broad spectrum of measurements was assigned to a working group of the ISO Technical Advisory Group on Metrology (ISO/TAG4/WG3) and led to publication of the ISO *Guide to the Expression of Uncertainty in Measurement* [2] (hereafter referred to as the *Guide*) in 1993. The *Guide* is now available as a BSI publication.

0.5 The uncertainty calculations given in this publication are consistent with the recommendations made in the *Guide* and with the European Accreditation of Laboratories (EAL) document R2 [3]. There are some areas where EAL R2 uses different terminology to that of the Guide. For example, EAL R2 uses the term "standard uncertainty of the output estimate" in place of "combined standard uncertainty" which is used in the *Guide*. Where such differences occur, this edition of M 3003 has been kept consistent with the *Guide* (see, however, the Note to Section 10). At the same time other changes have been made in this edition, as follows:

(a) To correct transcription errors that occurred in Edition 8 of NIS 3003.

(b) To improve the presentation and to broaden the document to include aspects of testing activities.

(c) Some of the uncertainty examples in Appendix H have been updated as a result of welcomed comments from the users of the document.

(d) Sections describing *Best Measurement Capability* and the use of calculators or spreadsheets have been added.

(e) The situation where uncertainty has to be expressed for a range of values has been addressed.

0.4 Appendices D, E, F and G in this publication provide details of common sources of error and uncertainty for a number of different fields of calibration and Appendix H provides examples of the application of M 3003 to the expression of uncertainty in these fields. Appendix I addresses the situation where an expression of uncertainty is required for a range of values; Appendix J deals with reporting "compliance with specification". An overview of the expression of uncertainties for testing activities is given in Appendix K. The use of calculators and spreadsheets for calculation of uncertainty is discussed in Appendix L.

0.5 As far as possible the terms and symbols used in this publication have been aligned with the *Guide*. A full list of symbols and their definitions is given in Section 10. Definitions of some of the general metrological terms are given in the *Guide*.

1 Overview

1.1 It has long been recognised that most measurements in calibration and testing work are subject to errors which are not perfectly quantifiable and that, therefore, there is uncertainty associated with the results of such measurements. A measurement result is therefore incomplete without a statement of the corresponding measurement uncertainty.

1.2 This section of M 3003 briefly examines methods of combining and expressing uncertainties and gives an overview of the approach taken herein, in the *Guide* [2] and in its derivatives such as EAL-R2 [3]. The detail of the methodology involved is presented in Sections 2 to 8 and Appendices B and C of this document.

1.3 It should be stated at this point that there is no fundamentally correct way of combining uncertainties. The *Guide* describes the accepted methods used by most laboratories and required by most accreditation bodies, but it should be remembered that it <u>is</u> a guide - a set of conventions, designed to produce a reasonably quantifiable uncertainty statement, mainly based on statistical techniques. We are trying to express the concept of impreciseness as consistently as possible!

1.4 By using a predetermined set of conventions, such as is presented in the *Guide*, accreditation bodies, laboratories and their clients are able to compare results from different sources in a meaningful manner. This approach also means that uncertainties, passed down from National Standards to the end user (often through a chain of several laboratories), are treated in a consistent and meaningful manner at each step of this process.

1.5 An easy and obvious way of combining uncertainty contributions would be simply to sum their limit values and quote the total as the overall uncertainty. This has certainly been the tradition in a number of fields of measurement but is no longer acceptable. Arithmetic summation has the advantages that it is quick, easy and generally yields a high level of confidence. In fact, at first sight, the use of arithmetic summation would appear to imply 100% confidence in the extent of the uncertainty associated with the measurement result.

1.6 Arithmetic summation does, however, have disadvantages. First, it is unlikely that, when errors from a number of different sources are combined, the sign and magnitude of the errors are such that they will all add up to the limit value produced by arithmetic summation. Arithmetic summation means, therefore, that an unduly pessimistic estimate of uncertainty would often be quoted.

1.7 Secondly, every experienced metrologist knows that if a measurement is repeated a number of times under the same conditions, a spread of values will be obtained. Furthermore, there will occasionally be a result that is so inconsistent with the others that the metrologist might ignore it, regarding it as an outlier or spurious result. But how far away does this outlier have to be from the mean value before it is ignored? This is where the experience of the metrologist comes in - he or she will, through knowledge and experience, have developed an almost intuitive "feel" for results that should be discounted. So, almost without thinking about it, the metrologist might impose statistical limits on the spread of values that are included in the overall uncertainty - and without any quantifiable justification for the extent of these limits.

1.8 It follows that if this intuitively truncated spread of values is added in with other sources of error, there is a small but finite possibility that the "true value" actually lies outside the stated uncertainties - in other words, the 100% confidence implied by arithmetic summation is not attainable. Using this approach, a statement of the confidence with which it is assumed that the "true value" does lie within the stated limits is not possible.

1.9 For these reasons, the *Guide* approaches the subject of measurement uncertainty using statistical methods. As with all methods of addressing this subject it is not perfect - it sometimes depends, for example, on imperfectly known probability distributions being assumed to have particular characteristics - but it enables reasonable values to be assigned to the overall uncertainty, complete with information regarding the confidence probability with which the uncertainty statement is associated.

1.10 It is now apparent that the subject of measurement uncertainty is complex. The *Guide* takes a logical approach to the subject and it is worth, at this point, summarising this approach:

(a) There is an input quantity to the measurement process, for example the imported calibration uncertainty, effects of ambient temperature or the repeatability of the process. This input quantity is represented by the symbol x. There will usually be more than one input quantity so the ith quantity can be represented by the symbol x_i.

(b) The standard uncertainty associated with the ith input quantity is represented by $u(x_i)$. The standard uncertainty is defined as one standard deviation and is derived from the uncertainty of the input quantity by dividing by a number associated with the assumed probability distribution. The divisors for the distributions most likely to be encountered are as follows:

Normal	1
Normal (k = 2)	2
Rectangular	$\sqrt{3}$
Triangular	$\sqrt{6}$
U-shaped	$\sqrt{2}$

(c) There is an output quantity for the measurement process - the estimated value of the measurand - and this is given the symbol y.

(d) In some cases the input quantity to the process may not be in the same units as the output quantity. For example, one contribution to a measurement of, say, a gauge block or a Weston cell, may be the effect of temperature. In these cases the input quantity is temperature, but the output quantity is dimensional or electrical. It is therefore necessary to introduce a sensitivity coefficient so that the output quantity (y), can be related to the input quantity (x_i). This sensitivity coefficient is referred to as c_i. The sensitivity coefficient is effectively a conversion factor from one unit to another.

(e) The relationship between the input quantity and the output quantity may not be linear. The partial derivative $\partial f/\partial x_i$ can be used to obtain the sensitivity coefficient and this is one of the reasons that mathematical modelling is used to describe measurement systems. In practice the derivation of the partial derivatives can be difficult and the effort involved is not always justified by the results obtained. A linear approximation such as the quotient $\Delta f/\Delta x_i$, where Δf is the change in f resulting from a change Δx_i in x_i, is often sufficient.

NOTE

See paragraph 2.12 for further definitions of f, x and y.

(f) Once the output quantities associated with each input quantity x_i have been obtained in the form of a standard uncertainty (one standard deviation) they are combined by taking the square root of the sum of the squares. This yields the combined standard uncertainty $u_c(y)$.

(g) The combined standard uncertainty, $u_c(y)$, is also in the form of one standard deviation. This does not give sufficient confidence, for most purposes, that the "true value" lies within the stated limits. For this reason $u_c(y)$ is then multiplied by a coverage factor, k, to provide an expanded uncertainty, U. For most purposes a value of 2 is used for k, yielding a confidence level of approximately 95%.

(h) The result of the measurement is then reported, normally in the form $y \pm U$. This statement is not complete without mention of the coverage factor used to obtain the expanded uncertainty. An indication of the confidence level obtained should also be included. Appropriate statements are described later in this document.

1.11 There are some exceptions to the general approach for the method of combining uncertainties described above. These are dealt with later in the document but can be summarised as follows:

(a) Where there are unreliable input quantities - for example, poor repeatability.

(b) Where there is a dominant systematic uncertainty contribution.

(c) Where there is correlation between two or more of the input quantities.

(d) Where the uncertainties are not bilateral, ie the positive uncertainty limit differs from the negative one.

2 Concepts

2.1 An expression of the result of a measurement is incomplete unless it includes a statement of the associated uncertainty. The uncertainty of a measurement result is a parameter that characterises the spread of the values that could reasonably be attributed to the measurand. It states the range of values within which the value of the measurand is estimated to lie within a stated level of confidence.

2.2 It is essential to distinguish the term 'error' (in a measurement result) from the term 'uncertainty'. Error is the measurement result minus the *true* value of the measurand. Whenever possible a correction equal and of opposite sign to an error is applied to the result. Because true values are never known exactly (else there is no need to make a measurement), corrections are always approximate and a residual error remains. The uncertainty in this residual error will contribute to the uncertainty of the reported result.

2.3 Given the meaning of 'error' it follows that uncertainty can also be defined as the range about zero in which the error is thought to lie. It can be characterised in terms of the spread of the probability distribution for the error. This error distribution may be derived from the observed random variation in results, from theoretical knowledge of the error mechanism, or in some other way. The CIPM recommends a standard deviation as a measure of the spread of the distribution (one 'standard deviation'), but this is not sufficient in most fields of measurement where a confidence interval needs to be defined. It can be said that the true value of the measurand (or the error) lies within the stated range with a certain level of confidence (eg 95% or 99%).

2.4 The basis for the treatment of uncertainty in this publication is the assumption that all uncertainty components can be treated in the same way irrespective of the nature of their associated errors. In particular, it is assumed that their associated error distributions can be combined using ordinary statistical procedures, whether they are fixed during a measurement process (systematic error) or vary randomly (random error).

2.5 Sometimes a normal (ie Gaussian) distribution will adequately describe an error. At other times, when information is lacking, it may be appropriate to model it with a rectangular distribution, assigning equal probabilities to values between extreme limits. It must be noted that there will sometimes be circumstances when this assumption would give rise to optimistically small uncertainty estimates, eg, when the distribution is 'U'-shaped; Appendix D9 gives an example. It may be pessimistic, such as when the distribution is trapezoidal. In cases of doubt the rectangular distribution can be assumed; this assumption should, however, always be recorded. Further information on the treatment of various distributions is given in the *Guide*. When a number of distributions of whatever form are combined it can be shown that, apart from in one exceptional case, the resulting probability distribution tends to the normal form in accordance with the Central Limit Theorem [6]. The importance of this is that it makes it possible to assign at least a minimum level of confidence in terms of probability to the total 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.

2.6 Faced with the task of identifying and evaluating uncertainties in any specific measurement process it is convenient to classify them in terms of their effect on that process. When a measurement is repeated a number of times under substantially the same conditions, then, provided the measurement process has sufficient sensitivity to resolve small differences, the results will not all be the same due to the cumulative effect of small independent random variables. It is because of the observable random effects that this indeterminacy is referred to as the random component of uncertainty.

2.7 Corrections for errors in a measurement process may have to be made to ensure the traceability of the mean value of a sample of results to the national standard. The residual errors in these corrections are systematic in their effect on the measurement process at the time of its use, and therefore the corresponding uncertainty components are sometimes classified as systematic by association.

2.8 In mentioning the traditional classification of describing components of uncertainty as 'random' or 'systematic', it has to be appreciated that this is only meant to apply to the specific measurement process under consideration. In a hierarchical national measurement system, as uncertainties are propagated down through the calibration laboratories in the system, the classification of an influence quantity as a random component of uncertainty at one echelon will change to being a systematic component of uncertainty at the next lower level. For example, a calibration laboratory undertaking a calibration of a reference standard for another laboratory in a lower echelon will report a single value of total measurement uncertainty that combines the calibration laboratory's random and systematic components of uncertainty for the measurand. When the lower echelon laboratory uses the calibrated standard, the total uncertainty of its value will then be a systematic component in its effect within the uncertainty budget of that laboratory's further measurements.

2.9 Because the nature of the effect of an uncertainty component can change the CIPM has advocated the grouping of uncertainty components according to the method used to estimate their numerical values:

Type A: those that are evaluated by statistical methods.
Type B: those that are evaluated by other means.

2.10 In paragraph 3.3.4 of the *Guide* it is said 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. Therefore any convention as to how they are classified does not affect the estimation of total uncertainty. But it should always be remembered that, in the present 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 some exceptions.

2.11 In general a measurement process can be regarded as having estimated input quantities, given the symbol x, which contribute to the estimated value of the measurand or output quantity, given the symbol y. Where, as in most cases, there are several input quantities these are represented by x_i and the standard uncertainty associated with the estimated value of each input quantity is represented by $u(x_i)$. Standard uncertainty and its evaluation are discussed in Sections 3 and 4.

2.12 The measurement process can usually be modelled by a functional relationship between the estimated input quantities and the output estimate in the form

$$y = f(x_1, x_2, \ldots\ldots, x_N). \quad (1)$$

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. In some cases the input quantities are not in the same units as the output quantity, as in the above example, and each input uncertainty will need to be multiplied by an appropriate factor (see section 5) before it is combined with the other uncertainties.

3 Type A evaluation of standard uncertainty

3.1 A Type A evaluation will normally be used to obtain a value for the repeatability or randomness of a measurement process exhibited on one particular occasion. 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

$$\bar{q} = \frac{1}{n}\sum_{j=1}^{n} q_j = \frac{q_1 + q_2 + q_3 + \ldots\ldots + q_n}{n}. \quad (2)$$

3.2 The spread in the results, ie the range, indicates the merit or repeatability of the measurement process and depends on the apparatus used, the method, and sometimes on the person making the measurement. A useful statistic is the standard deviation σ of the n values that comprise the sample, which is given by

$$\sigma = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(q_j - \bar{q})^2} \quad (3)$$

3.3 If further measurements are made, using the same experimental conditions as specified in para 3.2, then, for each sample of results considered, different values for the arithmetic mean and standard deviation will be obtained. For large values of n, these mean values approach a central limit value of a distribution of all possible values. This probability density distribution can often be assumed to have the normal form. In practice the measurement process may have limitations in response for large deviations from the mean value and this will cause the actual form of the distribution curve to be truncated to some extent.

3.4 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

$$s(q_j) = \sqrt{\frac{1}{(n-1)} \sum_{j=1}^{n} (q_j - \bar{q})^2} \ . \qquad (4)$$

It will be noted that this result differs from the standard deviation of the sample itself by the factor $1/(n\text{-}1)$ in place of $1/n$ under the square root sign, the difference becoming smaller as the number of measurements n is increased.

3.5 The estimated standard deviation of the *uncorrected* mean value of the measurand is given by

$$s(\bar{q}) = \frac{s(q_j)}{\sqrt{n}}. \qquad (5)$$

It may not always be practical to repeat the measurement many times during a calibration. In these cases a more reliable estimate of the standard deviation of a measurement system can be obtained from an earlier Type A evaluation, based on a larger number of readings. If a prior assessment of $s(q_j)$ is used then the value of n used in equation (5) to calculate the standard deviation of the mean is the number of repeat readings made for the calibration itself and not that used in equation (4) to obtain the estimated standard deviation.

3.6 Whenever possible at least two measurements should be made as part of the calibration procedure; however, it is acceptable for a single measurement to be made when it is known that the random contributions in the measurement, including those for the device being calibrated, are negligible. For some calibrations it may be desirable to make only one measurement on the device being calibrated, even though it is known to have imperfect repeatability, and to rely on a previous assessment of the repeatability of like 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, if only to establish that its effect can be ignored, and n in equation (5) is then 1.

NOTE: The degrees of freedom under these circumstances are related to the number of measurements used for the previous assessment and not to that for the calibration itself. Degrees of freedom are discussed in Appendix B.

3.7 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 the cause should be investigated before proceeding further.

3.8 Although no correction can be made for a random component of uncertainty, equation (5) shows the benefit of increasing the number of measurements even when using a previous good estimate for the standard deviation of the whole population of possible values. However, the benefit becomes progressively less as the number is increased, and it is usually not necessary to make more than about 10 measurements and often 4 measurements are sufficient, provided the guidelines mentioned in paragraph 7.3 for the required level of confidence are followed.

3.9 The above statistical analysis of measurement values is a Type A evaluation for a random component of uncertainty. However, 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 as described in paragraph 4.5 for the case of a Type B uncertainty when only upper and lower bounds can be assessed.

3.10 The term *standard uncertainty*, $u(x)$, is used for the uncertainty of the result of a measurement expressed as a standard deviation. As mentioned in 2.11, there will usually be more than one input quantity so any individual contribution is identified by a subscript or a unique symbol. Thus the standard uncertainty, $u(x_i)$, of an input quantity, x_i, evaluated by means of repeated measurements is obtained from

$$u(x_i) = s(\bar{q}), \tag{6}$$

where $s(\bar{q})$ is calculated in accordance with equation (5).

4 Type B evaluation of standard uncertainty

4.1 It is probable that systematic components of uncertainty, ie those that account for errors that remain constant while the measurement is made, will be obtained from Type B evaluations: the most important of these systematic components, for an instrument, will often be the uncertainties associated with the corrections to indicated values on a calibration certificate issued by a calibration laboratory in a higher echelon in the national calibration system. However, there can be, and usually are, other important contributions to systematic errors in measurement that arise in the instrument user's own laboratory. 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.

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

(a) the reported uncertainty for the reference standard and any drift or instability in its value or reading;

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

(c) the equipment being calibrated or measured, for example its resolution and any instability during calibration;

(d) the operational procedure;

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

More detailed guidance concerning sources of error and uncertainty is given in Appendices D, E, F and G for electrical calibrations, mass calibrations, temperature calibrations and dimensional calibrations, respectively.

4.3 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 occasion, to simplify the measurement process it may be preferable 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.

4.4 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 deviations 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 deviation can be obtained from a simple calculation. The distributions discussed below will be applicable in the majority of cases.

4.5 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, (See, however, Section 2.5.). Then, if a_i is the semi-range of variation the standard deviation, again referred to as the standard uncertainty, $u(x_i)$, is given by

$$u(x_i) = \frac{a_i}{\sqrt{3}}. \qquad (7)$$

The *Guide* gives some information on the derivation of this expression.

4.6 An uncertainty obtained from a calibration certificate where the level of confidence or a coverage factor (k) has been reported may, unless otherwise indicated, be treated as having a normal probability distribution and the standard uncertainty will be given by

$$u(x_i) = \frac{\text{expanded uncertainty}}{k}. \quad (8)$$

See paragraph 7.2 for a definition of expanded uncertainty. For example a calibration certificate issued by an accredited laboratory for an instrument will normally report an expanded uncertainty based on $k = 2$ and include a statement that relates this to a confidence level of approximately 95% (see Section 8.1). However, if only a level of confidence is given, for example "confidence probability not less than 95%", then a normal distribution can be assumed with a coverage factor $k = 2$. The following relationships apply, in the case of a normal distribution, to other levels of confidence:

$k = 2.58$ for 99% confidence
$k = 3$ for 99.7% confidence

If a value of k greater than 2 is given but the level of confidence is quoted as 95% (see Appendix B for details of when this will occur) then the value given for k must be used in equation (8). In this situation the effective degrees of freedom ν_{eff} will also need to be considered. This is also addressed in Appendix B.

4.7 When an instrument has been certified as conforming to specification then the uncertainty in the calibration should have been taken into account. Since it is not the normal practice of most instrument manufacturers to declare confidence levels for tolerances, rectangular probability distributions can be assumed (see, however, paragraph 2.5), ie

$$u(x_i) = \frac{\text{Tolerance limit}}{\sqrt{3}}$$

Note: If tolerance limits had, for example, been quoted with a level of confidence corresponding to 3 standard deviations of the manufacturer's production probability distribution then one would have taken the instrument's uncertainty contribution as

$$u(x_i) = \frac{\text{Tolerance limit}}{3}$$

4.8 In some cases an uncertainty may already be expressed as a standard uncertainty, eg when it is derived from a calibration performed on one of the component parts of a measurement system. Clearly, in this case, no further calculation is required.

5 Combined standard uncertainty

5.1 Once the standard uncertainties $u(x_i)$ of the input quantities x_i have been derived from both Type A and Type B evaluations, the standard uncertainty of the output quantity $y = f(x_1, x_2, \ldots\ldots, x_N)$, can be calculated as follows

$$u_c(y) = \sqrt{\sum_{i=1}^{N} c_i^2\, u^2(x_i)} \equiv \sqrt{\sum_{i=1}^{N} u_i^2(y)}\ , \qquad (9)$$

where c_i , a sensitivity coefficient, is the partial derivative $\partial f/\partial x_i$, or in some cases a known coefficient, such as the temperature coefficient of expansion and $u_i(y) \equiv |c_i|u(x_i)$. An example of the combined standard uncertainty is as follows

$$u_c(y) = \sqrt{[\frac{c_1 U_1}{k}]^2 + \frac{c_2^2 a_2^2 + c_3^2 a_3^2}{3} + c_4^2 u^2(x_4)}\ ,$$

where the error associated with U_1 has a normal probability distribution, a_2 and a_3 are limits with rectangular probability distributions, all obtained from Type B evaluation, and $u(x_4)$ is obtained from Type A evaluation.

5.2 The calculations required to obtain sensitivity coefficients by partial differentiation can be a lengthy process, particularly when there are many individual 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 quantity. This approach can also be used if f is known, but if f is not a straightforward function the determination of partial derivatives required is likely to be error-prone. In this approach the partial derivative $\partial f/\partial x_i$ can be replaced 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 carefully. It should be balanced between being sufficiently large to obtain adequate numerical accuracy (number of numerically significant figures) in Δf and sufficiently small to provide a mathematically sound approximation to the partial derivative.

5.3 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 in the same units as the output quantity and the partial derivatives will all be unity.

5.4 If the functional relationship is a product or quotient, ie 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 uncertainties, eg expressed in % or parts per million. The general form is $y = cx_1^{p_1}.x_2^{p_2}.....x_m^{p_m}$ where the exponents p_i are known positive or negative numbers. The standard uncertainty will then be given by

$$\frac{u_c(y)}{|y|} = \sqrt{\sum_{i=1}^{N}\left[\frac{p_i u(x_i)}{|x_i|}\right]^2}\,. \qquad (10)$$

This is of the same form as equation (9) but with the standard uncertainties expressed as relative values.

Some examples of the use of relative uncertainties are

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

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

$V = f(P,Z) = (P.Z)^{1/2}$ and $\frac{u(V)}{V} = \sqrt{\left[\frac{u(P)}{2P}\right]^2 + \left[\frac{u(Z)}{2Z}\right]^2}$.

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.

6 Correlated input quantities

6.1 The expressions given for the standard uncertainty of the output estimate, equations (9) and (10), 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, eg 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.

6.2 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.

6.3 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 (9) or (10). The *Guide* should be consulted for a more detailed approach to dealing with correlation based on the calculation of correlation coefficients.

6.4 An example of the treatment of correlated contributions is shown in paragraph H6.4.

7 Expanded uncertainty and level of confidence

7.1 In most fields of measurement there is a need for some statement of confidence that can be associated with a calculated total uncertainty. It is helpful in making valid comparisons of measurement results and in giving proper meaning to an uncertainty reported on a certificate in terms of probability that the reported value of a measurand with its associated (±) uncertainty provides a range of values that includes the true value. A further consideration is the choice of level of confidence, also referred to as coverage probability. Although the utmost confidence in a statement of total uncertainty for a measurement will always seem desirable, in a hierarchical national system of laboratories involving the propagation of uncertainties from one echelon to the next, this level of confidence is not possible for many measurements.

7.2 The *Guide* recognises the need for providing a level of confidence associated with an uncertainty and uses the term expanded uncertainty, U, which is obtained by multiplying the combined standard uncertainty by a coverage factor, k, thus

$$U = k\,u_c(y). \qquad (11)$$

7.3 UKAS, in line with EAL and generally accepted international practice, recommends that a coverage factor of $k = 2$ is used to calculate the expanded uncertainty. This value of k will give a level of confidence of approximately 95% (95.45%), assuming a normal distribution. However, if the random contribution to uncertainty is relatively large compared with other contributions and the number of repeat readings is small there is a possibility that the probability distribution will not be normal in form and a value of $k = 2$ will give a level of confidence of less than 95%. In these circumstances the procedure given in

Appendix B should be used to obtain a value for the coverage factor that maintains the level of confidence at approximately 95%. A criterion that can be used to determine whether or not to use the procedure in Appendix B is as follows:

Generally, if an uncertainty assessment involves only one Type A evaluation and the number of readings, n, is greater than 2 and the combined standard uncertainty is more than twice the Type A uncertainty then $k = 2$ will provide a coverage probability of approximately 95% and there is no need to use Appendix B to obtain a different value for the coverage factor.

Further detail on the justification for this criterion is given in paragraph B7.

This may be of particular relevance to testing laboratories where it may not be practical, or even possible, to carry out many repeat measurements. It might be noted that it was not considered necessary to use Appendix B for any of the calibration examples included in Appendix H.

7.4 It has been mentioned that the ideal level of confidence for all measurements would be the utmost level of confidence, that is approaching 100% probability. If each uncertainty contribution were based on a rectangular distribution, arithmetic summation of their ranges would yield such a level of confidence. However, this procedure leads to considerably increased values of total uncertainty with very small probabilities that the true values are actually near the range limits, when the input quantities are uncorrelated.

7.5 If a level of confidence of 95% is considered to be too low for a particular calibration then a higher coverage factor, eg $k = 3$, giving a level of confidence of approximately 99.7%, can be used.

7.6 A statement of confidence cannot in practice report a specific level of probability, such as 95%, as this requires a knowledge of the actual probability distribution for each quantity upon which the value of a measurand depends. Nevertheless the ability to report an approximate level of confidence does give a very valuable meaning to a measurement result.

7.7 In some circumstances the value of U calculated for a level of confidence of 95% using the procedure given in this document will be greater than the total uncertainty obtained by arithmetic summation and therefore represents an unrealistic result. This situation could occur where there is a dominant Type B contribution with a theoretical U-shaped or assumed rectangular probability distribution, in which case the procedure given in Appendix C should be followed to obtain the value of U.

8 Reporting of results

8.1 After the expanded uncertainty has been calculated for a level of confidence 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:

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

8.2 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:

"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".

8.3 The common use of the word "approximately" in this context has generated a considerable amount of debate about its specific meaning. For the purposes of this document "approximately" is interpreted as meaning "effectively" or "for most practical purposes".

8.4 In the special circumstances where a dominant Type B contribution occurs refer to Appendix C. If a two-part uncertainty is being reported, refer to Appendix I.

8.5 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.

8.6 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. Uncertainties should normally be rounded up to the appropriate number of figures but may be rounded down when this does not significantly reduce confidence in a measurement result.

8.7 The numerical value of the measurement result should normally be rounded to the least significant figure in the value of the expanded uncertainty quoted unless there are justifiable reasons for quoting more figures. The normal rules of rounding apply however if the rounding decreases the value of the uncertainty by more than 5% the rounded-up value should be used. Rounding of results and uncertainties should be carried out only at the final stages of the calculation in order to prevent cumulative rounding errors having a significant effect.

9 Step by step procedure for the determination 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 H4 in Appendix H. Although this example relates to a calibration activity, the process for testing activities is unchanged.

	General case	Example H4: Calibration of a weight of nominal value 10 kg of OIML Class M1
9.1	If possible determine the mathematical relationship between the input quantities and the output quantity: $y = f(x_1, x_2, \ldots\ldots, x_N)$. (1)	It will be assumed that the unknown weight, W_X, can be obtained from the following relationship: $W_X = W_S + D_S + \delta I_d + \delta C + Ab$.
9.2	Identify all corrections that have to be applied to the results of measurements of a quantity (measurand) for the 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 estimated, 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 H4: Calibration of a weight of nominal value 10 kg of OIML Class M1
9.3	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 evaluation, using equation (7) for assumed rectangular distributions: $u(x_i) = \frac{a_i}{\sqrt{3}},$ (7) or equation (8) for assumed normal distributions: $u(x_i) = \frac{expanded\ uncertainty}{k},$ (8) or refer to other references 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: $u(x_1) = u(W_s) = \frac{30}{2} = 15\,mg,$ $u(x_2) = u(D_s) = \frac{30}{\sqrt{3}} = 17.32\,mg,$ $u(x_3) = u(\delta C) = \frac{3}{\sqrt{3}} = 1.73\,mg,$ $u(x_4) = u(Ab) = \frac{10}{\sqrt{3}} = 5.77\,mg,$ $u(x_5) = u(\delta I_d) = \frac{10}{\sqrt{6}} = 4.08\,mg.$
9.4	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.	From previous knowledge of the measurement system it is known that there is a significant random component of uncertainty.
9.5	If a random component of uncertainty is significant make repeated measurements to obtain the mean from equation (2): $\bar{q} = \frac{1}{n}\sum_{j=1}^{n} q_j$ (2)	Three measurements were made of the difference between the unknown weight and the standard weight from which the mean difference was calculated: $\bar{W}_R = \frac{0.015 + 0.025 + 0.020}{3} = 0.020\,g.$

	General case	Example H4: Calibration of a weight of nominal value 10 kg of OIML Class M1
9.6	Either calculate the standard deviation of the mean value from equations (4) and (5): $$s(q_j) = \sqrt{\frac{1}{(n-1)}\sum_{j=1}^{n}(q-\bar{q})^2}\ , \qquad (4)$$ $$s(\bar{q}) = \frac{s(q_j)}{\sqrt{n}}\ , \qquad (5)$$ or refer to the results of previous repeatability measurements for a good estimate of $s(q_j)$ based on a larger number of readings.	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 (4), 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 (5): $$s(\bar{W}_R) = \frac{s(W_R)}{\sqrt{n}} = \frac{8.7}{\sqrt{3}} = 5.0\,mg.$$
9.7	Even when a random component of uncertainty is not significant, where possible always check the instrument indication at least once to minimise operator recording mistakes.	
9.8	Derive the standard uncertainty for the above Type A evaluation from equation (6): $$u(x_i) = s(\bar{q}). \qquad (6)$$	This is then the standard uncertainty, equation (6): $$u(x_6) = u(W_R) = s(\bar{W}_R) = 5.0\,mg.$$
9.9	Calculate the combined standard uncertainty for uncorrelated input quantities using equation (9) if absolute values are used: $$u_c(y) = \sqrt{\sum_{i=1}^{N} c_i^2\, u^2(x_i)} \equiv \sqrt{\sum_{i=1}^{N} u_i^2(y)}\ , \qquad (9)$$ where c_i is the partial derivative $\partial f/\partial x_i$, or a known sensitivity coefficient. Alternatively use equation (10) if the standard uncertainties are relative values: $$\frac{u_c(y)}{\lvert y\rvert} = \sqrt{\sum_{i=1}^{N}\left[\frac{p_i u(x_i)}{\lvert x_i\rvert}\right]^2}\ , \qquad (10)$$ 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, ie mg, 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).

	General case	Example H4: Calibration of a weight of nominal value 10 kg of OIML Class M1
9.10	If correlation is suspected use the guidance in paragraph 6.3 or refer to other referenced documents.	None of the input quantities is considered to be correlated to any significant extent; therefore equation (9) can be used to calculate the combined standard uncertainty: $u(W_X) = \sqrt{15^2 + 17.32^2 + 4.08^2 + 1.73^2 + 5.77^2 + 5.0^2}$ $= 24.55\ mg$.
9.11	Either calculate an expanded uncertainty from equation (11): $U = k\,u_c(y)$, (11) or, if there is a significant random contribution evaluated from a small number of readings (see section 7.3), use Appendix B to calculate a value for k_p and use this value to calculate the expanded uncertainty.	$U = 2 \times 24.55\,mg = 49.10\,mg$. Since $n > 2$ and $u(W_X)/u(W_R) > 2$ it was not considered necessary to use Appendix B to determine a value for k_p. In fact the effective degrees of freedom of $u(W_X)$ is greater than 5000 which gives a value for $k_{95} = 2.00$.
9.12	Report the expanded uncertainty in the value of the measurand in accordance with Section 8.	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 level of confidence of approximately 95%. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

10 Symbols

10.1 The symbols used are taken mainly from the *Guide*. The meanings have been given in the text, including the appendices, where they occur, but are repeated 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 total uncertainty.

c_i sensitivity coefficient used to multiply an input quantity x_i to express it in terms of the output quantity y.

f	Functional relationship between the measurand y and the input estimates x_i on 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 used to calculate expanded uncertainty U.
k_p	Coverage factor used to calculate an expanded uncertainty for a specified level of confidence p where a normal probability distribution cannot be assumed.
n	number of repeat readings or observations.
N	number of input estimates x_i on which the measurand depends.
p	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 n results from that population.
$s(\bar{q})$	Experimental standard deviation of arithmetic mean $\bar{q}$.
$t_p(\nu_{eff})$	Student t-factor for ν_{eff} degrees of freedom corresponding to a given 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 provides a confidence interval $Y = y \pm U$.
ν	Degrees of freedom (general).
ν_i	Degrees of freedom of standard uncertainty $u(x_i)$ of input estimate x_i.
ν_{eff}	Effective degrees of freedom of $u_c(y)$ used to obtain $t_p(\nu_{eff})$.
q_j	jth repeated observation of randomly varying quantity q.
$\bar{q}$	Arithmetic mean or average of n repeated observations of randomly varying quantity q.

x_i Estimate of input quantity X_i .

y Estimate of the measurand Y.

NOTE: The *Guide* uses the symbols q_k and $s(q_k)$ where q_j and $s(q_j)$ are used here. M 3003 uses the subscript j instead of k in order to avoid any possible confusion with the coverage factor k.

11 References

1 CIPM at its 70th meeting in 1981 [Recommendation 1 (C1-1981), published in Metrologia 18 (1982), p.44].

2 BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML. *Guide to the Expression of Uncertainty in Measurement.* International Organisation for Standardization, Geneva, Switzerland. ISBN 92-67-10188-9, First Edition 1993.

BSI Equivalent:

BSI PD 6461: 1995, *Vocabulary of metrology, Part 3. Guide to the Expression of Uncertainty in Measurement.* BSI ISBN 0 580 23482 7

3 EAL-R2, *Expression of uncertainty of measurement in calibration,* Edition 1, April 1997, EAL.

4 NIS 80, *Guide to the Expression of Uncertainty in Testing,* Edition 1, September 1994, UKAS.

5 NIS 81, *The Treatment of Uncertainty in EMC Measurements,* Edition 1 May 1994, UKAS.

6 DIETRICH, C.F. *Uncertainty, Calibration and Probability,* London, Adam Hilger, Edition 2 1990. Chapter 4 discusses the combination of rectangular probability distribution and compares the results with normal distributions of equal standard deviation. Chapter 7 discusses the general theory of uncertainty including the combination of a Student's 't' distribution and a normal distribution.

7 WELCH, B.L. Biometrika, 1947, Vol 34, p.28. and ASPIN, A.A. Biometrika, 1949, Vol 36, p.290.

8 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.

9 NIS 4303, *Coaxial connectors in radio frequency and microwave measurement,* Edition 1, December 1991, UKAS.

10 NIS 0416, *Guidance on Weighing in NAMAS Accredited Laboratories,* Edition 1, June 1990, UKAS.

11 KRAGTEN, J. *Calculating standard deviations and confidence intervals with a universally applicable spreadsheet technique,* Analyst, 119, 2161-2166, 1994.

Appendix A

Best measurement capability

A1 Best 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. The best measurement capability shall be calculated according to the procedures given in this document and shall normally be quoted as an expanded uncertainty using a coverage factor of $k = 2$. An accredited laboratory cannot report an uncertainty lower than their accredited best measurement capability but may report an equal or higher uncertainty. Since the magnitude of the uncertainty reported on a certificate of calibration will often depend on properties of the device being calibrated any definition of best measurement capability should not include uncertainties that are dependent on this device. It may also be the case that an accredited laboratory can achieve a particular uncertainty if conditions are optimum but cannot achieve this uncertainty routinely.

A2 In order to promote harmony between accredited laboratories and between accreditation bodies the EAL has adopted the following definition of best measurement capability: *"the smallest uncertainty of measurement a laboratory can achieve within its scope of accreditation, when performing more or less routine calibrations of nearly ideal measurement standards intended to define, realize, conserve or reproduce a unit of that quantity or one or more of its values, or when performing more or less routine calibrations of nearly ideal measuring instruments designed for the measurement of that quantity"*. In other words "best measurement capability" is the smallest uncertainty a laboratory can achieve when performing more or less routine calibrations on a nearly ideal device being calibrated.

A3 A nearly ideal 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 these devices that are considered to be nearly ideal will depend on the field of calibration but may include an instrument with very low random fluctuations, negligible temperature coefficient, very low voltage reflections coefficient etc. The uncertainty budget that is intended to demonstrate the best measurement uncertainty should still include contributions from the properties of the device being calibrated that are considered to be nearly ideal but the value of the uncertainty can be entered as zero or a negligible value. Where necessary the laboratory's schedule of accreditation will include a remark that describes the conditions under which the best measurement capability can be achieved.

A4 By "more or less routine calibrations" it is meant that the laboratory shall be able to achieve the stated capability in the normal work that it performs under its accreditation and, by implication, using the procedures, equipment and facilities that were the subject of the assessment. Where a lower uncertainty can be achieved by taking a large number of readings this should be considered when arriving at the budget for the best measurement capability and would therefore be within the "more or less routine" conditions.

A5 It is acceptable for a laboratory to be accredited for a measurement uncertainty that is larger than they can actually achieve, if this is requested. Under these circumstances the uncertainty will still be described in the laboratory's schedule as their best measurement capability and the laboratory will not be permitted to report a smaller uncertainty on accredited certificates. Clearly if the principles of this document are followed when constructing the uncertainty budget the resulting expanded uncertainty will reflect what the laboratory can actually achieve, however, this may be lower than the uncertainty the laboratory wishes to be accredited for and report on their certificates of calibration. A possible solution to this dilemma is to make a greater allowance for one of the contributions in the uncertainty budget, for example for drift or long term instability in a reference standard.

A6 In some cases the best measurement capability quoted in a laboratory's Schedule has to cover a two 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 quantity. In these cases the best measurement capability 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 expression. Guidance about the expression of uncertainty over a range of values is presented in Appendix I.

Appendix B

Deriving a coverage factor for unreliable input quantities

B1 Coverage factor

B1.1 In the majority of measurement situations it will be possible to evaluate Type B uncertainties with high reliability. Further, 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 level of confidence 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 Use of the t-distribution

B2.1 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 level of confidence being significantly less than 95% if a coverage factor of $k = 2$ is used. In these situations the value of k, or more strictly k_p where p is the confidence probability in percentage terms, eg 95, 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 level of confidence at approximately the required level p.

B3 Derivation of a value from the t-distribution

B3.1 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 *Guide* 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

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

B4 Degrees of freedom

B4.1 The degrees of freedom, ν_i, for contributions obtained from Type A evaluations is n-1, viz., the number of readings used to evaluate $\bar{q}$ minus 1. If a value for the standard deviation is obtained from a previous evaluation, as mentioned in section 3.5, then the degrees of freedom would be calculated from the number of readings used to make this assessment rather than number of readings made during the calibration. However, it is recommended that when the prior evaluation is carried out a sufficient number of readings are made in order to ensure that $\nu_{eff} > 30$, ie $k < 2.09$.

B4.2 It is normally possible to take the degrees of freedom, ν_i, of Type B contributions as infinite, that is, their value is known with a very high degree of certainty. If this is done and there is only one contribution obtained from a Type A evaluation, usually the repeatability of the process, then equation B(1) can be simplified - see the example at B6.

B4.3 However, if a Type B contribution is an expanded uncertainty based on a *t*-distribution rather than a normal distribution, as described in this Appendix, then this will be an example of a Type B uncertainty that does not have infinite degrees of freedom. For this eventuality the degrees of freedom will be that quoted on the calibration certificate, see paragraph 8.2, or it can be obtained from the *t*-distribution table below for the appropriate value of k_{95}.

B5 Use of the t-distribution tables

B5.1 Having obtained a value for ν_{eff} the *t*-distribution table is used to find a value of k_p. The table below gives some values for k_{95}, ie those appropriate to a level of confidence of 95%; values for other levels of confidence can be found in the *Guide*.

ν_{eff}	1	2	3	4	5	6	7	8	10	12	14	16
k_{95}	13.97	4.53	3.31	2.87	2.65	2.52	2.43	2.37	2.28	2.23	2.20	2.17

ν_{eff}	18	20	25	30	35	40	45	50	60	80	100	∞
k_{95}	2.15	2.13	2.11	2.09	2.07	2.06	2.06	2.05	2.04	2.03	2.02	2.00

NOTE: A coverage factor of $k = 2$ actually relates to a level of confidence of 95.45% for a normal distribution. For convenience this is approximated to 95% which relates to a coverage factor of $k = 1.96$. However, the difference is not generally significant since, in practice, the level of confidence is based on conservative assumptions and approximations to the *true* probability distributions. The values given in the table are for a level of confidence of 95.45%.

B5.2 Normally ν_{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 $\nu_{eff} > 3$; higher-order interpolation should be used otherwise. Alternatively, use the next lower value.

B5.3 The value of k_{95} obtained from the table is that required to calculate the expanded uncertainty, U_{95}, from

$$U_{95} = k_{95}\, u_c(y). \qquad \text{B(3)}$$

B6 Example

B6.1 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 from equation B(1)

$$\nu_{eff} = \frac{5.7^4}{\frac{3.5^4}{(4-1)} + 0 + 0 + 0 + 0 + 0} = \frac{5.7^4}{3.5^4} \times 3 = 21.1.$$

The value of ν_{eff} given in the above table immediately lower than 21.1 is 20 which gives a value for k_p of 2.13 and this is the value that should be used to calculate the expanded uncertainty.

B7 Criterion for use of this Appendix

B7.1 The criterion given in paragraph 7.3 to determine the need to use this Appendix is based on the conclusion that if $u_c(y)/s(\bar{q}) > 2$ and $n > 2$ and all the other contributions are assumed to have infinite degrees of freedom, then $\nu_{eff} \geq 30$, giving a value for k_p of less than 2.09, which can be approximated by $k = 2$.

Appendix C

Dominant systematic component of uncertainty

C1 Dominant component

C1.1 In some measurement processes there can be one component of uncertainty derived from 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 may be assumed that the arithmetic sum of these contributions would be for a level of confidence approaching 100%, there is a degree of pessimism in following the normal recommended procedure given by equations (9) or (10) and (11). Consequently special consideration needs to be given to the situation in which the calculated expanded uncertainty fails to meet the criterion

$$U \leq \text{Arithmetic sum of limit values of all contributions.} \qquad \text{C(1)}$$

C2 Rectangular and U-shaped distributions

C2.1 In most cases the criterion of equation C(1) will be met, but, for example, when making rf and microwave electrical measurements mismatch can be a dominant contribution. Because the probability distribution is U-shaped (see Appendix D) the standard uncertainty is $\frac{a_i}{\sqrt{2}}$ rather than $\frac{a_i}{\sqrt{3}}$ for a rectangular distribution of the same semi-range limit a_i. The increased standard uncertainty makes it less likely that the criterion of equation C(1) will be met. When the criterion is not met then the dominant contribution, a_d, should be extracted and a new value of the expanded uncertainty calculated as follows

$$U = a_d + U'. \qquad \text{C(2)}$$

where U' is calculated from the remaining contributions using equations (9) or (10) and (11).

C2.2 This situation can occur whenever there is a dominant Type B uncertainty where the divisor associated with the assumed probability distribution is less than the coverage factor used to calculate U. This applies to the commonly assumed rectangular distribution but also to others such as U-shaped or triangular distributions.

C3 Derivation of confidence level

C3.1 When this situation has occurred the derivation of a confidence level with associated probability involves the convolution of probability distributions. As it would not be appropriate to attempt to present this in this publication, a level of confidence of 95%, based on a coverage factor of $k = 2$ can only apply to calculating a value for the term U' in equation C(2). In these special circumstances, the statement of confidence given in paragraph 8.1 is to be replaced by, for example:

"*the uncertainty quoted is dominated by the uncertainty due to the resolution of the instrument being calibrated for which a rectangular probability distribution has been assumed*".

or

"*the uncertainty quoted is dominated by the uncertainty due to mismatch for which a U-shaped probability distribution has been assumed*".

Appendix D

Some sources of error and uncertainty in electrical calibrations

The following is a list of the more common sources of systematic error and uncertainty (after correction) with brief comments about their nature. Further, more detailed, advice is given in other UKAS publications on calibration, the IEE Monograph Series of Titles, as well as other sources.

D1 Instrument calibration

D1.1 The uncertainties assigned to the values on a calibration certificate for the calibration of an instrument, whether measuring equipment or a reference standard.

D2 Secular stability

D2.1 The performance of all instruments 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 needs to be displayed in a graphical form and a curve fitting procedure that gives a progressively greater weight to each of the more recent calibrations has to be followed 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 periodicity of calibration.

D2.2 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.

D3 Measurement (or service) conditions

D3.1 If the laboratory measurement environment is different from that required for a calibration, then due allowance has to be made for any influence condition that could affect the measurement results and possibly determine the need for re-calibration. 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 also affect the values of unsealed mica capacitors and inductors. 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 precision components. Devices that incorporate thermal compensation, 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.

D3.2 It is also necessary to be aware of the possible effects of electrical operating conditions, such as power dissipated, 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.

D3.3 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.

D4 Interpolation of calibration data

D4.1 When an instrument with a broad range of measurement capabilities is calibrated there are practical and economic factors which limit the number of calibration points. Consequently the value of the quantity to be measured and 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.

D4.2 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 performance aberrations due to manufacturing imperfections.

D5 Resolution

D5.1 The limit to the ability of an instrument to respond to small changes in the quantity being measured, referred to as resolution, is treated as a systematic component of uncertainty. In a digital instrument it can be considered as ± ½ the least significant digit to which the instrument responds on the range in use, providing it can be assumed that there is no hysteresis with respect to the input quantity. In an analogue instrument it is determined by the practical ability to read the position of a pointer on a scale. The presence of electrical noise causing fluctuations in instrument readings will commonly determine the usable resolution.

D6 Lay-out of apparatus

D6.1 The physical lay-out 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.

D7 Thermal emfs

These are generated at junctions of dissimilar materials if there is a temperature difference; they are significant in dc measurements when low voltages are being measured. In ac/dc transfer measurements of voltage the polarity of the dc supply is reversed and an arithmetic mean is taken of two sets of dc measurements. Generally an allowance has to be made as a systematic component of uncertainty for the presence of thermal emfs.

D8 Loading and lead impedance

D8.1 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.

D8.2 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.

D8.3 For capacitance measurements, the inductive property 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.

D9 RF mismatch errors and uncertainty

D9.1 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, since the phases of voltage reflection coefficients are not usually known.

D9.2 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 by [8] to be related to the actual power, P_L, absorbed in a wattmeter terminating the line by the equation

$$P_O = \frac{P_L}{1 - |\Gamma_L|^2}\left(1 - 2|\Gamma_G||\Gamma_L|\cos\varphi + |\Gamma_G|^2|\Gamma_L|^2\right). \qquad D(1)$$

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

$$P_O = \frac{P_L}{1 - |\Gamma_L|^2}\left(1 - 2|\Gamma_G||\Gamma_L|\cos\varphi\right). \qquad D(2)$$

D9.3 When φ is unknown this expression for absorbed power can have limits

$$P_O(limits) = \frac{P_L}{1 - |\Gamma_L|^2}\left(1 \pm 2|\Gamma_G||\Gamma_L|\right). \qquad D(3)$$

D9.4 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 [8] show that the distribution is U-shaped with a standard deviation, given by

$$u(mismatch) = \frac{2|\Gamma_G||\Gamma_L|}{\sqrt{2}} = 1.414|\Gamma_G||\Gamma_L|. \qquad D(4)$$

D9.5 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 [8] have shown that the standard deviation of mismatch, M, expressed in dB is approximated by

$$M \approx \frac{8.686}{\sqrt{2}}[|\Gamma_G|^2(|s_{11a}|^2 + |s_{11b}|^2) + |\Gamma_L|^2(|s_{22a}|^2 + |s_{22b}|^2) + |\Gamma_G|^2 . |\Gamma_L|^2(|s_{21a}|^4 + |s_{21b}|^4)]^{0.5} \quad D(5)$$

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. In Reference [8] Harris and Warner concluded that the distribution for M will approximate to that of a normal distribution due to the combination of its component distributions. (Note: the above expression is not the same as that given in NIS 3003 Edition 8 which gave an expression for the limit values of mismatch. However, the above expression is considered to be more in keeping with the principles of the *Guide* and is therefore preferred.)

D9.6 The values Γ_G and Γ_L used in equation D(4) and D(5) and the scattering coefficients used in equation D(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 mismatch uncertainty is 0.036.

D10 Directivity

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; eg a directivity of 30 dB is equivalent to an uncertainty of ±0.0316 VRC. As with D9.6 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 $(0.016^2 + 0.01^2)^{0.5} = 0.019$ (34.4 dB).

D11 Test port match

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 D9. As with D9.6 and D10, the measured value of test port match will have an uncertainty which should be taken into account by using quadrature addition.

D12 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 the repeatability of particular connector pairs in use can be evaluated by connecting and disconnecting the device, 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 [9] provides advice on the specifications and use of coaxial connectors including guidance on the repeatability of the insertion loss of connector pairs.

Appendix E

Some sources of error and uncertainty in mass calibrations

The following is a list of 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.

E1 Reference weight calibration

E1.1 The reference weights will have uncertainties stated on the certificate of calibration issued by either a UKAS accredited laboratory, a national standards laboratory or other body acceptable to UKAS.

E2 Secular stability of reference weights

E2.1 It is also necessary to take into account the likely change in mass of the reference weights since their previous 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 (eg, ill-fitting screw knobs), surface finish, unstable adjustment material, physical wear and damage, 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 need to be based on the stability of the weights.

E3 Weighing machine/weighing process

E3.1 The performance of the weighing machine used for the calibration must be assessed to estimate the contribution it makes to the overall uncertainty of the weighing process. The performance assessment needs to 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 will need to include some or all of the following:

(a) repeatability of measurement;

(b) linearity within the range used;

(c) digit size/weight value per division, ie readability;

(d) eccentricity (off centre load), especially if groups of weights are placed on the weighing pan simultaneously; magnetic effects (eg magnetic weights, or the effect of force balance motors on cast iron weights);

(e) temperature effects, eg differences between the temperature of the weights and the weighing machine;

(f) length of arm error.

Guidance on the assessment of the above can be found in NAMAS publication NIS 0416 [10].

E4 Air buoyancy effects

E4.1 The accuracy with which air buoyancy corrections can be made depends on how well the density of the weights are known, and how well the air density can be determined. The density of weights can be determined by some laboratories, but for most mass work assumed figures are used. The air density is usually calculated from an equation (see NIS 0416 [10]) 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.

E4.2 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 as much as ±3 parts in 10^6 and for aluminium as much as ±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.

E4.3 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 which make use of stainless steel, plated brass, German silver or gunmetal weights this effect may be as much as ±16 parts in 10^6. For cast iron weights the figure may be as much as ±18 parts in 10^6 and for aluminium weights as much as ±45 parts in 10^6.

E4.4 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.

E5 Environment

E5.1 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 F

Some sources of error and uncertainty in temperature calibrations

The more common sources of systematic error and uncertainty (after correction) are listed below. Each source may have several uncertainty components.

F1 Calibration of reference thermometer

F1.1 The uncertainty assigned to the calibration of the reference thermometer(s). This will be stated on the certificate of calibration.

F2 Measuring instruments

F2.1 The uncertainty assigned to the calibration of any electrical or other instruments used in the measurements, eg standard resistors and digital multimeters.

F3 Further influences

F3.1 Additional uncertainties in the measurement of the temperature using the reference thermometers:

(a) Drift since the last calibration of instruments in F1 and F2;

(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, eg, calibration bath or furnace, and must include 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 must 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 eg, 'in air' or in stirred liquid.

F4 Contributions associated with the thermometer to be calibrated

F4.1 These may include the electrical factors in F2 above as well as some of the components listed in F3. When thermocouples are being calibrated any uncertainty introduced by compensating leads and reference junctions must be taken into account. Similarly any thermal emfs introduced by switches or scanner units should be investigated. When partial immersion liquid-in-glass thermometers are to be calibrated an additional uncertainty factor to account for effects arising from differences in depth of immersion should be included even when the emergent column temperature is measured.

F5 Mathematical interpretation

F5 Uncertainty arising from mathematical interpretation, eg in applying scale corrections or deviations from a reference table, or in curve-fitting to allow for scale non-linearity, should be assessed.

Appendix G

Some sources of error and uncertainty in dimensional calibrations

The following is a list of the more common sources of errors and uncertainties in dimensional measurements.

G1 Reference standards and Instrumentation

G1.1 The uncertainties assigned to the reference standards and those for the measuring instruments used to make the measurements.

G2 Thermal effects

G2.1 The uncertainties associated with differences in temperature between the gauge being calibrated and the reference standards and measuring instruments used. 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 thermometer itself.

G3 Elastic compression

G3.1 The uncertainties associated with differences in elastic compression between the materials from which the gauge being calibrated and the reference standards were manufactured. These are likely to be most significant in the more precise calibrations and in cases involving dissimilar materials, and 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 properties of the materials involved.

G4 Cosine errors

G4.1 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 which 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.

G5 Geometric errors

G5.1 Errors in the geometry of the gauge being calibrated, any reference standards used or critical features of the measuring instruments used to make the measurements can introduce additional uncertainties. Typically these will include small errors in the flatness or sphericity of stylus tips, the straightness, flatness, parallelism or squareness of surfaces used as datum features, and the roundness or taper in cylindrical gauges and reference standards. Such errors are often most significant in cases where perfect geometry has been wrongly assumed and where the measurement methods chosen do not capture, suppress or otherwise accommodate the geometric errors that prevail in a particular case.

Appendix H

Examples of application for calibration

NOTES

(i) The contributions and values given in the following examples are not intended to imply mandatory or preferred requirements. Laboratories should determine the uncertainty contributions for the particular measurement they are performing and report the estimated uncertainty on the certificate that is issued.

(ii) In order to gain familiarity with the principles set out in this document the reader may find it useful to examine these examples and repeat the calculations presented, referring to the cited equations as necessary. Although only calibration examples are presented in this edition of M 3003 the principles involved apply equally to testing activities, therefore staff of testing laboratories may also find this of benefit.

H1 Calibration of a 10 kΩ resistor by voltage intercomparison

H1.1 A high-resolution digital voltmeter is used to measure the voltage developed across a standard resistor and an unknown resistor of the same nominal value as the standard, when the series-connected resistors are supplied by from a constant current source. The value of the unknown resistor, R_X , is given by

$$R_X = (R_S + R_D + R_T) \frac{V_X}{V_S} ,$$

where R_S = Calibration value for the standard resistor,
R_D = Relative drift in R_S since the previous calibration,
R_T = Relative change in R_S due to the temperature of the oil bath,
V_X = Voltage across R_X ,
V_S = Voltage across R_S .

H1.2 The calibration certificate for the standard resistor reported an uncertainty of ±1.5 ppm at a level of confidence of not less than 95% (k = 2).

H1.3 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 ±2.0 ppm.

H1.4 The relative difference in resistance due to temperature variations in the oil bath was estimated to have limits of ±0.5 ppm.

H1.5 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.

H1.6 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 (2), Mean value $\bar{V} = +10.5\ ppm$

From equations (4) and (5) $u(V) = s(\bar{V}) = \dfrac{0.158}{\sqrt{5}} = 0.0706\ ppm$

H1.7 *Uncertainty budget*

Symbol	Source of uncertainty	value ±ppm	Probability distribution	Divisor	c_i	$u_i(R_X)$ ±ppm	ν_i or ν_{eff}
R_S	Calibration of standard resistor	1.5	normal	2.0	1.0	0.75	∞
R_D	Uncorrected drift since last calibration	2.0	rectangular	√3	1.0	1.155	∞
R_T	Effect of the temperature of oil bath	0.5	rectangular	√3	1.0	0.289	∞
V_S	Voltmeter across R_S	0.2	rectangular	√3	1.0	0.115	∞
V_X	Voltmeter across R_X	0.2	rectangular	√3	1.0	0.115	∞
V	Repeatability	0.071	normal	1.0	1.0	0.071	4
$u(R_X)$	Combined standard uncertainty		normal			1.418	>500
U	Expanded uncertainty		normal (k=2)			2.836	>500

H1.8 *Reported result*

Measured value of 10 kΩ resistor is: 10 000.11 Ω ± 0.03 Ω

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

NOTE: This example illustrates that even when the random component of uncertainty is observable it may not be significant.

H2 Calibration of a power sensor at a frequency of 18 GHz

H2.1 The measurement involves the calibration of an unknown power sensor against a standard power sensor by substitution on a stable, monitored source of known low source impedance. The measurement is made in terms of Calibration Factor, defined as

$$Calibration\ Factor = \frac{Incident\ power\ at\ reference\ frequency}{Incident\ power\ at\ calibration\ frequency}$$

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).

H2.2 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.

H2.3 None of the uncertainty contributions is considered to be correlated to any significant extent.

H2.4 Mismatch Uncertainties = 200 $\Gamma_G\Gamma_S$ % and 200 $\Gamma_G\Gamma_X$ %, 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 Appendix D9.6.

H2.5 The Standard Power Sensor was calibrated by an accredited laboratory 6 months before use, the expanded uncertainty (±1.1%) was quoted for a coverage factor k = 2.

H2.6 The long-term stability of the standard sensor was estimated from the results of 5 annual calibrations to have rectangular limits not greater than ±0.4% per year. A value of ±0.2% is assumed as the previous calibration was within 6 months.

H2.7 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.

H2.8 The measured results were as follows:

No.	Calibration Factor
1	93.45%
2	92.20%
3	93.95%
4	93.02%

From equation (2), Mean value $\bar{K}_X = 93.16\%$.

From equations (4) and (5) $u(K_R) = s(\bar{K}_X) = \dfrac{0.7415}{\sqrt{4}} = 0.3707\%$.

H2.9 *Uncertainty budget*

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	∞
δREF	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		√2	1.0	0.06	
M_3	Standard sensor at 18 GHz	1.40		√2	1.0	0.99	
M_4	Unknown sensor at 18 GHz	1.68		√2	1.0	1.19	
K_R	Repeatability	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

H2.10 *Reported result*

The Calibration Factor at 18 GHz is 93.2 % ± 3.4 %.

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a level of confidence 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 it may be necessary to increase the total uncertainty if the random component of uncertainty exceeds that during calibration due to the use of other connector pairs (see paragraph D12).

H3 Calibration of a 30 dB coaxial attenuator

H3.1 The measurement involves the calibration of a coaxial step attenuator at 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_a = Indicated attenuation with unknown attenuator set to zero
- A_b = Indicated attenuation with unknown attenuator set to 30 dB
- A_{IF} = Calibration of reference IF attenuator
- D_{IF} = Drift in reference 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

H3.2 The result is corrected for the calibration of the IF attenuator using the results from a calibration certificate, which gave an uncertainty (±0.01dB) at a level of confidence of 95% (k = 2).

H3.3 No correction is made for drift but the limits (±0.002dB) were estimated from the results of the previous three calibrations.

H3.4 No correction is made for mixer linearity; the uncertainty was estimated from a series of linearity measurements over the dynamic range of the system to be ±0.002 dB/10dB. This gives an uncertainty of ±0.006 dB at 30 dB for which the probability distribution is assumed to be rectangular.

H3.5 The resolution of the detection system was estimated to cause possible rounding errors of one-half of one least significant recorded digit ie ±0.005 dB.

H3.6 No correction is made for mismatch; the mismatch uncertainty (±0.022dB) is calculated using equation D(5), where $\Gamma_L = \Gamma_G = 0.03$, $s_{11a} = 0.05$, $s_{11b} = 0.09$, $s_{22a} = 0.05$, $s_{22b} = 0.01$, $s_{21a} = 1$ (0 dB) and $s_{21b} = 0.031$ (30 dB). These values include the uncertainty in the measurement of Γ_L , Γ_G , s_{11} and s_{22} - see paragraph D9.6.

H3.7 Four measurements were made which involved setting the reference level with the step attenuator set to zero and then measuring the 30 dB setting, the results were

No.	Attenuation
1	30.04 dB
2	30.07 dB
3	30.03 dB
4	30.06 dB

From equation (2), Mean value $\bar{A}_X = 30.05\,dB$.

From equations (4) and (5) $u(A_R) = s(\bar{A}_X) = \dfrac{0.018}{\sqrt{4}} = 0.009dB$.

H3.8 *Uncertainty budget*

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	0.01	normal	2.0	1.0	0.0050	∞
D_{IF}	Drift since last calibration	0.002	rectangular	√3	1.0	0.0012	∞
L_M	Mixer linearity	0.006	rectangular	√3	1.0	0.0035	∞
R_D	Resolution of detector system	0.005	rectangular	√3	1.0	0.0029	∞
M	Mismatch	0.022	normal	1	1.0	0.0220	∞
A_L	Leakage	0.001	rectangular	√3	1.0	0.0006	∞
A_R	Repeatability	0.009	normal	1.0	1.0	0.0090	3
$u(A_X)$	Combined standard uncertainty		normal			0.0248	>100
U	Expanded uncertainty		normal (k=2)			0.0495	>100

H3.9 *Reported result*

The measured value of the 30 dB attenuator at 10 GHz is 30.05 dB ± 0.05 dB.

The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor $k = 2$, providing a level of confidence 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: 0.1dB 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 relative values eg percentage etc.

(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; see Section 9 of Appendix D.

(3) In a subsequent use of an attenuator it may be necessary to increase the total uncertainty if the random component of uncertainty exceeds that during calibration due to the use of other connector pairs (see Appendix C, Section 9).

H4 Calibration of a weight of nominal value 10 kg of OIML Class M1

H4.1 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, $W_X = W_S + D_S + \delta I_d + \delta C + Ab$.

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,
Ab = Correction for air buoyancy.

H4.2 The calibration certificate for the standard mass gives an uncertainty of ±30 mg at a level of confidence of approximately 95%.

H4.3 The monitored drift limits for the standard mass have been set equal to the $k = 2$ (approximately 95% confidence level) uncertainty of its calibration, and are ±30 mg.

H4.4 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.

H4.5 The linearity error of the comparator over the 2.5 g range permitted by the laboratory's quality system for the comparison was estimated from previous measurements to have limits of ±3 mg.

H4.6 A previous Type A evaluation of the repeatability of the measurement process (10 comparisons between standard and unknown) gave a standard deviation, $s(W_R)$, of 8.7 mg. This test replicates the normal variation in positioning single weights on the comparator, and therefore includes effects due to eccentricity errors.

H4.7 No correction is made for air buoyancy, for which the uncertainty limits were estimated to be ±1 ppm of nominal value ie ±10 mg.

H4.8 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 + 0.02 g

mass of standard 10 000.005 g
calibration result 10 000.025 g

H4.9 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

$$u(W_R) = s(\bar{W}_R) = \frac{s(W_R)}{\sqrt{n}} = \frac{8.7}{\sqrt{3}} = 5.0\,mg.$$

H4.10 *Uncertainty budget*

Symbol	Source of uncertainty	value ±mg	Probability distribution	Divisor	c_i	$u_i(W_X)$ ±mg	v_i or v_{eff}
W_S	Calibration of standard weight	30.0	normal	2.0	1.0	15.0	∞
D_S	Uncorrected 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 linearity	3.0	rectangular	√3	1.0	1.73	∞
Ab	Air buoyancy (1 ppm of nominal value)	10.0	rectangular	√3	1.0	5.77	∞
W_R	Repeatability	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)			49.10	>500

H4.11 *Reported result*

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 level of confidence of approximately 95%. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

H5 Calibration of a weighing machine of 205 g capacity by 0.1 mg digit

H5.1 The calibration is carried out using weights of OIML Class E2. Tests 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, $I_X = W_S + D_S + \delta I_d + A_b + I_R$

where: W_S = Weight of the standard,
D_S = Drift of standard since last calibration,
δI_d = The rounding of the value of one digit of the indication,
A_b = Correction for air buoyancy,
I_R = Repeatability of the indication.

H5.2 The calibration certificate for the stainless steel 200 g standard mass gives an uncertainty of ±0.1 mg at a level of confidence of approximately 95% (k = 2).

H5.3 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.

H5.4 No correction is 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.

H5.5 No correction is made for air buoyancy. As the span of the weighing machine was adjusted with its internal weight before calibration, the uncertainty limits were estimated to be 1 ppm of the nominal value ie= 0.2 mg.

NOTE: For aluminium weights, a likely value for these uncertainty limits would be ±30 ppm of the nominal value.

H5.6 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 is 9, ie (n - 1).

H5.7 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, then

$$u(I_R) = s(\bar{W}_R) = \frac{s(W_R)}{\sqrt{n}} = \frac{0.05}{\sqrt{1}} = 0.05\ mg$$

H5.8 *Uncertainty budget*

Symbol	Source of uncertainty	value ± mg	Probability distribution	Divisor	c_i	$u(I_X)$ ± mg	ν_i or ν_{eff}
W_S	Calibration of standard weight	0.1	normal	2.0	1.0	0.05	∞
D_S	Uncorrected drift of standard weight since last calibration	0.1	rectangular	√3	1.0	0.058	∞
δI_d	Digital rounding error	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
$u_c(I_X)$	Combined standard uncertainty		normal			0.150	>500
U	Expanded uncertainty		normal (k=2)			0.300	>500

H5.9 *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 level of confidence of approximately 95%. The uncertainty evaluation has been carried out in accordance with UKAS requirements.

H6 Calibration of a Grade 2 gauge block of nominal length 10 mm

H6.1 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 20°C
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 20°C 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

H6.2 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).

H6.3 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.

H6.4 The thermal expansion coefficient applicable to each gauge was assumed to have a value, α, of 11.5 µm m^{-1} °C^{-1} with limits of ± 1 µm m^{-1} °C^{-1}. Combining these two rectangular distributions the difference in thermal expansion coefficient between the two blocks, $\delta\alpha$, is ± 2 µm m^{-1} °C^{-1} with a triangular distribution. For L = 10 mm this corresponds to ±20nm/°C. This difference will have two influences:

(a) The temperature difference δt between the two gauge blocks was estimated to be zero with limits of ±0.08°C, giving rise to a length uncertainty of ±1.6 nm.

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

As the influence of $\delta\alpha$ appears directly in both these uncertainty contributions they are considered to be correlated and, in accordance with paragraph 6.3, the corresponding uncertainties have been added before being combined with the remaining contributions. This is included in the uncertainty budget as $\delta_{Ts,x}$.

H6.5 The error due to discrimination and linearity of the comparator D_C was taken as zero with limits of ±0.05µm 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.

H6.6 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 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.

H6.7 The repeatability of the calibration process (L_R) was established from previous measurements using gauge blocks of similar type and nominal length. This Type A evaluation, based upon 11 measurements, yielded an experimental standard deviation $s(L_R)$ as follows:

From equation (4) $s(L_R) = 0.0160\,\mu m.$

H6.8 The calibration of the unknown gauge 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 equations (5) and (6)

$$u(L_R) = s(\bar{L}_R) = \frac{s(L_R)}{\sqrt{n}} = \frac{0.0160}{\sqrt{1}} = 0.0160\,\mu m.$$

The measured result for the unknown gauge block was 9.99994 mm.

H6.9 *Uncertainty budget*

Symbol	Source of uncertainty	value ± nm	Probability distribution	Divisor	c_i	$u_i(L_X)$ ±nm	v_i or v_{eff}
L_S	Calibration of the standard gauge block	30	normal	2.0	1.0	15.0	∞
D_L	Change in value of standard gauge block with time	15	triangular	√6	1.0	6.1	∞
D_C	Discrimination and linearity of comparator	50	rectangular	√3	1.0	28.9	∞
δC	Difference in elastic compression	5.0	rectangular	√3	1.0	2.9	∞
$\delta_{Ts,x}$	Temperature effects	5.6	triangular	√6	1.0	2.3	∞
$L_{V(x)}$	Variation in length of unknown gauge block	30	rectangular	√3	1.0	17.3	∞
L_R	Repeatability	16	normal	1	1.0	16.0	10
$u(L_X)$	Combined standard uncertainty		normal			40.8	>400
U	Expanded uncertainty		normal (k=2)			81.6	>400

H6.10 *Reported result*

The measured length of the gauge block is 9.99994 mm ±0.08 µm.

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

NOTE: This uncertainty value compares with the Class C uncertainties accepted by UKAS of ±0.08 µm for gauges of this size and is in line with the associated uncertainty of measurement of ±0.08 µm as given in BS 4311:1993.

H7 Calibration of a Type N thermocouple at 1000°C

H7.1 A Type N thermocouple is calibrated against two reference standard Type R thermocouples in a horizontal furnace at a temperature of 1000°C. 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 0°C. The test thermocouple is connected to the reference point using compensating cables.

H7.2 The temperature t_x of the hot junction of the test thermocouple is given by

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

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

The voltage $V_x(t)$ across the test thermocouple wires with the cold junction at 0 °C during the calibration is

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

where

$t_S(V)$ = Temperature of the reference thermometer in terms of voltage with the cold junction at 0 °C. 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} = Voltage corrections for rounding errors due to the resolution of the microvoltmeter,

δV_R = Voltage correction due to contact effects of the reversing switch,

δt_{0S} , δt_{0X} = Temperature corrections due to the reference temperatures,

C_S , C_X = Sensitivity coefficients of the thermocouples for voltage at the measuring temperature of 1000 °C,

C_{S0} , C_{X0} = Sensitivity coefficients of the thermocouples for voltage at the reference temperature of 0 °C,

δt_D = Drift of the reference thermometers since last calibration,

δt_F = Temperature correction due to non-uniformity of the furnace,

t = Temperature at which the test thermocouple is to be calibrated (calibration point),

$\Delta t = t - t_X$ = Deviation of the temperature of the calibration point from the temperature of the furnace,

δV_{LX} = Voltage correction due to the compensation leads.

H7.3 The reported result is the output emf of the test thermocouple at the temperature of the hot junction. Because 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 been split into two parts.

H7.4 The Type R reference thermocouples are supplied with calibration certificates that relate the temperature of their hot junctions with their cold junctions at 0 °C to the voltage across their wires. The expanded uncertainty U is ±0.3 °C with a coverage factor k = 2. No correction is made for drift since the last calibration but an uncertainty of ±0.3° C has been estimated from previous calibrations.

H7.5 The voltage sensitivity coefficients of the reference and test thermocouples have been obtained from reference tables as follows:

Thermocouple	1000 °C	0 °C
reference	C_S = 0.077 °C/µV	C_{S0} = 0.189 °C/µV
unknown	C_X = 0.026 °C/µV	C_{X0} = 0.039 °C/µV

H7.6 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} , due to resolution of ±0.5 µV for each indication.

H7.7 Corrections were made to the microvoltmeter readings by using data from the calibration certificate. Drift and other influences were considered negligible therefore only the calibration uncertainty U of ± 2.0 µV (k = 2) is to be included in the uncertainty budget.

H7.8 Residual parasitic offset voltages due to the switch contacts have been estimated to be zero within ±2 µV.

H7.9 The temperature of the reference point of each thermocouple is known to be 0 °C within ±1 °C.

H7.10 The temperature gradients inside the furnace had been measured. At 1000 °C deviations from non-uniformity of temperature in the region of measurement are within ±1 °C.

H7.11 The compensation leads had been tested in the range 0 °C to 40 °C. Voltage differences between the leads and the thermocouple wires were estimated to be less than 5 µV.

H7.12 The sequence of measurements is as follows: 1st Standard, Test Thermocouple, 2nd Standard, 2nd Standard, Test Thermocouple, 1st Standard. The polarity is then reversed and the sequence 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	1st standard thermocouple	Test thermocouple	2nd standard thermocouple
Voltage, after any correction for the digital voltmeter calibration	+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 µV
Temperature of hot junctions	1000.4°C		1000.6°C
Mean temperature of furnace	1000.5 °C		

H7.13 The reported thermocouple output emf will be corrected for the difference between the nominal temperature of 1000 °C and the measured temperature of 1000.5 °C. The reported thermocouple output will be

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

H7.14 In this example it is assumed that the procedure requires that the difference between the two standards must not exceed 0.3 °C. If this is the case then the measurement must be repeated and/or the reason for the difference investigated.

H7.15 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. 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 which gave pooled estimates of the standard deviation for the temperature of the furnace and the voltage of the thermocouple to be calibrated.

The resulting standard uncertainties were as follows.

From equations (4) and (5) $u(t_S) = s_p(\bar{t}_S) = \dfrac{s_p(t_S)}{\sqrt{n}} = \dfrac{0.1}{\sqrt{1}} = 0.10\,°C,$

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

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

H7.16 *Uncertainty budget (temperature of the furnace)*

Symbol	Source of uncertainty	value ±	Probability distribution	Divisor	c_i	$u_i(T)$ ±°C	ν_i or ν_{eff}
δt_S	Calibration of standard thermocouples	0.3 °C	normal	2.0	1.0	0.150	∞
δt_D	Drift in standard thermocouples	0.3 °C	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 °C	rectangular	√3	1.0	0.058	∞
t_S	Repeatability	0.1°C	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 °C	rectangular	√3	1.0	0.577	∞
$u_c(T)$	Combined standard uncertainty		normal			0.642	>500
U	Expanded uncertainty		normal (k=2)			not used	

H7.17 *Uncertainty budget (emf of test thermocouple)*

Symbol	Source of uncertainty	value ±	Probability distribution	Divisor	c_i	$u_i(V)$ ±μV	ν_i or ν_{eff}
Δt_X	Uncertainty of correction for furnace temperature (from previous uncertainty budget)	0.642 °C	normal	1.0	38.5	24.7	>500
δV_{LX}	Effects due to compensation 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 °C	rectangular	√3	25.6	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	∞
$u_c(T)$	Combined standard uncertainty		normal			25.0	>500
U	Expanded uncertainty		normal (k=2)			50.0	>500

H7.18 *Reported result*

The type N thermocouple shows, at the temperature of 1000.0 °C, with its cold junction at a temperature of 0 °C, an emf of 36 230 μV ± 50 μV.

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

Appendix I

Expression of uncertainty for a range of values

I1 Introduction

I1.0 On occasions it is convenient to provide a statement of uncertainty which describes a range of values rather than a single result.

I1.1 The *Guide* 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 readings. 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.

I1.2 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.

I2 Principles

I2.1 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 (ie can be expressed in the same units as the measurand) and some are relative in nature (ie can be expressed in terms of percentage, parts per million etc.).

I2.2 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.

I2.3 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 fractional terms. However if the user has only been provided with a single value of uncertainty for each reported value, it would be difficult 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, and additionally is often more representative of the way instrument specifications are expressed.

I2.4 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, effectively, a separate uncertainty evaluation is carried out for each. These evaluations are carried out in the manner already described in this publication.

I2.5 The results of these evaluations are then expressed as separate absolute and relative terms. Traditionally this has been expressed in the form

±(*a* [relative units] + *b* [absolute units]).

This linear addition of quantities is not in accordance with the principles embodied in the *Guide*, unless there happens to be a high degree of correlation between the absolute and relative terms - which is not usually the case, or one of the components is dominant (see Appendix C). 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 I3.

I3 Example of uncertainty evaluation for a range of values

I3.1 In this example a 6½-digit electronic multimeter is calibrated on its 1 V dc range using a multi-function calibrator.

I3.2 The calibrations were carried out 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.

I3.3 No corrections were made for known errors of the calibrator as these were identified as being small relative to other sources of uncertainty (see 4.3). An appropriate allowance for uncorrected errors has been therefore been included in the uncertainty evaluation.

I3.4 The reading on the multimeter under test, V_{DVM}, can be described as follows:

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

where V_{CAL} = 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.

I3.5 The calibration uncertainty was taken 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).

I3.6 The manufacturer's 1-year performance specification for the calibrator was deemed to include 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.

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

I3.7 The effects of thermoelectric voltages, V_T, for the particular connecting leads used had been evaluated on a previous occasion. This was considered to be an absolute uncertainty contribution and a value of ±1 µV was assigned, with a rectangular distribution.

I3.8 Effects due to common-mode signals, V_{CM}, had also been the subject of a previous evaluation and a value of ± 1 µV, with a rectangular distribution, was assigned. This contribution was considered to be absolute in nature.

I3.9 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.

I3.10 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. The value of $s(q_j)$ (Equation 4) was found to be 2.5 ppm for both the 1 V and 1.9 V scale points. As only one measurement is made when the calibration is carried out this is divided by $\sqrt{1}$ (Equation 5) to give a value of 2.5 ppm for $s(\bar{q})$.

I3.11 *Uncertainty budget*

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	N	2	1	1.4	0.25	∞
V_{SPEC}	1-year specification of multifunction calibrator	8.0	2.0	R	√3	1	4.6	1.15	∞
V_T	Thermoelectric voltages	-	1.0	R	√3	1	-	0.58	∞
V_{CM}	Effect of common-mode voltages	-	1.0	R	√3	1	-	0.58	∞
δV_{RES}	Rounding due to multimeter resolution	-	0.5	R	√3	1	-	0.29	∞
V_R	Repeatability	2.5	-	N	1	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

I3.12 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: ± 3 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 level of confidence 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 units as appropriate.

Appendix J

Statements of compliance with a specification

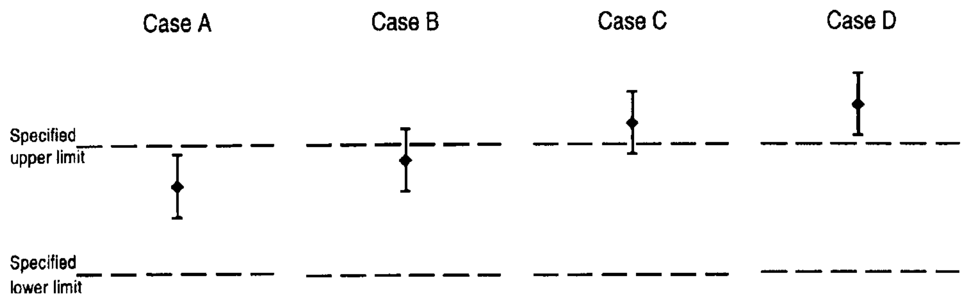
J1 Introduction

J1.1 In many situations it will be necessary to make a statement in calibration certificates or test reports as to whether or not the reported result complies with a given specification. 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 interest to the user.

J1.2 The need to report compliance with a specification may arise as the result of a request from a client or may be required by the standard or method describing the particular test or calibration.

J2 Assessment of compliance with a specification

J2.1 It will normally be necessary to consider the uncertainty of measurement when making a statement of compliance (or non-compliance) with a stated specification. The diagram below shows four cases of a result, with expanded uncertainty limits, relative to the specification under consideration.



J2.2 In Case A, the reported value (indicated by ♦), extended by the uncertainty of measurement, lies within the specification limits. A statement of compliance can therefore be made for the confidence level stated.

In Case B, the reported value lies within the specification limits. However the uncertainty overlaps one specification limit and therefore a statement of compliance cannot be made for the confidence level stated. The result does, however, mean that compliance with the specification is more likely than non-compliance.

In Case C, the reported value lies outside the specification limits. However the uncertainty overlaps one specification limit and therefore a statement of non-compliance cannot be made for the confidence level stated. The result does, however, mean that non-compliance with the specification is more likely than compliance.

In Case D, the reported value, extended by the uncertainty of measurement, lies outside the specification limits. A statement of non-compliance can therefore be made for the confidence level stated.

J2.3 In cases A and D above it is clear that a statement of compliance or non-compliance can be made because the uncertainty, at the given level of confidence, does not compromise such a statement. For cases B and C the situation is not so straightforward. Two suggested solutions are as follows:

(a) If possible use a more accurate method in order to reduce the uncertainty.

(b) Report the result and uncertainty with a statement that compliance (or non-compliance) could not be demonstrated. This could also give an indication regarding the likelihood of compliance (or non-compliance) being achieved. A suggested statement is as follows:

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 level of confidence. However the result indicates that compliance (non-compliance) is more probable than non-compliance (compliance) with the specification limit.

J2.4 In cases such as those described in J2.3(b) it is essential that the client is made aware of the situation because the end-user is taking some of the risk that the test item may not meet the specification.

J3 Reporting compliance with specification

J3.1 If compliance or non-compliance with a specification is clearly demonstrated, as in cases A and D above, then a statement to this effect can be made in certificates and 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" is not acceptable. A suggested statement of compliance is as follows:

The equipment complies with the stated specification at the measured points, 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.

J3.2 When making compliance or non-compliance statements, the specification and the relevant clauses within it shall be unambiguously identified in the calibration certificate or test report.

J3.3 There may be cases where the uncertainty attainable for a given test or calibration is larger than the specification for the item under consideration. A laboratory shall not normally contract to perform work on this basis; however reported uncertainties greater than the specification is permissible providing such measurements form a very small part of the overall measurements. A statement of compliance shall not be given under these circumstances but a note shall be included indicating those uncertainties associated with the measured values that are greater than the required specification.

J3.4 Specific requirements for reporting compliance with specification in a calibration certificate are given in NAMAS publication M12, *Conditions for reporting Calibration Results*.

Appendix K

Uncertainties for test results

K1 Requirements for uncertainties in testing

K1.1 It is a requirement for testing laboratories to quote a value of the uncertainty of their results under certain circumstances, as described in paragraph 0.2.

K1.2 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 UKAS policy 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.

K1.3 Testing laboratories are therefore required to have a defined policy covering the provision of estimates of the uncertainties of the tests performed. The laboratory should use documented procedures for the evaluation, treatment and reporting of the uncertainty.

K1.4 Testing laboratories should consult UKAS for any specific guidance that may be available for the testing field concerned.

K1.5 The methodology for estimation of uncertainty in testing is no different from that in calibration and therefore the procedures given in this document apply equally to testing results. It is recommended that readers new to this subject consult Section 1 of this document and UKAS document NIS 80, *Guide to the Expression of Uncertainties in Testing* [4], to gain familiarity with the concepts involved.

K2 Objectives

K2.1 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

K2.2 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.

K3 Sources of uncertainty

K3.1 There are many possible sources of uncertainty. As these will depend on the technical discipline 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, eg 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, eg when using a stopwatch.

(f) Instrument resolution or discrimination threshold, or errors in graduation of a scale.

(g) Values assigned to measurement standards (both reference and working) and reference materials.

(h) Changes in the characteristics or performance of a measuring instrument since the last calibration.

(i) Values of constants and other parameters used in data evaluation.

(j) Approximations and assumptions incorporated in the measurement method and procedure.

(k) 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 local environment, eg temperature, humidity and air pressure; variability in the performance of the person carrying out the test.

K3.2 These sources are not necessarily independent and, in addition, unrecognised systematic effects may exist that cannot be taken into account but contribute to error. It is for this reason that UKAS encourages - and sometimes insists on - participation in inter-laboratory comparisons, participation in proficiency testing schemes, and internal cross-checking of results by different means.

K3.3 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.

K3.4 Definitions are given in paragraph 2.9 for Type A and Type B evaluations and further detail on the means of evaluation is given in Sections 3 and 4.

K3.5 It is recognised that in certain areas of testing it may be known that a contribution to uncertainty exists but it is not possible to quantify it. In such cases, if a statement of uncertainty is required, this contribution can be omitted providing a note is associated with the uncertainty statement to this effect (see Section 9.2 of NIS80 [4]).

K3.6 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. Such instruments have to be shown to comply with the relevant specifications, normally by calibration, but where this specification has a minimal effect on the overall uncertainty then the expanded uncertainty can, if required, be evaluated using the statistical processes described in Section 3 and, where relevant, Appendix B of this document.

K3.7 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 most of the process can be considered to be negatively correlated (see paragraph 6.2) and the uncertainty of measurement can be evaluated from the resolution and repeatability of the process; matrix effects may also have to be considered.

K4 Process

K4.1 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.

(b) Assign values to these using information such as described in K3.3, or in the case of Type A evaluations, calculate the standard deviation using equations 4 and 5.

(c) Express each uncertainty value as the equivalent of a standard deviation (paragraphs 4.4 to 4.8).

(d) Consider each uncertainty component and decide whether any are interrelated and whether a dominant component exists (see Section 6 and Appendix C respectively).

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

(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 9).

(h) Multiply the combined standard uncertainty by a coverage factor (k), selected on the basis of the confidence level required, to provide the *expanded uncertainty* (U) (equation 11).

(g) Report the result, expanded uncertainty, coverage factor and confidence level in accordance with Section 8.

K4.2 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 repeatability of the system is significant, and its effects are considered by using a Type A evaluation, it may be necessary to use the procedure in Appendix B.

Appendix L

Use of calculators and spreadsheets

L1 Introduction

L1.1 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.

L2 Use of calculators

L2.1 Most scientific calculators are easily capable of all the calculations required for the estimation of measurement uncertainty. It is recommended that readers of this document gain familiarity with the functions involved by repeating the calculations presented in Appendix H; practise with this can give rise to a better understanding of the process as well as giving the users confidence in their own abilities.

L2.2 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 L4).

L3 Use of spreadsheets

L3.1 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.

L3.2 It is recommended that a sample of results produced by a spreadsheet programme is checked manually using a scientific calculator to ensure that correct results are being generated.

L3.3 Further information on the use of spreadsheets can be found in reference [11].

L4 Calculation of standard deviations

L4.1 Many scientific calculators include statistical functions for calculation of standard deviations in accordance with Equation 4. The calculator key associated with this function will usually be marked σ_{n-1} or, sometimes, s.

L4.2 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

The value of s that is obtained from this data, using Equation 4, should be 0.00416. Most calculators will either:

(a) Display an error message, or

(b) Display a value of zero, or

(c) Display an incorrect non-zero value.

L4.3 The reason that many calculators cannot evaluate these results correctly can be seen by examining the equation for the estimated standard deviation

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

The mean value $\bar{q}$ is subtracted from each of the observed values q_j of the quantity q. These differences are quite small numbers (5×10^{-6} relative to the 1000.025 value in the example above). Each is then squared; the example here gives a relative result of 25×10^{-12}. This is beyond - or at least comparable to - the internal resolution of most calculators and therefore errors can result.

L4.4 The 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
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, ie 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.

L4.5 It is possible that electronic spreadsheets will also suffer from this form of rounding error; therefore suitable checks should be devised to evaluate any such effects.