

ANNEX 4: TECHNICAL NOTE ON MEASUREMENT UNCERTAINTY

1 Introduction

The principal aim of analytical work is to gain information on the material under investigation. This information always constitutes a probability distribution determined by a random error and a systematic error inherent in the analytical procedure used. A systematic error can act as an additive or as a multiplicative shift. Systematic errors are superimposed by the random error. Analytical practice shows that there is always some doubt about the correctness of a stated result, even when all the suspected sources of error have been taken into account and the appropriate corrections have been applied. This is due to the uncertainty regarding the correction factors and the uncertainty arising from random effects, which cannot be eliminated, although they can be reduced by increasing the number of observations. The result of an analysis after careful consideration of all error sources may by chance be very close to the true value. However, the uncertainty can still be very large, simply because the analyst is very unsure of how close that result is to the true value. Consequently, a measurement cannot be properly interpreted without the knowledge of the uncertainty associated with the result.

The concept of expressing or estimating the uncertainty of measurements was developed to inform the final users of the analytical data concerning how much allowance must be made for the possibility that repetition of the test will give a different value (Horwitz, 1998). This information is particularly necessary when analytical results are not used by the data originator, as is the regular case in the assessment of data from environmental monitoring programmes.

This technical note provides information on how the uncertainty of measurement of the analytical methods used in the COMBINE programme of HELCOM can be estimated, so that it would be possible to judge whether or not the accuracy (trueness and precision) of the method meets the requirements of this programme. It should be taken into account that the requirements on accuracy depend on the aims and the purpose of the monitoring programme.

2 Definitions

In accordance with the current version of the International Vocabulary of Basic and General Standard Terms in Metrology (ISO, 1993), the ISO Guide (ISO, 1995) defines **Measurement Uncertainty** as a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand. The following definitions apply (EURACHEM/CITAC, 2000):

Standard uncertainty

$u(x_i)$ uncertainty of the result x_i of a measurement expressed as a standard deviation.

Combined standard uncertainty

$u_C(y)$ standard uncertainty of the result y of a measurement when the result is obtained from the values of a number of other quantities, equal to the positive square root of a sum of terms, the terms being the variances or covariances of these other quantities weighed according to how the measurement results vary with these quantities.

Expanded standard uncertainty

U quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand, an expanded uncertainty is calculated from a combined standard uncertainty u_C and a coverage factor k using $U = k \times u_C$

Coverage factor

k numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty, the choice of the factor k is based on the level of confidence desired; $k = 2$ for an approximate level of confidence of 95 %.

3 Procedures to Estimate Uncertainty

Generally, there are two main approaches to estimate the uncertainty of analytical measurements or an analytical procedure, respectively.

According to the original ISO approach (Horwitz, 1998), the uncertainty of analytical results is derived by listing all of the possible errors in the form of standard deviations. Using this error budget model, the combined standard uncertainty can then be calculated as the square root of the sum of squares of the individual error components. This "bottom-up" approach of ISO

assumes that an analytical method can be structured into small, simple steps, and that an individual standard uncertainty can be attributed to all of these steps, sometimes based on a best guess of experienced analysts.

A “top-down” view on estimating the combined standard uncertainty through intercomparison tests was developed by the Analytical Methods Committee of the Royal Society of Chemistry of the UK (1995). Both systematic and random errors of individual laboratories become random in an intercomparison test or in a laboratory proficiency testing scheme, provided that identical and homogeneous samples are analysed. Following ISO 5725 (1994), the within- and between-laboratory variance can be calculated and combined in the reproducibility of the intercomparison test. Only in the case that the same analytical procedure is used, can the calculated reproducibility of the intercomparison test be considered as the standard uncertainty of measurement.

A further “top-down” approach is provided by the Nordic Committee on Food Analysis (1997) based on experimental data generated in the individual laboratory. According to this proposal, the combined standard uncertainty is characterized as the internal reproducibility standard deviation calculated from method validation data or using information from internal quality assurance measures (analysis of certified reference materials, participation in proficiency testing schemes or use of reference methods). In the case that analytical results are compared to those of a reference method, BCR information can be used to estimate the uncertainty (BCR, 1998).

A comprehensive description of all of these approaches can be found in a recently released EURACHEM/CITAC Guide (2000).

3.1 The “bottom-up” procedure

The “bottom-up” approach to estimate the uncertainty of analytical results seems to be rather impractical (Horwitz, 1998), because it does not include the outcome from intercomparison exercises or from laboratory proficiency testing schemes in marine monitoring that are available today. Therefore, if information on the uncertainty of analytical data generated in the COMBINE programme is needed, the “top-down” approach should be preferred. The procedure by which the uncertainty of measurements is calculated depends on the requirements of the monitoring programme.

3.2 The “top-down” model using data from intercomparison exercises

As stated in the EURACHEM/CITAC Guide (2000), the reproducibility standard deviation of intercomparison exercises which are carried out according to ISO 5725 (1994), or Thompson and Wood (1993) can be used as the combined standard uncertainty for methods operating within their defined scope, provided that contributions from additional factors (i.e., sampling error, inhomogeneous distribution of the analyte) have been shown to be negligible. The sampling error and its inclusion in the uncertainty budget requires special investigations. The same is true for the uncertainty of co-factor determinations carried out in the COMBINE Programme to support monitoring results. Attachment 1 gives a practical example of the uncertainty estimation using data from an intercomparison exercise.

3.3 The “top-down” model for individual laboratories

In case an individual laboratory is requested to provide information on uncertainty of measurement in connection with reported data, the approach proposed by the Nordic Committee on Food Analysis (1997) should be followed. The so-called internal reproducibility standard deviation can be derived easily from internal quality control charts or from special investigations carried out to determine the internal reproducibility.

According to the EURACHEM/CITAC Guide (2000), an uncertainty estimation based on validation studies and routine quality control requires the best available estimate of the overall bias and the precision together with a quantification of any uncertainties associated with effects incompletely accounted for in the method performance studies (i.e., matrix effects, different concentration levels, calibration model used). Information on the overall bias and precision should be available from routine quality control charts, which are to be established in the COMBINE monitoring laboratories. For example, the overall bias is best estimated by the analysis of an appropriate certified reference material (CRM), analysed several times over a period of time or regularly between routine monitoring samples. In that case, the combined standard uncertainty is simply the combination of the standard deviation associated with these determinations and the uncertainty in the CRM value (see Attachment 2 for an example). To study the influence of additional parameters on their contribution to the uncertainty budget, special investigations are needed. These are normally carried out in the form of robustness or

ruggedness testing during method validation studies (Youden and Steiner, 1975). Robustness studies provide information on the variability of analytical results caused by small deviations from the experimental design prescribed by the analytical method.

4 Reporting of Uncertainty

The expanded uncertainty (estimated combined standard uncertainty multiplied by the coverage factor k , usually $k = 2$ for a confidence limit of 95 %) should be reported for individual monitoring parameters in the form of a standard deviation or confidence interval together with information on how it was determined.

5 References

- Analytical Methods Committee. 1995. *Analyst*, 120: 2303–2308.
- BCR. 1998. *Metrology in chemistry and biology: a practical approach*.
- EURACHEM/CITAC. 2000. *EURACHEM/CITAC Guide, Quantifying uncertainty in analytical measurement. Second Edition*.
- Horwitz, W. 1998. Uncertainty—a chemist's view. *J. AOAC Int.*, 81: 785–794.
- ISO. 1993. *International Vocabulary of Basic and General Standard Terms in Metrology*.
- ISO. 1994. *ISO 5725. Accuracy (trueness and precision) of measurement methods and results*.
- ISO. 1995. *Guide to expression of uncertainty in measurements*.
- Nordic Committee on Food Analysis. 1997. *Estimation and expression of measurement uncertainty in chemical analysis. NMKL Procedure No. 5*.
- Thompson, M., and Wood, R. 1993. *The International Harmonized Protocol for the Proficiency Testing of (Chemical) Analytical Laboratories (Technical Report)*. *Pure and Applied Chemistry*, 65: 2123–2144.
- Youden, W.J., and Steiner, E.H. 1975. *Statistical Manual of the AOAC*. AOAC, Arlington, VA.

Attachment 1.

1. Protocol for evaluation of measurement uncertainty from in-house quality control measurements

Estimating measurement uncertainty can be done by identifying all possible sources of uncertainty associated with a method, quantifying uncertainty components (estimate the magnitude of the uncertainty associated with each potential source) and calculating total uncertainty by combining the individual uncertainty components following appropriate mathematical rules (“bottom-up” approach, see, e.g., EURACHEM Guide (2003)).

Another approach uses data from routinely undertaken internal quality control measurements, e.g., results of the replicate analyses of certified reference materials (CRM), without identifying all potential sources of uncertainty associated with the method and quantifying uncertainty components (“top-down” approach).

This document gives guidance how uncertainty estimates for a method can be obtained from replicate quality control measurements of a representative certified reference material. It is assumed that these measurements comprise the total analytical procedure and have been carried out with appropriate frequency and during a sufficiently long time period. In that way, it can be ensured that most relevant uncertainty components associated with the method will be covered (starting from the laboratory sample or analysis sample, excluding contributions associated with sampling and sample handling prior to analysis).

Following these assumptions, the total uncertainty of the method is composed of:

- a contribution from the precision of the method, and
- a contribution from the trueness of the method (recovery of the analyte from the CRM).

Both contributions can be easily quantified using data from routinely performed quality control measurements.

The mathematical equations 1 to 6, outlined below, can be applied for calculating measurement uncertainty on the condition that the relative uncertainty of measurement of the method expressed in percent is approximately constant within the working range.

This condition might apply in most cases as long as the lower limit of the working range is well above the limit of quantification (see Section 3, below).

This implies further that

- (1) the precision expressed as relative standard deviation (RSD) is approximately constant within the working range considered. This denotes that the absolute standard deviation increases proportionally with increasing concentration of the analyte in the sample.
- (2) the relative uncertainty of the recovery of the analyte from the CRM $u(\bar{R}_m)$ is independent of the concentration of the analyte. This denotes that it is approximately constant, e.g., $\pm 5\%$ of the determined concentration.

Note: If this condition does not apply, modified mathematical equations adjusted to the specific circumstances need to be used. For details see (Barwick, 2000).

Then, the relative combined uncertainty, $u_c(y)$, of the method is obtained using the following equation:

$$u_c(y) = \sqrt{RSD^2 + u(\bar{R}_m)^2} \quad \text{Eq. 1}$$

The relative standard deviation is given by:

$$RSD = \frac{s_{obs}}{\bar{C}_{obs}} \quad \text{Eq. 2}$$

where \bar{C}_{obs} is the mean of replicate analyses of the CRM and s_{obs} is the standard deviation of the results from the replicate analyses of the CRM.

The relative uncertainty of the recovery, $u(\bar{R}_m)$, is calculated using:

$$u(\bar{R}_m) = \bar{R}_m \times \sqrt{\left(\frac{s_{obs}^2}{n \times \bar{C}_{obs}^2}\right) + \left(\frac{u(C_{CRM})}{C_{CRM}}\right)^2} \quad \text{Eq. 3}$$

where C_{CRM} is the certified concentration of the analyte in the CRM, n is the number of replicates ($n \geq 10$, see [2]) and $u(C_{CRM})$ is the standard uncertainty of the certified concentration for the CRM with a mean recovery, \bar{R}_m , given by:

$$\bar{R}_m = \frac{\bar{C}_{obs}}{C_{CRM}} \quad \text{Eq. 4}$$

It is assumed that \bar{R}_m does not differ significantly from 1 and, hence, no correction for recovery is made. To determine whether the recovery is significantly different from 1, a significance test is used. The test statistic t is calculated using the following equation

$$t = |1 - \bar{R}_m| / u(\bar{R}_m) \quad \text{Eq. 5}$$

If the degrees of freedom associated with $u(\bar{R}_m)$ are known, t is compared with the two-tailed critical value, t_{crit} , for the appropriate number of degrees of freedom at 95% confidence.

If t is less than the critical value, then \bar{R}_m is not significantly different from 1.

If the degrees of freedom associated with $u(\bar{R}_m)$ are unknown, for example, if there is a contribution from the uncertainty in the certified value of a reference material, t is compared with k , the coverage factor that will be used in the calculation of the expanded uncertainty (see Eq. 6) (Barwick and Ellison, 2000).

If $|1 - \bar{R}_m| / u(\bar{R}_m) < k$, the recovery is not significantly different from 1.

If $|1 - \bar{R}_m| / u(\bar{R}_m) > k$, the recovery is significantly different from 1 and results are corrected for recovery.

Guidance on how to proceed is given in (Barwick and Ellison, 2000).

To calculate the combined uncertainty, $u_c(y)$, both relative standard uncertainties RSD and $u(\bar{R}_m)$ are combined following equation 1.

The expanded uncertainty, $U(y)$, is obtained by multiplying the combined standard uncertainty, $u_c(y)$, by an appropriate coverage factor, k , (Eq. 6). For most cases, a coverage factor of 2 is recommended, which gives an interval containing approximately 95% of the distribution of values:

$$U(y) = k \times u_c(y) = 2 \times u_c(y) \quad \text{Eq. 6}$$

The result y of an analytical measurement should be stated together with the corresponding expanded uncertainty, $U(y)$, in the following form:

(result): $x \pm U$ [units]

The stated uncertainty is an expanded uncertainty, calculated using a coverage factor of 2. This corresponds approximately to the 95% confidence interval (EURACHEM, 2003).

2 ESTIMATION OF MEASUREMENT UNCERTAINTY USING REPRODUCIBILITY DATA FROM INTERLABORATORY STUDIES

In principle, it is possible to use the relative reproducibility standard deviation, CV_R , obtained in intercomparison studies as a basis for estimating the uncertainty of a method in a particular laboratory, if there is no significant difference between the relative repeatability standard deviation seen in the interlaboratory study and that observed in the laboratory. If so, this indicates that the precision achieved in the particular laboratory is similar to that obtained by the participants of the interlaboratory study (EURACHEM 2003).

For estimating the laboratory's expanded uncertainty, the relative reproducibility standard deviation CV_R obtained in the interlaboratory study is assumed to be an estimate of the combined standard uncertainty of the laboratory and multiplied with the coverage factor $k = 2$.

The uncertainty for a method, $U(y)$, obtained in that way can only be considered as a rough estimate for getting an idea about the order of uncertainty, but cannot replace estimating uncertainty from own measurements of, e.g., certified reference materials.

3 TRANSITION TO CONSTANT ABSOLUTE UNCERTAINTY OF MEASUREMENT AT LOW CONCENTRATIONS

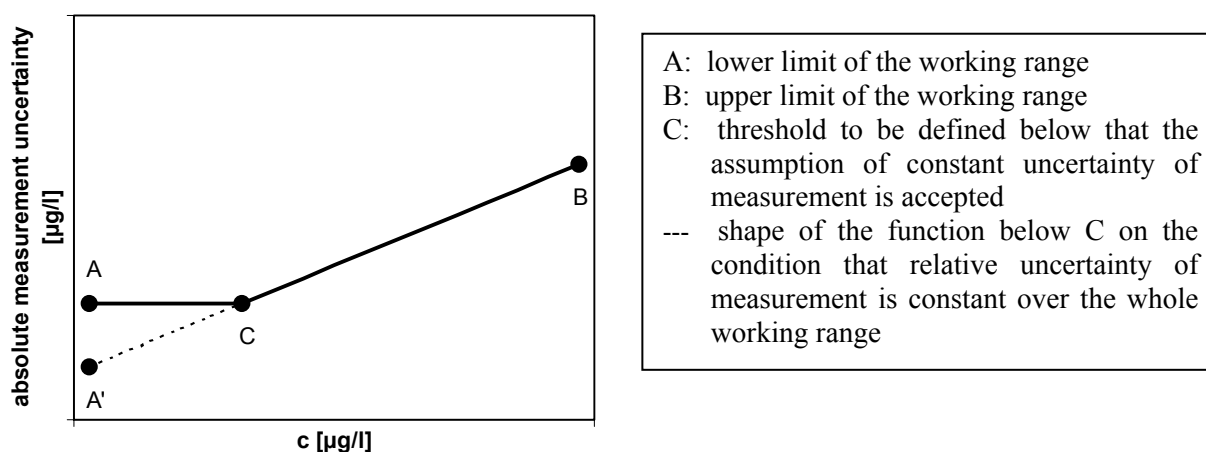


Figure 1. Graphical representation of the absolute uncertainty of measurement as a function of analyte concentration

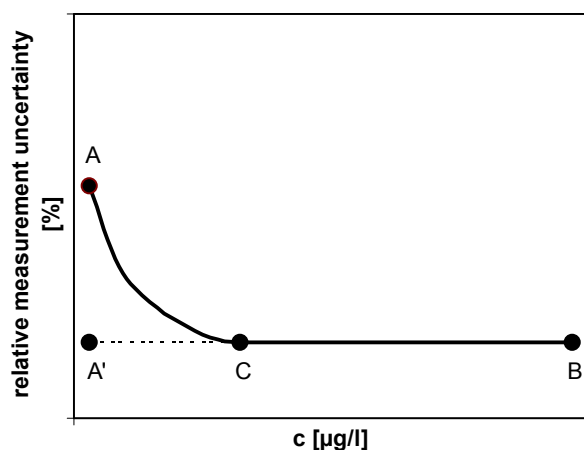


Figure 2. Graphical representation of the relative uncertainty of measurement as a function of analyte concentration

4 INTRODUCTION TO ADDITIONAL UNCERTAINTY COMPONENT IN CASE OF SIGNIFICANT DEVIATION FROM 1 OF THE RECOVERY OF THE ANALYTE FROM CRM

If the recovery \bar{R}_m of the analyte from the reference material differs significantly from 1 (t -test, $t \geq t_{krit}$), an additional uncertainty component is introduced.⁶ Instead of Eq. 1, Eq. 7 and 8 apply.

$$u_c(y) = \sqrt{RSD^2 + u(\bar{R}_m)^2} \quad \text{Eq. 1}$$

$$: \quad u_c(y) = \sqrt{RSD^2 + u(\bar{R}_m)^2 + \Delta^2} \quad \text{Eq. 7}$$

$$\Delta = \bar{C}_{obs} - C_{CRM} \quad \text{Eq. 8}$$

where Δ is the deviation of the measured concentration in the CRM from the reference value

REFERENCES

- Barwick, V.J., and Ellison S.L.R. 2000. Development and Harmonisation of Measurement Uncertainty Principles – Part (d): Protocol for uncertainty evaluation from validation data; VAM Project (Version 5.1, Jan. 2000), <http://www.vam.org.uk/publications> (look under publications for the first author).
- EURACHEM. 2003. EURACHEM/CITAC Guide: Quantifying Uncertainty in Analytical Measurement, www.uni-stuttgart.de/eurachem/pdf

⁶ If the recovery \bar{R}_m of the analyte from the reference material differs significantly from 1, the analytical procedure is to be checked for the reason of the bias and, where applicable, the method has to be modified. But, in some cases, if the uncertainty of the certified concentration of the analyte in the CRM is extremely small, significant differences in the t -test can be observed even when the recovery, \bar{R}_m , is close to 1.

EXAMPLE 1: ESTIMATION OF MEASUREMENT UNCERTAINTY USING THE RESULTS OF REPLICATE ANALYSES OF A CRM

During routine analyses of phosphate in seawater samples, a certified reference material was regularly analysed (30 times) as an AQC sample over a period of three months. The certified phosphate concentration in the reference material was $2.43 \pm 0.41 \mu\text{mol l}^{-1}$ and assumed to be representative for the working range of the method.

According to manufacturer's specifications, the confidence interval of the phosphate concentration in the CRM was calculated using the reproducibility standard deviation obtained in the certification interlaboratory study multiplied by three. Hence, the standard uncertainty of the phosphate concentration, $u_c(\text{PO}_4)$, in the CRM is given by $0.41 \mu\text{mol l}^{-1} / 3 = 0.14 \mu\text{mol l}^{-1}$.

Note: Be aware that depending on the producer of the CRM, different modes of calculation for the confidence interval of the certified concentration are in use. This must be taken into account when calculating the standard uncertainty of the certified concentration, $u_c(y)$, in the CRM.

Certified concentration of the Phosphate in the reference material $C_{CRM} = 2.43 \mu\text{mol l}^{-1}$

Standard uncertainty of the certified Phosphate concentration $u(C_{CRM}) = 0.14 \mu\text{mol l}^{-1}$

From the results of the replicate analyses of the CRM, the following values can be determined directly:

Mean of replicate analyses of the CRM, $\bar{C}_{obs} = 2.34 \mu\text{mol l}^{-1}$

Standard deviation of the results from the replicate analyses of the CRM, $s_{obs} = 0.12 \mu\text{mol l}^{-1}$

Then, the relative standard deviation of the mean of the Phosphate concentration, RSD_{PO_4} , is given by:

$$RSD = \frac{s_{obs}}{\bar{C}_{obs}} = \frac{0.12 \mu\text{mol l}^{-1}}{2.34 \mu\text{mol l}^{-1}} = 0.051$$

and the recovery, \bar{R}_m , is given by:

$$\bar{R}_m = \frac{\bar{C}_{obs}}{C_{CRM}} = \frac{2.34 \mu\text{mol l}^{-1}}{2.43 \mu\text{mol l}^{-1}} = 0.963$$

To calculate the relative standard uncertainty of the recovery, $u(\bar{R}_m)$ equation 3 is used:

$$u(\bar{R}_m) = \bar{R}_m \times \sqrt{\left(\frac{s_{obs}^2}{n \times \bar{C}_{obs}^2}\right) + \left(\frac{u(C_{CRM})}{C_{CRM}}\right)^2} = 0.963 \times \sqrt{\left(\frac{0.12^2}{30 \times 2.34^2}\right) + \left(\frac{0.14}{2.43}\right)^2} = 0.056$$

To test whether the observed recovery is significantly different from 1, a statistical significance test (t-test) is performed following equation 5:

$$t = \frac{|1 - \bar{R}_m|}{u(\bar{R}_m)} = \frac{1 - 0.963}{0.056} = 0.661$$

If $t < k$ (coverage factor), it can be assumed that the recovery is not significantly different from 1 [2]. Since 0.661 is less than 2, the significance test indicates no significant difference between the observed recovery (0.963) and 1.

The relative combined standard uncertainty $u_c(\text{PO}_4)$ is then estimated as:

$$u_c(\text{Phos}) = \sqrt{RSD^2 + u(\bar{R}_m)^2} = \sqrt{0.051^2 + 0.056^2} = 0.076$$

Using the recommended coverage factor $k = 2$, the expanded uncertainty, $U(\text{PO}_4)$, is given by:

$$U(\text{Phos}) = k \times u_c(y) = 2 \times 0.076 = 0.152$$

Result: the relative expanded uncertainty, $U(\text{PO}_4)$, for the determination of phosphate in seawater samples within the considered working range is 0.152 and 15.2%, respectively.

This denotes for a theoretical result of $10.0 \mu\text{mol l}^{-1}$ Phosphate:

Phosphate concentration: $10.0 \pm 1.5 \mu\text{mol l}^{-1}$,
the stated uncertainty is an expanded uncertainty, calculated using a coverage factor of 2 (this corresponds approximately to the 95% confidence interval).

EXAMPLE 2: ESTIMATION OF MEASUREMENT UNCERTAINTY USING REPRODUCIBILITY DATA FROM INTERLABORATORY STUDIES

The results of the three QUASIMEME exercises on determination of phosphate in seawater carried out in 2001 were as follows:

Relative reproducibility standard deviation: 4.67 / 4.47 / 6.30 %
Phosphate concentration - Assigned value: 9.71 / 13.08 / 1.88 $\mu\text{mol l}^{-1}$

Using this information, the averaged relative reproducibility standard deviation expressed as coefficient of variation, $CV(\text{PO}_4)$, for the intercomparison study on phosphate determination in seawater is, which can be equated with combined standard uncertainty, $u_c(\text{PO}_4)$, is given by:

$$CV(\text{PO}_4) = \sqrt{\sum x_i^2 / n} = \sqrt{(4.67^2 + 4.47^2 + 6.30^2) / 3} = 5.2 \%$$

Using the recommended coverage factor $k = 2$, the expanded uncertainty, $U(\text{PO}_4)$, is 10.4%.

This result is in satisfactory agreement with the estimated expanded uncertainty, $U(\text{PO}_4)$, of 15.2% obtained using the results of replicate analyses of a certified reference material.

Attachment 2

Measurement Uncertainty in Environmental Chemical Analytical Practice—Uncertainty Associated with Bias

Environmental chemical analytical practice shows that systematic errors (bias) can be assessed by the analysis of appropriate certified reference materials (CRMs).

The reference material used should be typical of the test material being analysed, in terms of both the chemical and physical properties of the matrix and the concentration of the analyte. If these conditions are met, it is possible to infer (within the limits set by uncertainty) that a bias apparent in the results obtained on the analysis of certified reference material is also present in the results of a test material under investigation.

Commonly, bias is estimated during the process of method validation. The respective value obtained is then held to apply during subsequent use of the analytical method. To justify this assumption, all routine runs of the method must contain appropriate reference material to act as an internal quality control. This would help to ensure that the analytical system does not change in any significant way that would invalidate the initial estimate of the bias.

The principle of uncertainty is a helpful tool in planning and executing bias experiments. The intent of the following is to indicate how uncertainty in bias is estimated and included in the combined standard uncertainty in a measurement result.

The difference between the measurement results x_q and the accepted reference value x_r of a certified reference material (CRM) is a measure to estimate the analytical bias.

The concentration value x_q of the CRM measurement is compared with the assigned value x_r and a correction (or recovery) factor is calculated according to the equation

$$f_r = \frac{x_q}{x_r} \quad (1)$$

The critical value

$$T = \frac{|1 - f_r|}{u(f_r)} \quad (2)$$

is used to test whether the correction factor f_r deviates significantly from unity.

Such a test considers the question “is $|1 - f_r|$ greater than $u(f_r)$, the uncertainty in the determination of f_r , at a selected confidence level?”

For $T > 2$ it is concluded that a bias exists at an approximate level of confidence of 95 %, and the corrected measurement result x_{corr} is calculated according to equation

$$x_{corr} = \frac{x_p}{f_r} \quad (3)$$

x_p is the concentration value of the analyte in the test material analysed with the uncertainty $u(x_p)$ derived from an instrumental response.

If $T < 2$ a bias cannot be detected at the selected confidence level mentioned and x_p remains uncorrected.

The standard uncertainty of f_r is estimated according to the equation

$$u(f_r) = f_r \sqrt{\left(\frac{u(x_q)}{x_q}\right)^2 + \left(\frac{u(x_r)}{x_r}\right)^2} .$$

(4)

The result of a measurement is only an approximation or estimate of the value of the analyte and thus is complete only when accompanied by a statement of the combined standard uncertainty of that estimate.

The uncertainty associated with the correction factor needs to be included in the uncertainty budget regardless of whether or not f_r deviates significantly from 1. Accordingly, the combined uncertainty of the sample value is

$$u(x_{corr}) = x_{corr} \sqrt{\left(\frac{u(x_p)}{x_p}\right)^2 + \left(\frac{u(f_r)}{f_r}\right)^2} .$$

(5)