

# Principles of the EURACHEM/CITAC guide “Use of uncertainty information in compliance assessment”

Alex Williams

Received: 7 May 2008 / Accepted: 4 June 2008 / Published online: 9 July 2008  
© Springer-Verlag 2008

**Abstract** Many, possibly most, analytical measurements are carried out to assess compliance with a specification or a regulation, for example in the control of contaminants in food or the detection of performance enhancing substances in sport. When making an assessment of compliance the presence of unavoidable measurement uncertainty introduces the risk of making incorrect decisions, that is of accepting a batch of material which is outside the specification or rejecting one that is within. This often leads to controversy over whether or not the compliance decision is correct. How to make reliable assessment decisions is described in the EURACHEM/CITAC Guide “Use of uncertainty information in compliance assessment”. The key is the use of decision rules that lead to an unambiguous interpretation of the measurement result and its uncertainty. These decision rules need to be designed to ensure that requirements of the specification or regulation are met and that the risk of making an incorrect decision is acceptable. Ideally they should form part of the specification or regulation.

**Keywords** Measurement uncertainty · Assessment of compliance

---

Presented at the Measurement Uncertainty Symposium, April 2008, Berlin, Germany.

---

A. Williams (✉)  
Chairman EURACHEM/CITAC Working Group on  
Measurement Uncertainty and Traceability,  
19 Hamesmoor Way, Mytchett, Camberley,  
Surrey GU16 6JG, UK  
e-mail: aw@camberley.demon.co.uk;  
alex\_williams@camberley.demon.co.uk

## Introduction

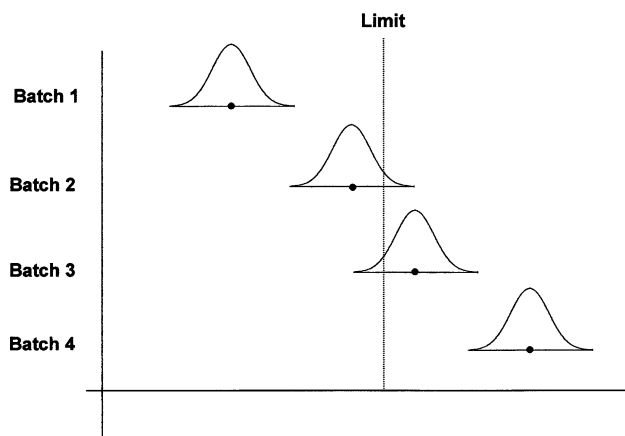
It is well known that when making an assessment of compliance, measurement uncertainty introduces the risk of making incorrect decisions, that is of accepting a batch of material which is outside the specification or rejecting one that is within. The probability of making a wrong decision depends both upon the size of the measurement uncertainty and on how the uncertainty is taken into account when assessing compliance.

Figure 1 shows the results of measurements of the concentration of an analyte in samples from different batches of a material. Superimposed on the value of the result and its uncertainty is a curve indicating the probability distribution of the likely values of the concentration. These results are to be used for an assessment of compliance with a specification that sets an upper limit  $L$  on the concentration.

The results show clearly that Batch 1 is within specification and that Batch 4 is outside, since in both cases the difference between the value of the result and the limit is much larger than the measurement uncertainty.

The result for Batch 2 shows that although the value of the concentration is likely to be below the limit there is a small, but depending upon the circumstances, perhaps a significant probability that it is above. Similarly for Batch 3 the value of the concentration is likely to be above the limit but the probability that it is below might be significant. Without further information on how the uncertainty should be taken into account, it is not possible to assess whether or not these batches are within specification.

When EURACHEM/CITAC gave the working group on Measurement Uncertainty and Traceability the task of writing guidance on this topic there was a strong



**Fig. 1** Compliance with an upper limit

recommendation that the guidance should cover how to deal unambiguously with cases such as Batches 2 & 3.

When we reviewed the literature on how to take into account measurement uncertainty in the assessment of compliance we found that a great deal of work had been done on the provision of guidance for the assessment of electrical and mechanical products and the guidance provided utilises the concept of “Decision Rules”. Rather surprisingly we found that this concept could be applied to analytical measurements and our Guide is based on the standard “Guidelines for Decision Rules: Considering Measurement Uncertainty in Determining Conformance to Specification [1] published by the American Society of Mechanical Engineers (ASME). A decision rule gives a prescription for the acceptance or rejection of a product based on the measurement result, its uncertainty and the specification limit or limits, taking into account the acceptable level of the probability of making a wrong decision. On the basis of the decision rule, an “Acceptance zone” or a “Rejection Zone” is determined, such that if the measurement result lies in the acceptance zone the product is accepted or if in the rejection zone it is rejected.

### Decision rules

Decision rules are designed to use the information about the measurement uncertainty to give an acceptable level of confidence of making a correct decision on compliance.

It is important to note that the uncertainty is a parameter that characterises the dispersion of the values that could reasonably be attributed to the measurand. Thus it is not a parameter that characterises the dispersion of the results but one that characterises the “degree of belief”, in the value attributed to the measurand: this is discussed by Kacker et al. [2].

Because of this uncertainty it is not possible to state definitely that the value of the measurand is, for example, above a certain limit. It is only possible to make a statement about the probability  $P$  that the value of the measurand is above this limit, based on the distribution of the values that could be attributed to the measurand. This is shown in Fig. 2, where  $P$  is the un-shaded area of the curve and is the probability that the value of the measurand exceeds the limit.

Thus for assessment against an upper limit the decision rule could be that the batch is to be judged to be in compliance if on the basis of the measurement result and its uncertainty the probability that the value of the measurand (i.e. the concentration) is above the limit is greater than  $P$ , where  $P$  is typically 0.95 or 0.99.

This might be the decision rule used when a strong case is required to reject the batch, for example, if a prosecution is to be brought for the breach of a regulation.

Thus depending on the value of  $P$ , using the above decision rule Batch 3 in Fig. 1 would be in compliance and Batch 4 would not.

In general the decision rules may be more complicated but the basic requirements for deciding whether or not to accept or reject a product are the same viz.:

1. A valid traceable result, with its uncertainty.
2. A specification giving the specific values of the characteristics (measurands) being controlled and the upper and/or lower bounds of their permissible values.
3. A decision rule that describes how the measurement uncertainty will be taken into account with regard to accepting or rejecting a product according to its specification and the result of a measurement.
4. The upper and lower bounds of the acceptance or rejection zone (i.e. the range of results) derived from the decision rule, which leads to the acceptance or rejection when the measurement result is within the appropriate zone.

Ideally the product specification or the regulation should also contain the decision rules, but unfortunately this is not often the case. Where the decision rules are not included in the specification then preferably they should be drawn up, in discussion with the customer for the analysis, as part of the definition of the analytical requirement. In any case the laboratory should always make clear the decision rules used in determining compliance.

A decision rule should have a well-documented method of determining the location of acceptance and rejection zones, ideally including the minimum acceptable level of the probability that the measurand lies within the specification limits. It may also give the procedure for dealing with repeated measurements and “outliers”. The laboratory will normally carry out the determination of the acceptance/rejection zones,

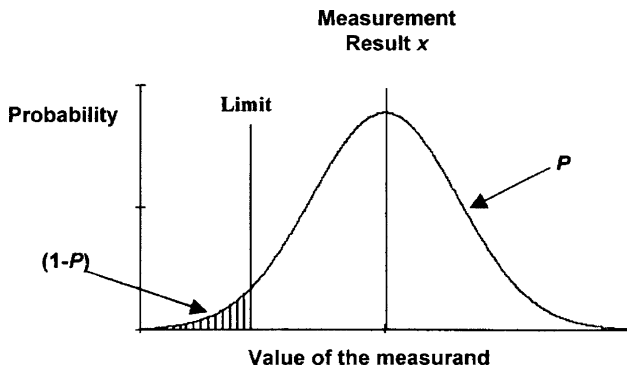


Fig. 2 Probability that the value is above the limit

based on the decision rule and the information available about the uncertainty.

An example of such a decision rule contained in a regulation is that given for implementing Directive 93/23/EC [3] viz.

1. The result of an analysis shall be considered non-compliant if the decision limit of the confirmatory method for the analyte is exceeded.
2. If a permitted limit has been established for a substance, the decision limit is the concentration above which it can be decided with a statistical certainty of  $1 - \alpha$  that the permitted limit has been truly exceeded.

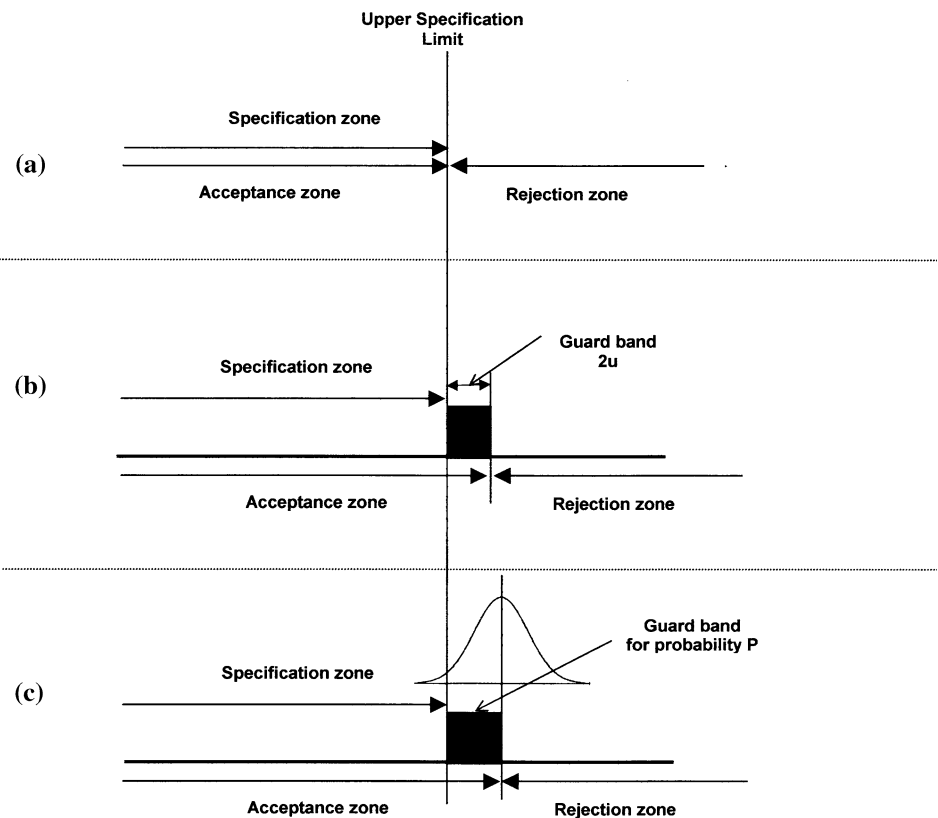
3. If no permitted limit has been established for a substance, the decision limit is the lowest concentration level at which a method can discriminate with a statistical certainty of  $1 - \alpha$  that the particular analyte is present.
4. For substances listed in Group A of Annex I to Directive 96/23/EC, the  $\alpha$  error shall be 1% or lower. For all other substances, the  $\alpha$  error shall be 5% or lower.

This is a decision rule for non-compliance or rejection. From this decision rule a rejection zone can be defined as shown in Fig. 3c. The start of the rejection zone is at the specification limit  $L$  plus an amount  $g$  (called the Guard band). The size of the guard band,  $g$ , is chosen so that for a measurement result of  $L + g$  there is a statistical certainty of  $(1 - \alpha)$  that the permitted limit has been exceeded. In general  $g$  will be a multiple of the standard uncertainty  $u$  as described in the next section.

**Use of decision rules**

The use of these rules is described in the following examples, which cover a range of decision rules and their application. The examples are limited to the case where the decision rule is utilised with a specification that sets an

Fig. 3 Acceptance/rejection zones for different decision rules



upper limit for the value of the measurand. Application of decision rules to specifications that set a lower limit or an upper and lower limit follows similar lines.

1. The batch will be considered in compliance if the value of the result is less than or equal to the limit, if it is above the limit it will be judged non-compliant.

In this case, as is shown in Fig. 3a the “Acceptance Zone” is the same as the “Specification Zone”. This type of decision rule would normally also state that the measurement uncertainty should be less than a certain percentage of the limit. For example it is used when the uncertainty is so small compared with the limit that the risk of making a wrong decision is acceptable. It could also be used when a standard method of measurement is prescribed in the specification or regulation and the measurement uncertainty has been taken into account in setting the limit. To use such a rule without specifying the maximum permitted value of the uncertainty would mean that the probability of making a wrong decision would not be known. Using this decision rule for the results in Fig. 1, Batches 1 & 2 are in compliance and Batches 3 & 4 are not.

2. The batch will be considered non-compliant if the value of the result exceeds the limit by more than twice the standard uncertainty  $u$ .

In this example, as shown in Fig. 3b, the start of the rejection zone is at the specification limit  $L$  plus an amount  $2u$ . For the case where the distribution of the values attributable to the measurand is approximately normal, this decision rule corresponds to a batch being assessed as non-compliant if the probability of the value of the measurand being greater than the limit is above about 97.5%. If it is assumed that the expanded uncertainty shown on the results in Fig. 1 correspond to twice the standard uncertainty  $u$ , then utilising this decision rule only Batch 4 would be judged non-compliant.

3. The batch will be considered to be non-compliant if the probability of the value of the measurand being greater than the limit exceeds  $P$ . (For this example  $P$  will be taken as 95%, but it is easy to apply to other levels of probability)

In order to implement this decision rule it is necessary to have information about the probability distribution function (PDF) of the values attributable to the measurand on the basis of the result and its uncertainty. Figure 3c shows the “Acceptance and Rejection zones” and the size of the guard band depends on the probability distribution.

In many cases it is generally sufficient to assume a normal PDF. The basis for making this assumption and the conditions under which it might be appropriate are given in Annex G of GUM [4]. The assumption is based on the use of the

Central Limit Theorem and section G 2.3 points out that “...if the combined standard uncertainty  $u$  is not dominated by a standard uncertainty component obtained from Type A evaluation based on just a few observations, or by a standard uncertainty component obtained from a Type B evaluation based on a rectangular distribution, a reasonable first approximation to calculating the expanded uncertainty  $U$  that provides an interval with a level of confidence  $P$  is to use for  $k$  the value from the normal distribution”.

Thus the start of the rejection zone is at the specification limit  $L + ku$ , where  $k$  is chosen to give the required value of  $P$ . For a value of  $P$  of 95% the value of  $k$  for a normal distribution is 1.65, thus the size of the guard band will be  $1.65u$ , and the start of the rejection zone is at  $L + 1.65u$ .

When the standard uncertainty is based on an effective number of degrees of freedom  $\nu$ , then  $P$  can be derived from the  $t$ -distribution and the size of the guard band will be  $t_{0.95} \cdot u$ ; for example if  $\nu = 5$  then the guard band will be  $2u$ . An alternative to using the effective number of degrees of freedom is given in Ref. [2, 5].

In order to implement decisions at other levels of confidence, then a value of  $k$  obtained from tables for the normal or the  $t$ -distribution at the appropriate level of  $P$  can be used.

However, in GUM section G 1.2 it is pointed out that since the value of  $U$  is at best only approximate it does not make sense to try to distinguish between say a 94% and a 96% level of confidence. In addition, GUM indicates that obtaining intervals with levels of confidence of 99% or greater is especially difficult.

In some cases the PDF of the values attributable to the measurand can be obtained from the PDFs of the input variables in the measurement model equation, either analytically or by Monte Carlo Simulation [6–8]. This enables the probability  $P$  required to assess compliance to be calculated directly.

In the above examples it has been implicitly assumed that the uncertainty is independent of the measured value  $x$ . The situation is a little more complicated when  $u$  is proportional to the value  $x$  of the measured variable. This arises for example in the case of many of trace level measurements, for the control of contaminants in food or of banned substances in sport, which have large uncertainties that are approximately proportional to the level of the analyte.

Consider the following two decision rules for compliance with an upper limit  $L$ .

1. For the value of the concentration of the analyte equal to the limit  $L$ , calculate the size of the guard band,  $g_1$ , such that the probability of obtaining a value of  $x$  greater than  $L + g_1$  is 5%. A value of  $x$  greater than  $L + g_1$  indicates non compliance
2. Calculate the size of the guard band,  $g_2$ , so that for an observed value of  $x$  equal to  $L + g_2$  the probability

that the value of the measurand is less than the limit is 5%. A value of  $x$  greater than  $L + g_2$  indicates non compliance.

Both of these decision rules require knowledge of the appropriate probability distribution; decision rule 1 requires the PDF,  $f(x,a,u)$ , of observed values  $x$ , for a given value of the measurand  $a$  and decision rule 2 the PDF,  $h(a,x,u)$ , of the values of the measurand for a given observed value  $x$ . When the uncertainty is not a function of the value of the measurand and  $f(x,a,u)$  is normal, then as is shown in Ref. [2],  $h(a,x,u)$  is also normal. In these circumstances the two decision rules are identical. This is not the case when the uncertainty is proportional to the value of the measurand when as will be shown below if  $f(x,a,u)$  is normal then  $h(a,x,u)$  is asymmetric with larger values of the measurand being more probable than for a normal distribution and therefore the size of the guard band for this decision rule may be different.

Decision rule 1, has the advantage that it utilises the value of  $u$  at a fixed value of  $a$ , that is at the limit  $L$ . Thus when  $f(x,a,u)$  is normal, the size of  $g_1$  can be determined using tables of the normal distribution. For a value of  $P_1$  of 5%,  $g_1 = 1.65 \cdot u_L = 1.65 \cdot L \cdot u_{rel}$ , where  $u_{rel}$  is the relative uncertainty; that is for  $L = 2$  and  $u_{rel} = 0.2$ ,  $g_1 = 0.66$ .

For decision rule 2, in order to determine the probability  $P_2$ , that the value of the measurand is less than the limit, it is necessary to derive the probability  $h(a,x,u)da$  that the value of the concentration  $a$  lies in the interval  $a + da$  when a value  $x$  has been observed. This can be done by using Bayes theorem, as described by Kacker et al. [2], which for non-informative prior is

$$h(a,x,u)da = \frac{f(x,a,u)da}{\int_0^\infty f(x,a,u)da} \tag{1}$$

Thus for an observed value  $L + g_2$  and taking  $f(x,a,u)$  as normal

$$P_2 = \frac{\int_0^L \frac{1}{u_a \sqrt{2\pi}} \exp\left[-\frac{(L+g_2-a)^2}{2u_a^2}\right] da}{\int_0^\infty \frac{1}{u_a \sqrt{2\pi}} \exp\left[-\frac{(L+g_2-a)^2}{2u_a^2}\right] da} \tag{2}$$

In order to compare these two decision rules it is helpful to write  $P_1$  in terms of an integral from 0 to  $L$ , taking advantage of the fact that the normal distribution is symmetrical, that is

$$P_1 = \int_0^L \frac{1}{u_L \sqrt{2\pi}} \exp\left[-\frac{(L+g_1-x)^2}{2u_L^2}\right] dx \tag{3}$$

If  $u$  is independent of the concentration, that is  $u_a = u_L = u$ , and  $f(x,a,u)$  is normal then the integrals in Eqs. 2 and 3 used to derive  $P_1$  and  $P_2$  are identical and

therefore, as was pointed out above  $g_1 = g_2$ . When  $u$  is proportional to the concentration  $a$ , i.e  $u_a = u_{rel} \cdot a + u_o$ , where  $u_o$  is the uncertainty at zero concentration, it is necessary to calculate the value of  $g_2$  that gives the required value of  $P_2$  by numerical integration. For  $P_1 = P_2 = 0.05$  a relative uncertainty of 0.1, a limit of 2.0 and taking  $f(x,a,u)$  as normal, both decision rules give guard bands of almost identical size. Even for a relative uncertainty of 0.2 the difference is not really significant, decision rule 1 gives  $g_1 = 0.66$  and decision rule 2 gives  $g_2 = 0.59$ .

Figure 4 shows the PDFs used to calculate the values of  $g_1$  and  $g_2$  for  $P_1 = P_2 = 0.05$  using Eqs. 2 and 3 for a relative uncertainty of 0.2 and a limit of 2.0. It can be seen that, although  $h(a,x,u)$  is asymmetric, and PDFs differ for larger values of  $x$  and  $a$  they are very similar up to the limit  $L$ . Calculations at other values of the uncertainty show that the asymmetry increases as the uncertainty increases but the size of guard bands do not differ significantly.

In practice, since both of these decision rules give essentially the same result, decision rule 1 might be preferred since it is much simpler to implement and it is also a suitable decision rule to use when  $u$  does not vary with concentration. However, basing the decision rule on the uncertainty at the measured value leads to a different result.

### Design of decision rules

It is important that decision rules are clear and unambiguous and that they provide the required level of confidence in the assessment decision.

The decision rule in example 2 in the previous section are clear, unambiguous and easy to apply, but the level of

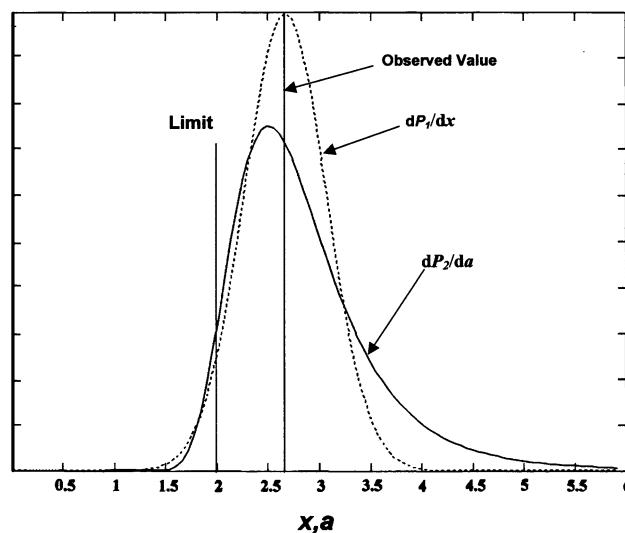


Fig. 4 PDFs used to calculate  $g_1$  and  $g_2$

confidence that the limit has been exceeded depends on the PDF of the values attributable to the measurand and this has not been taken into account. In the case of this decision rule, for a PDF that is normal, a sample would be taken to indicate compliance if the probability of the measurand being greater than the limit was 97.5%. However as was pointed out in the example if  $u$  were based on just 5 degrees of freedom the probability level would be reduced to 95%. The important point is that the decision rule needs to meet the requirements of the regulation or specification.

In addition the decision rule may need to specify a maximum value for  $u$  since the larger the value of  $u$  the larger is the proportion of the samples that will be judged incorrectly. The smaller the value of  $u$ , in general, the higher the cost of analysis will be. Thus, ideally  $u$  should be chosen to minimise the cost of the analysis plus the cost of the wrong decision. However, the information needed to do this is very rarely available. A common approach is to carry out screening measurements using a relatively inexpensive method with a comparatively large uncertainty and to follow this using a method with a small uncertainty for those samples for which the screening result implies non-compliance.

## Summary

In order to decide whether or not to accept/reject a product requires.

- (a) a specification giving the specific values of the characteristics (measurands) being controlled and the upper and/or lower bounds of their permissible values, and
- (b) a decision rule that describes how the measurement uncertainty will be taken into account with regard to accepting or rejecting a product according to its specification and the result of a measurement.

The decision rule should have a well-documented method of determining the location of acceptance and

rejection zones. Ideally it should include the minimum acceptable level of the probability that the measurand lies within the specification limits. It should also cover other items appropriate to the test being carried out such as the maximum allowable value of the uncertainty and how to deal with repeat measurements and outliers.

Utilising the decision rule the size of the acceptance or rejection zone is determined by adjusting the limits by means of appropriate guard bands. The size of the guard band is calculated from the value of the measurement uncertainty and the minimum acceptable level of the probability that the measurand lies within the specification limits, as described in the section Use of Decision Rules. Special attention is required when the measurement uncertainty is proportional to concentration.

In addition a reference to the decision rules used should be given when reporting on compliance.

## References

1. Guidelines for Decision Rules: Considering Measurement Uncertainty in Determining Conformance with Specifications ASME (New York) B89.7.3.1-2001
2. Kacker R, Jones A (2003) On use of Bayesian statistics to make the guide to the expression of uncertainty in measurement consistent. *Metrologia* 40(5):235–248. doi:10.1088/0026-1394/40/5/305
3. COMMISSION DECISION of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results (2002/657/EC) <http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=OJ:L:2002:221:0008:0036:EN:PDF>
4. Guide to the Expression of Uncertainty in Measurement ISO, Geneva (1993) ISBN 92-67-10188-9
5. Williams A (1999) *Accred Qual Assur* 4:14–17. doi:10.1007/s007690050303
6. Cox MG, Harris PM (2003) *Accred Qual Assur* 8:375–379. doi:10.1007/s00769-003-0646-4
7. Cox MG, Harris PM (2001) Measurement uncertainty and the propagation of distributions. Technical report. National Physical Laboratory, Teddington
8. Gonzalez AG, Herrador MA, Asuero AG (2005) *Accred Qual Assur* 10:149–154. doi:10.1007/s00769-004-0896-9