Available online at www.sciencedirect.com

SCIENCE DIRECT®

Talanta 65 (2005) 829–837

Talanta

www.elsevier.com/locate/talanta

Attempts to include uncorrected bias in the measurement uncertainty

Václav Synek*

Faculty of the Environment, Jan Evangelista Purkyně University, Králova výšina 7, CZ-40096 Ústí nad Labem, Czech Republic

Received 25 January 2004; received in revised form 1 July 2004; accepted 26 July 2004
Available online 8 September 2004

Abstract

In ISO *Guide* it is strictly recommended to correct results for the recognised significant bias, but in special cases some analysts find out practical to omit the correction and to enlarge the expanded uncertainty for the uncorrected bias instead. In this paper, four alternatively used methods computing these modified expanded uncertainties are compared according to the levels of confidence, widths and layouts of the obtained uncertainty intervals. The method, which seems to be the best, because it provides the same uncertainty intervals as in the case of the bias correction, has not been applied very much, maybe since these modified uncertainty intervals are not symmetric about the results. The three remaining investigated methods maintain their intervals symmetric, but only two of them provide intervals of the kind, that their levels of confidence reach at least the required value (95%) or a larger one. The third method defines intervals with low levels of confidence (even for small biases). It is proposed a new method, which gives symmetric intervals just with the required level of confidence. These intervals are narrower than those symmetric intervals with the sufficient level of confidence obtained by the two mentioned methods. A mathematical background of the problem and an illustrative example of calculations applying all compared methods are attached.

Keywords: Bias; Coverage probability; Level of confidence; Systematic error; Uncertainty

1. Introduction

ISO *Guide to the Expression of Uncertainty in Measurement* (ISO *Guide*) [1] strongly recommends to correct results for the recognised significant bias and not to take this bias into account by enlarging the uncertainty. In spite of this fact in specific cases some authors [2–12] find reasonable not to correct individually each result for the bias but to include the bias into the uncertainty assigned to all the results instead. Phillips and Eberhardt [5] compared two of these methods with their newly proposed method in 1997 and Maroto et al. [12] summarized four of these methods in 2002.

This paper mentions proceedings of the bias evaluation and summarizes some ways of the bias incorporation in the overall combined uncertainty, which have been published. It investigates four of these methods, compares them providing that the measurement results are normally distributed with a constant or relative systematic error and proposes a new method. At the end of this paper a mathematical background of the problem and illustrative practical examples of applications of the investigated methods are appended.

2. Estimation of bias and calculation of uncertainty

Systematic error, often termed bias, affects the measurement results in repeated observations always in the same way hence it does not bring about the result variability. In spite of it, this type of errors also contributes to the result uncertainty, because the correction of a result for its systematic error is known with some uncertainty, which should be incorporated into the combined uncertainty of the result. The systematic error of an analytical procedure is estimated in a validation process, when the trueness is assessed as described in *Quantifying Uncertainty in Analytical Measurements* [13] and in a rank of papers, e.g. [6,9–11,14,15]. The overall value of the systematic error can be evaluated by replicate analysis of a

* Tel.: +420 475 284151; fax: +420 475 284185.
E-mail address: synek@fzp.ujep.cz.

0039-9140/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2004.07.038

relevant CRM or spiked samples or by replicate analysis of a suitable sample performed with the tested method and a standard method.

When the systematic error is supposed to have a constant absolute value[1], the estimate of its value, b, is calculated by using the mean of the results obtained with the tested method, $\bar{x}_r$, and the reference value, x_{ref}:

$$b = \bar{x}_r - x_{ref} \quad (1)$$

Provided that a CRM was used in the bias estimation, the bias uncertainty, $u(b)$, can be calculated:

$$u(b) = \left[\frac{s_r^2}{p} + u^2(x_{ref})\right]^{1/2} \quad (2)$$

where $u^2(x_{ref})$ is the standard uncertainty in the certified value of the CRM and s_r is the sample standard deviation of the p replicate results of the CRM analysis. (Some ways of the bias uncertainty estimation by using spiked samples or standard analytical methods are given e.g. in [10,11,13–15].) The found bias b is tested for its significance; b is not significant when:

$$|b| \leq ku(b) \quad (3)$$

where k is usually a value of the coverage factor (see Eq. (9)) if the number of effective degrees of freedom is high enough [8,9,11,15]. In some works [6,7,10,12–14] it is more precisely required to replace k by the two-tailed value from Student's distribution for the effective degrees of freedom associated to $u(b)$ at a given level of confidence.

When the systematic error is assumed to be relative, recovery, R, is calculated:

$$R = \frac{x_r}{x_{ref}} \quad (4)$$

and its difference from 1 is tested on significance:

$$|R - 1| \leq ku(R) \quad (5)$$

where $u(R)$ is the uncertainty of R, which can be estimated by analogy with Eq. (2), (see e.g. [6,10,11,14,15]).

When the systematic error is statistically significant all results, x, measured on routine samples should be corrected by using the determined value b or R (before the analyst should have tried to revise the analytical procedure in order to eliminate the significant error):

$$x_c = x - b \quad (6)$$

or

$$x_c = \frac{x}{R} \quad (7)$$

where x_c denotes the measurement results after the correction. The standard uncertainty of b or R is then included in the evaluation of the combined standard uncertainty, u_c, of the corrected result. When Eq. (6) is applied for the correction, $u(b)$ is combined with the other uncertainty components, $u(x_1)$, $u(x_2)$..., in this way:

$$u_c = [u^2(b) + u^2(x_1) + u^2(x_2) + \cdots]^{1/2} \quad (8)$$

In Eq. (8) there should be at least one uncertainty component besides $u(b)$; it is a standard deviation estimating intermediate precision for the routine samples[2].

In the case of a relative systematic error the relative uncertainties are used in Eq. (8). (The calculations of combined uncertainties are explained in detail e.g. in [1,13].) Provided that the value of a significant correction itself or the value of its uncertainty is negligible with respect to the overall combined standard uncertainty, they may be ignored [1]. When the systematic error is not statistically significant, usually no correction is applied, but the uncertainty of its evaluation, $u(b)$ or $u(R)$, is again included into the overall combined uncertainty in the measurement, since the insignificancy is also known with this uncertainty. So that in the case of both significant and insignificant biases the combined standard uncertainty u_c of the measurement is calculated in the same way.

After the evaluation of u_c the expanded uncertainty, U, can be computed:

$$U = ku_c \quad (9)$$

where k is a coverage factor. When the probability distribution of measurement results is approximately normal with a standard deviation equal to u_c and the number of effective degrees of freedom for u_c is of a significant size, a coverage factor of $k = 2$ ($k = 3$) produces an uncertainty interval having a level of confidence of approximately 95% (99%) [1].

3. Are there any sensible reasons to enlarge the uncertainty intervals instead of the correction?

As was stated in the introduction, in specific cases some analysts do not follow the recommendation of ISO *Guide* [1] and do not correct each individual result for the recognised significant bias, but instead they modify the expanded uncertainty for this bias. It might seem to be paradoxical to estimate the bias of a measurement first and not to use it for correction in the end. The papers dealing with this problem,

[1] *Note* 1: In literature there are two types of systematic errors usually distinguished: constant and proportional [14,15]. Both of them are supposed to be constant in a concentration range – the former expressed in an absolute value and the latter expressed in a relative value. (There are also systematic errors, which cannot be thought constant in both absolute and relative values, e.g. errors arising from gradually changing temperature [13].) These types of errors are termed also additive and multiplicative [2] or constant absolute and constant relative [9] or even translational and rotational systematic errors.

[2] *Note* 2: Before this calculation some authors [7,10] require the bias uncertainty to be modified: $u(b)t/k$, where t is the appropriate value from Student's distribution and k is from Eq. (9).

which have been published [2–12], prove that such proceeding seems many analysts to be sensible.

Philips and Eberhardt [5] stress that “paper and pencil” corrections to each measurement value can be time consuming, particularly under high measurement throughput situations, so that it may be more economically reasonable to simply account for the bias by enlarging the uncertainty value that is attached to every measurement result. Ellison [7], Kurfürst [8], and Hässelbarth [9] give some reasons for involving uncorrected bias in uncertainty. The author’s experience, which was gained from evaluation of analytical results personally measured on biological, environmental and food samples for many years, makes the author emphasise mainly these of them: There are situations when the bias is statistically significant and its size is small in comparison with the overall expanded uncertainty U but not entirely negligible. So the analyst supposes the correction of each result to be rather impractical but he does not want to omit it completely. Another and very important reason is that the systematic error of an analytical method is frequently determined using a single RM. The correction value determined corresponds to the interferences arising from its matrix and at the given level of the analyte. Consequently, the analyst is not sure of the correction value in the case of unknown samples that are to be analysed, which are marked with a wider range of the analyte levels and variability of their matrix compositions. (The matrix compositions of the CRM and the real samples are mostly not totally identical; often they are only similar.) More precise studies of the bias value and its uncertainty would demand additional investigations by using other reference materials or spiked samples. Frequently this requirement cannot be fulfil because of lack of trustworthy reference materials and also because of some insufficiencies of spiking studies (e.g. they give information only on relative systematic errors) or because of want of time. In this situation, when the analyst is not sure that the correction does increase the trueness of all the results measured, he rather decides to enlarge the common uncertainty interval for the unsure bias than to correct individually each result for it. (Owing to type II errors in testing the bias for significance, Maroto et al. [12] recommend also the inclusion of an insignificant bias into the expanded uncertainty when its $u(b)$ is larger than 30% of u_c.)

4. Methods incorporating uncorrected bias into the expanded uncertainty

There are more methods incorporating the correction for systematic effects into the uncertainty interval, but no one is perfect. Their defects are e.g. that the levels of confidence associated with such enlarged uncertainty intervals are frequently confused and these intervals are uselessly too wide, since their extensions do not bring about a substantial extension of the level of confidence (the intervals asymmetrically expand mainly into one of the tails of the result probability distribution).

The first method enlarges the uncertainty interval by adding the absolute value of the correction to the expanded uncertainty. Then the modified uncertainty, $U(\text{bias})$[3], is:

$$U(\text{bias}) = U + |b| \tag{10}$$

This oldest method was published by Taylor [2] and Grabe [3]. However, they applied the standard deviation of the mean $\bar{x}$ instead of the overall combined uncertainty and the corresponding value of the Student’s distribution as a value of k. In addition Grabe did not work with an estimation of the correction b, but instead he uses the half-width of the rectangular distribution (symmetric to zero) of possible values of the bias. This 1st method was also mentioned in ISO *Guide* [1] and in other articles [6,8,12].

The second approach [5,9,12] (denoted by RSSu) combines the systematic error correction as a standard uncertainty with the other but genuine uncertainty components; it adds the square bias in the usual root-sum-of-squares:

$$U(\text{RSS}u) = k({u_c}^2 + b^2)^{1/2} \tag{11}$$

The third method (denoted by RSSU) [5,7,10–12] combined the square bias with square expanded uncertainty:

$$U(\text{RSS}U) = (U^2 + b^2)^{1/2} \tag{12}$$

This method represents an enlargement of the bias uncertainty $u(b)$ to a value, with which the determined bias have to be found out insignificant by the t-test (see Eq. (3)) and due to the bias correction is not necessary. In modified Eq. (12) $U(\text{RSS}U) = 2({u_c}^2 + b/2^2)^{1/2}$ the term $b/2$ can represent a judgement estimate of the uncertainty component from the matrix variability and then this approach is very similar to the one mentioned in note 4.

The fourth method (proposed by Phillips and Eberhardt [5] and termed SUMU is based on the principle that the bounds of the normal expanded uncertainty associated with an uncorrected result are corrected for the bias instead of this result. In this way, a new right (upper) and left (lower) uncertainty sides, denoted by U_+ and U_-, result:

$$U_+ = U - b \quad \text{and} \quad U_- = U + b \tag{13}$$

Then the uncorrected result is not in the middle of this uncertainty interval. If $|b| > U$, the lowest applied value of U_+ or U_- is zero (i.e. negative values are replaced by zero). It means that when $|b| \le U$, these intervals are equivalent to the intervals defined by the normal expanded uncertainty about the corrected results. These intervals are asymmetric about the uncorrected results but symmetric about the corrected results, which of course are not calculated.

Another method, published by Eckschlager et al. [4], calculates the expanded uncertainty modified for bias as a confidence interval with a strictly known level of confidence.

[3] *Note* 3: The designations of the given methods are taken from papers [5,12]. They seem to be rather confusing. Consequently in this paper the methods are also denoted with their sequence numbers.

The method considers the confidence level with respect to the mean of the uncorrected result distribution. This attitude is different from the one given in the papers by Phillips and Eberhardt [5] and Maroto et al. [12] and also in this paper, where the confidence level is studied with respect to the mean of the corrected result distribution – this mean is supposed to be the true value (see below). Because of this different attitude this method is not compared with others. (The question, which attitude is better, can be under discussion.) The uncertainty interval obtained with this method is expressed by using the non-central t-distribution, $t(v, \alpha, \delta)$, and the standard deviation of the mean, $s_{\bar{x}}$, (combined standard uncertainty was not used):

$$U_{+} = s_{\bar{x}} t(v, \alpha, \delta) \quad \text{and} \quad U_{-} = -2\delta - s_{\bar{x}} t(v, \alpha, \delta) \tag{14}$$

these equations are valid for positive values of δ, which denotes the true value of the bias (δ is supposed to be determined without any uncertainty). The values of the non-central t-distribution can be computed for given values of δ, v (degrees of freedom) and α (level of significance) according to the equations given in [4,16] (The both published equations differ; the correct equation seems to be only the one in [16].) This uncertainty interval is again asymmetric about the uncorrected result but symmetric about the corrected result.[4]

When the expanded uncertainty modified for the uncorrected bias by one from the four given method is computed, the bias uncertainty $u(b)$ is again included into the overall combined uncertainty (see Eq. (8)) [5,7,10–12]. It means, that in Eqs. (10)–(13) the b functions replace only the unperformed corrections of results (they do not replace this correction uncertainty $u(b)$, though b has been tested for its significance by using $u(b)$ in Eq. (3)). However, Hässelbarth [9] maintains an opposite attitude in his computation [second method – U(RSSu)]. Also in works [2,4] the bias uncertainty is not considered, but because combined uncertainty was not yet applied in chemometrics in that time. If a constant relative systematic error is supposed, Eqs. (10)–(13) are to be taken with relative uncertainties and with $(1-1/R)$ instead of b (if $R \approx 1$ then $1-1/R \approx R-1$).

5. Comparison of the methods

The methods incorporating the uncorrected bias into the expanded uncertainty can be compared according to their above-mentioned defects: levels of confidence (coverage probabilities) and the widths of the uncertainty intervals defined by them [5]. These comparisons are shown in Fig. 1 and Fig. 3. They involved the four mentioned methods and the cases when no bias correction and no modification of the uncertainty for the bias are performed (i.e. intervals defined by the normal expanded uncertainty about uncorrected biased results). These methods are compared with respect to the intervals defined by the normal expanded uncertainties (a coverage factor of 2 – exactly 1.96) about the corrected results. In the appendix of this paper, it is explained that these comparisons can be expressed generally with the only normalized parameter b/u_c. The range of b/u_c is considered from 0 to 2. Larger values are not studied since the author assumes, that in cases of so large biases the results should be either inescapably corrected (when the bias is supposed basically invariable) or maybe treated by the worst-case method (when the bias is supposed variable in wide limits). These comparisons assume, that the distribution of measurement results after the correction for b is normal with a mean equal to the unknowable true value, μ, of the measured quantity and with a standard deviation equal to the overall combined uncertainty u_c.

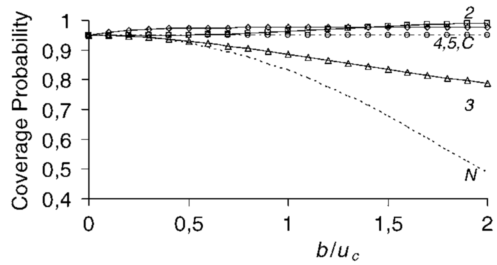

Fig. 1. Coverage probabilities of the uncertainty intervals vs. the bias normalized by u_c: 1 – the first [U(bias)] method; 2 – the second [RSSu] method; 3 – the third [RSSU] method; 4, 5, C – common line for the fourth [SUMU], fifth [U_e(95%)] method and U about corrected results; N–U about uncorrected results.

In Fig. 1, it can be seen relationships between the coverage probability and the parameter b/u_c for the uncertainty intervals of the given methods. In the case of the coverage factor equal to 2, the level of confidence (coverage probability) for the uncertainty intervals defined by the normal expanded uncertainties (U – Eq. (9)) about corrected results (x_c – Eq. (6)) is constantly 95%. The same level is valid also for the intervals given by the fourth (SUMU) method while the levels for the intervals defined by the first [U(bias)] and second [U(RSSu)] are higher than 95%. For the first method the level is higher than 95% from very low values of b/u_c; for b/u_c larger than approximately 0.7 the level reaches a maximum of 97.5% and then it does not change any more. For the second method the level is close to 95% at first, but when b/u_c exceeds a value of about 1.4, the level of confidence oversteps the level for the first method. The level of confidence for the third (RSSU) method is lower than 95%, especially when b/u_c exceeds a value of about 0.5. Of course, the lowest level of confidence

[4] *Note* 4: There is another method of incorporating the bias into the uncertainty instead of a correction for it but the author takes it to be basically different from the previous methods and hence he did not involve it in this summary. In contrast to those previous methods the bias is not assumed constant for a group of samples, but according to the matrix effects of the measured samples it is assumed to vary from zero to a maximum value, which corresponds to the worst-case from possible matrix effects [10,13]. The bias evaluated on a worst-case sample is supposed to be the extreme of the probability distribution (it is supposed rectangular or triangular) of the possible bias values. This method treats systematic errors, constant only for individual samples, like randomly variable errors.

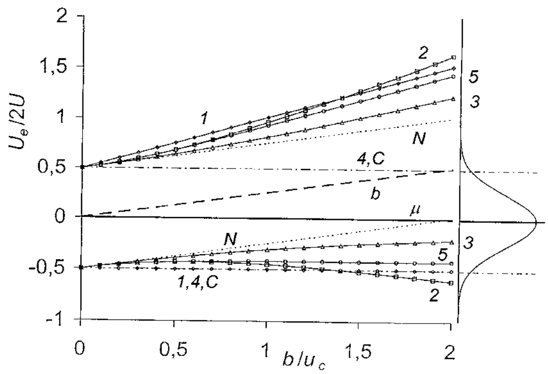

Fig. 2. Relationships between the layouts of the uncertainty intervals and the bias normalized by u_c; the values on the y-axis are centralized by μ and divided by $2U$. The area delimitated by the bounds of an interval investigated on the probability distribution of corrected results with mean μ (the density function of this distribution is plotted along the y-axis) illustrates the probability that this interval contents μ. The plotted lines represent: μ – mean of the probability distribution of the corrected results (zero line); b– function $\mu + b$, about which the modified expanded uncertainties are plotted (after the centralization it equals b); 1 – upper bounds of the first [U(bias)] method; 1, 4, C – common lower bounds of the first [U(bias)] and the fourth method [SUMU] and U about corrected results, 2 – bounds of the second [RSSu] method; 3 – bounds of the third [RSSU] method; 4, C – common upper bounds of the fourth [SUMU] method and U about corrected results, 5 – bounds of the fifth [U_e(95%)] method; N – bounds of U about uncorrected results.

is obtained for the intervals defined by the normal expanded uncertainties about uncorrected results. When the methods are compared from this point of view, it should be stressed, that the level of confidence associated with the interval defined by the normal expanded uncertainty is usually rather uncertain and also that in most practical situations it does not make sense to try to distinguish between an interval having a level of confidence of 95% and either a 94% or 96% interval (see ISO *Guide* [1]).

In Fig. 2 it can be seen how the uncertainty intervals studied are spread about the means μ (i.e. true values) of the probability distributions of corrected results and what parts of these distributions they cover[5]. These distributions are represented by the common standardised normal distribution, whose density function is plotted along the y-axis. The quantities plotted along the y-axis are centralized by subtraction of the means μ, hence the zero line represents μ. Their values are expressed in comparison with the widths of the normal uncertainty intervals, i.e. divided by $2U$ ($2U$ is thought to be a reference value). The line denoted by b represents values of $\mu + b$ (after the centralization only values of b are plotted). This plot shows the cases of positive values of the biases (in the cases of negative values a mirror picture would be drawn). The function $\mu + b$ expresses the means of the uncorrected result distributions (see Appendix A – Mathematical background), about which the lower and upper bounds of the compared intervals are plotted, when the correction is omitted (see Eqs. (27) and (28): upper bounds $(b + U_{e+})/2U$ and lower bounds $(b-U_{e-})/2U$). These bounds delimitate areas from the whole unit area under the drawn density distribution function; these areas indicate the coverage probabilities of the uncertainty intervals displayed in Fig. 1).

The intervals for the fourth method (SUMU) are identical with the normal uncertainty intervals of corrected results, which are defined by U about the line μ (zero line) – see Eqs. (26) and (30). These intervals are symmetric about μ and hence asymmetric about $\mu + b$. As the value of b/u_c increases, the left sides of these intervals expand and cover always an increasing part of the corrected result distribution, while the right sides decrease until b/u_c equals 2 (exactly 1.96), when the right side disappeared and the interval lies only under b. The intervals for other methods are symmetric about b, but their coverage probabilities are spread asymmetrically about b. When the values of b/u_c increase the common middle (line b) of these intervals deviate from μ and consequently the intervals expand mainly into the right (upper) tail of the probability distribution, so their expansions are not connected with considerable enlargements of the coverage probability. In the cases of the entirely omitted correction, the intervals ($b \pm U$) shift very quickly from the centre of the probability distribution as the values of b/u_c increase. Hence when b/u_c approaches 2, the coverage probability goes down 50%.

In Fig. 3 the total widths of the intervals are displayed (expressed again in comparison with $2U$). For a range of b/u_c from 0 to 1.4 the widest intervals are obtained by the first method [U(bias)], for larger values by the second method [RSSu]. When b/u_c oversteps 2, the intervals given by the latter method become considerably wider (it is not illustrated). The intervals of the third method [RSSU] are fairly narrower owing to their low values of the coverage probability. The intervals for the fourth method [SUMU], as they are identical with the normal uncertainty intervals, are narrowest and have the invariable width equal to 1 (expressed relatively to

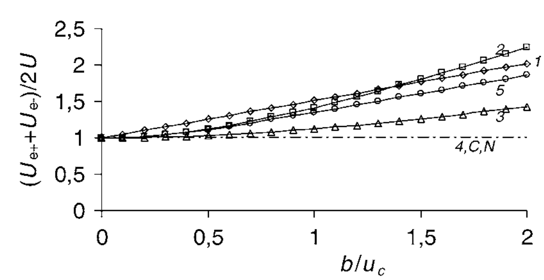

Fig. 3. Widths of the uncertainty intervals normalized by $2U$ vs. the bias normalized by u_c: 1 – the first [U(bias)] method; 2 – the second [RSSu] method; 3 – the third [RSSU] method; 4, C, N – common line for the fourth [SUMU] method, U about corrected results and U about uncorrected results, 5 – the fifth [U_e(95%)] method.

[5] *Note* 5: The uncertainty intervals are spread about measured values (x_i) and the content with a high probability (e.g. 95%) the true values (μ): $x_i-U_{e-} \leq \mu \leq x + U_{e+}$ (see Appendix A). However, the figures used to illustrate this problem and to show the layouts of these intervals express an opposite case (see e.g. [18]), i.e. $\mu-U \leq x \leq \mu + U$. The probabilities of both inequalities are identical (see Appendix A – Mathematical background).

Table 1

b/u_c	E	b/u_c	E
0.1	0.098	1.1	0.714
0.2	0.193	1.2	0.738
0.3	0.284	1.3	0.758
0.4	0.368	1.4	0.775
0.5	0.443	1.5	0.790
0.6	0.509	1.6	0.803
0.7	0.566	1.7	0.815
0.8	0.613	1.8	0.825
0.9	0.653	1.9	0.834
1.0	0.686	2.0	0.842

Relationship between the coefficient E enlarging expanded uncertainty for the bias, and the values of b/u_c; the modified expanded uncertainty defined by the fifth method: $U_e(95\%) = 1.96\ u_c + E\ |b|$ about uncorrected results provides uncertainty intervals with a coverage probability of 95%.

2U), as well as the intervals defined by U about uncorrected results of course (entirely omitted correction).

6. Proposal of a new method

A new (fifth) method can be proposed, which defines uncertainty intervals enlarged for the bias, symmetric about the measured uncorrected results and having just 95% level of confidence with respect to the mean of the corrected result distribution. For given values of b and u_c the expanded uncertainty defined by this method, $U_e(95\%)$, can be expressed like this:

$$U_e(95\%) = U + E|b| = ku_c + E|b| \quad (15)$$

where the coefficient E depends on b/u_c and the level of confidence (for the 95% level of confidence see Table 1). The forms of this mathematical expression can be different but with the same widths and layouts of these intervals. The author chose the form of Eq. (15), where $U_e(95\%)$ is calculated as a function of b, because it expresses the causality. This method need not be regarded as a new method, but only as an optimum when the 95% coverage probability is taken as the optimality criterion for the uncertainty intervals enlarged for the uncorrected bias and symmetric about uncorrected results.

For $b/u_c > 0.8$ the right bounds of these intervals are so high that practically 100% of the values of the distributions for corrected results are lower (see Fig. 2), so that the left bounds must define only 5% of lowest values, in order that the intervals between them would encompass 95% of all possible values. Hence, the left sides of these intervals must consist of the correction for the bias (b) and the 95% quantille of the corrected result distribution (i.e. 1.65 u_c):[6]

$$U_e(95\%) = 1.65u_c + |b| \quad (16)$$

[6] *Note* 6. For similar reason the first [U(bias)] method defines intervals with 97.5% coverage probabilities when b/u_c oversteps a value of 0.7 (97.5% quantille is 1.96 u_c), see Fig. 1.

The layouts of these uncertainty intervals and their total widths change with b/u_c as displayed in Figs. 2 and 3. It can be seen that these intervals are always narrower than the ones given by the first [U(bias)] and second [RSSu] methods. However, for values of b/u_c lower then 0.5 the intervals obtained by the fifth and second methods are practically identical. For values of b/u_c large than 0.8 the differences of the widths of the first and fifth method intervals are about (1.96–1.65)u_c and so the larger values of b/u_c are, the smaller are the relative differences of these widths. (When b/u_c equals 0.8 and 2, the relative difference is 12.9 and 8.6%, respectively – the uncertainty of the fifth method taken as 100%.)

7. Conclusions

The method modifying expanded uncertainties for uncorrected bias, which provides narrowest uncertainty intervals with 95% levels of confidence, is the fourth (SUMU) method, so that it is the best, as Phillips and Eberhardt [5] have already stated. Maybe because these uncertainty intervals are asymmetric, it has not been often applied in evaluating uncertainties of analytical results. From the methods compared, the first and second [U(bias) and RSSu] methods give widest intervals with 95% and higher levels of confidence. These two methods should be applied in combination, the former for values of b/u_c lower than 1.4, the latter for larger values. The third (RSSU) method is suitable only for very low values of b/u_c, otherwise provides intervals with levels of confidence considerably less than 95%. The combination of RSSu and U(bias) methods can be supplied by the fifth, newly proposed method [$U_e(95\%)$], which seems to be better than the given combination, especially for values of b/u_c from 1 to 2.

The results uncorrected for a significant bias are biased point estimates. However, the interval estimates defined about these points by an uncertainty modified for the uncorrected bias can cover the true estimated values with a sufficiently high probability. This probability can be evaluated by using the values of b and u_c or the interval can be constructed in order that its coverage probability would be just 95%. This problem of coverage probability continues when an uncorrected result with such a modified uncertainty is used in computation of a follow-up result and its combined uncertainty. The value of the standard uncertainty, which would be recovered from an uncertainty enlarged for bias by dividing this uncertainty with the coverage factor ($u = U_e/k$), is not an appropriate estimate of the standard deviation of the distribution of the uncorrected result. The follow-up result obtained will be biased and its uncertainty interval defined by the expanded uncertainty computed by using the usual approach (combining that problematic standard uncertainty with other uncertainty contributions as a square root of the sum of variances) will have a questionable value of coverage probability again. In order that an uncertainty interval with a determinable coverage probability could be obtained, the original standard uncertainty [$u(b)$] and the bias (b) would have to be used again. This proceeding would be sensible if the obtained uncertainty interval were associated with a large group of results.

There is a question whether these uncorrected results are traceable. Traceability can be established e.g. by the procedures mentioned in the chapter dealing with the estimation of bias, by using CRM, by spiking a sample or by using an accepted, closely defined procedure. I think that if a corrected result is traceable to a particular reference with an unbroken chain of comparisons also the corresponding uncorrected result with its uncertainty interval enlarged for the correction is. However, the links in the chain are less close since this uncertainty is larger. The traceability allows the results to be compared one with other or one with standard values (values stated without any uncertainty). The second comparison means to find out whether the uncertainty interval contents the standard value (analogy with the one sample t-test). It would be without any problem in the cases of these enlarged uncertainty intervals. But the comparison of two results would be more difficult (it is analogous with the t-test comparing two samples). It would be necessary again to return to the origin standard uncertainty and bias.

Appendix A. Mathematical background

The graphs plotted in Figs. 1–3 express function relations from the fraction b/u_c to the coverage probabilities, layouts and widths, respectively, of the uncertainty intervals investigated. These relations were obtained provided that both measured quantity x and bias b have normal distributions. They can be expressed generally, since each normal distribution, and also the normal distribution of the quantity x_c (corrected result) with a population mean μ (expectation) and variance u_c^2 (this distribution is denoted by $N[\mu, u_c^2]$), can be modified to the common standard normal distribution with a zero mean and a unit variance ($N[0,1]$). The standardizing equation is (see e.g. [17,18])

$$z = \frac{x_c - \mu}{u_c} \tag{17}$$

Values of the distribution function (cumulative) for a normal distribution, F, can be obtained using the distribution function for the standardised normal distribution, Φ, which are given in statistical tables (e.g. in [17]), since $F(x_c) = \Phi(z)$ for a pair of the corresponding values of x_c and z.

First it is necessary to create an error model of results measured. The quantity x is measured on a routine sample with a constant systematic error δ. The value of an ith result observed, x_i, is

$$x_i = \mu + \delta + \varepsilon_i \tag{18}$$

where μ is the mean of the distribution of the corrected results (the unknowable true value of the measurand of the sample) and ε_i is the random error of the ith observation, which is taken from a distribution $N[0, s^2]$. The variance s^2 can be estimated using the square uncertainty components given in Eq. (8) but without the component $u^2(b)$ (this component will join after the result correction for the systematic error – see below):

$$s^2 = u^2(x_1) + u^2(x_2) + \cdots \tag{19}$$

A correction for the systematic error is performed according to Eq. (6) employing a determined value of b (see Eq. (1)). This value can be just the jth measured value, b_j, randomly taken from a distribution of all possibly determinable values of b ($N[\delta, u^2(b)]$):

$$b_j = \delta + \eta_j \tag{20}$$

where η_j is the jth value of the random error, η, of the bias determination; the distribution of η is $N[0, u^2(b)]$.

The corrected result, x_{cij}, is obtained from values x_i and b_j according to Eq. (6):

$$x_{cij} = x_i - b_j \tag{21}$$

which can be modified using Eqs. (18) and (20) into this form:

$$x_{cij} = \mu + \delta + \varepsilon_i - \delta - \eta_j = \mu + \varepsilon_i - \eta_j \tag{22}$$

Because the errors ε_i and η_j are from the distributions $N[0, s^2]$ and $N[0, u^2(b)]$, their difference $(\varepsilon_i - \eta_j)$ is from the distribution $N[0, u_c^2]$, (u_c see Eq. (8)). It can be seen that the distribution of the corrected results is normal with the mean μ and variance u_c^2 ($N[\mu, u_c^2]$). Since according to Eqs. (21) and (22) $x_i = x_{cij} + b_j = \mu + b_j + \varepsilon_i - \eta_j$, the distribution of the uncorrected results x_i has a mean of $\mu + b_j$ and a variance of u_c^2.

Now the measured quantity after the correction, x_c, can be standardized (see Eqs. (17) and (22)); for the corrected result x_{cij} it means:

$$z_{ij} = \frac{x_{cij} - \mu}{u_c} = \frac{\varepsilon_i - \eta_j}{u_c} \tag{23}$$

Using the corrected result x_{cij} and the expanded uncertainty U (Eq. (9)) one can define the normal uncertainty interval $x_{cij} \pm U$, in which μ lies with a probability p (coverage probability):

$$p = P\{x_{cij} - U \le \mu \le x_{cij} + U\} \tag{24}$$

The inequality in the brackets can be modified and this equation results:

$$p = P\left\{\frac{-U}{u_c} \le \frac{x_{cij} - \mu}{u_c} \le \frac{U}{u_c}\right\} \tag{25}$$

Since the distribution of the function $(x_c - \mu)/u_c$ is $N[0, 1]$ (see Eq. (23)) and U/u_c equals k (see Eq. (9)), the probability p can be computed:

$$p = \Phi(k) - \Phi(-k) \tag{26}$$

(Confidence interval is defined and explained in each elementary statistics and also in [1,18]).

Now the uncertainty intervals defined by uncorrected results and the expanded uncertainties modified for biases can be investigated. Because the widths of the left and right sides of these uncertainty intervals can be different, they are denoted by U_{e-} and U_{e+} (generally for all the method investigated). The coverage probability p that an uncertainty interval modified for bias covers μ can be expressed:

$$p = P\{x_i - U_{e-} \le \mu \le x_i + U_{e+}\} \quad (27)$$

The inequality in the brackets can be successively modified into this form: $p = P\{-b_j/u_c + U_{e-}/u_c \ge (x_{cij}-\mu)/u_c \ge -b_j/u_c - U_{e+}/u_c\}$. After replacement of $(x_{cij}-\mu)/u_c$ by z_{ij} and multiplication by −1 one can obtain: $p = P\{b_j/u_c - U_{e-}/u_c \le -z_{ij} \le b_j/u_c + U_{e+}/u_c\}$; the quantity z has distribution $N[0,1]$ that is symmetric about zero, so that $-z$ has the same distribution, consequently:

$$p = P\left\{\frac{b_j}{u_c} - \frac{U_{e-}}{u_c} \le z_{ij} \le \frac{b_j}{u_c} + \frac{U_{e+}}{u_c}\right\} \quad (28)$$

The coverage probability p can be computed:

$$p = \Phi\left(\frac{b_j}{u_c} + \frac{U_{e+}}{u_c}\right) - \Phi\left(\frac{b_j}{u_c} - \frac{U_{e-}}{u_c}\right) \quad (29)$$

The general uncertainties U_{e+} and U_{e-} in Eqs. (28) and (29) can be replaced by the specific ones according to (Eqs. (10)–(13)) and (Eq. (15)) for the methods compared and by U in the case of the completely omitted correction (interval $x_i \pm U$). The left and right terms in the inequality (Eq. (28)) define the interval bounds in the standardized form. For positive values of b the following bounds can be obtained (the sign is important for the first and fifth method):

First method [U(bias)]:

$$\frac{b}{u_c} - \frac{(ku_c + b)}{u_c} \quad \text{and} \quad \frac{b}{u_c} + \frac{(ku_c + b)}{u_c},$$

$$\text{i.e.} \; -k \quad \text{and} \quad \frac{2b}{u_c} + k$$

Second method [RSSu]:

$$\frac{b}{u_c} - \frac{k({u_c}^2 + b^2)^{1/2}}{u_c} \quad \text{and} \quad \frac{b}{u_c} + \frac{k({u_c}^2 + b^2)^{1/2}}{u_c},$$

$$\text{i.e.} \; \frac{b}{u_c} - k\left[1 + \left(\frac{b}{u_c}\right)^2\right]^{1/2} \quad \text{and} \quad \frac{b}{u_c} + k\left[1 + \left(\frac{b}{u_c}\right)^2\right]^{1/2}$$

Third method [RSSU]:

$$\frac{b}{u_c} - \frac{[(ku_c)^2 + b^2]^{1/2}}{u_c} \quad \text{and} \quad \frac{b}{u_c} + \frac{[(ku_c)^2 + b^2]^{1/2}}{u_c},$$

$$\text{i.e.} \; \frac{b}{u_c} - \left[k^2 + \left(\frac{b}{u_c}\right)^2\right]^{1/2} \quad \text{and} \quad \frac{b}{u_c} + \left[k^2 + \left(\frac{b}{u_c}\right)^2\right]^{1/2}$$

Fourth method [SUMU]:

$$\frac{b}{u_c} - \frac{ku_c + b}{u_c} \quad \text{and} \quad \frac{b}{u_c} + \frac{ku_c - b}{u_c},$$

$$\text{i.e.} \quad -k \quad \text{and} \quad k \quad (30)$$

Normal uncertainty about an uncorrected result:

$$\frac{b}{u_c} - \frac{ku_c}{u_c} \quad \text{and} \quad \frac{b}{u_c} + \frac{ku_c}{u_c}, \quad \frac{b}{u_c} - k \quad \text{and} \quad \frac{b}{u_c} + k$$

Fifth method: [U_e(95%)] :

$$\frac{b}{u_c} - \frac{ku_c + Eb}{u_c} \quad \text{and} \quad \frac{b}{u_c} + \frac{ku_c + Eb}{u_c},$$

$$\text{i.e.} \quad \frac{b}{u_c}(1 - E) - k \quad \text{and} \quad \frac{b}{u_c}(1 + E) + k$$

It can be seen that all the bounds obtained are functions only of b/u_c and k.

The graph of the coverage probabilities versus b/u_c (Fig. 1) was calculated using (Eq. (29)) and the specific bounds given above. The graph of the dispositions of the studied intervals (Fig. 2) was obtained using Eq. (28). The y-axis was centralized by the distribution mean μ and normalized by $2U$ [upper bounds $(b + U_{e+})/2U$ and lower bounds $(b - U_{e-})/2U$]. The widths of the studied uncertainty intervals illustrated in Fig. 3 equal $U_{e-} + U_{e+}$. The values of the coefficient E (Table 1) were computed numerically using Eq. (29).

Appendix B. Illustrative application of the equations for the methods studied

Cadmium in drinking water samples was determined in 10 runs during about 6 month using ET AAS method. CRM Trace Elements in Water 1643d was tested with each run for AQC. The data obtained: certified content of Cd in the CRM x_{ref} = 6.47 ng ml^{-1} with expanded uncertainty U_{ref} = 0.37 ng ml^{-1}, mean of p = 10 determined results[7] $\overline{\chi}$ = 6.154 ng ml^{-1} with sample standard deviation s_r = 0.221 ng ml^{-1} (s_r is intermediate precision).

The calculated results:

Bias (Eq. (1)) b = 6.154−6.47 =−0.316 ng ml^{-1}; standard uncertainty of certified content $u(x_{ref}) = U_{ref}/k$ = 0.37/2 = 0.185 ng ml^{-1}; bias uncertainty (Eq. (2)) $u(b) = (0.221^2/10 + 0.185^2)^{1/2}$ = 0.198 ng ml^{-1}; overall combined standard uncertainty (Eq. (8)) $u_c = (0.198^2 + 0.221^2)^{1/2}$ = 0.297 ng ml^{-1}; test for b significance (Eq. (2)) $|b|$ = 0.316 $\le k\, u(b) = 2 \times 0.198$ = 0.396. The bias is insignificant hence it is not necessary to correct future measured results for it. However, $u(b)$ is larger than 30% of u_c and Maroto et al. [12] recommends enlarging the expanded uncertainty in this case. Due to the obtained results can be used in illustrative computations of the expanded uncertainties enlarged for the bias (k = 2):

Normal expanded uncertainty (Eq. (9)): U = 2·0.297 = 0.59 ng ml^{-1}

[7] *Note* 7: The results were obtained by the author and his colleagues in the analytical laboratory of Regional Hygiene Station in Usti nad Labem from February to July 2000.

The first method [U(bias)] (Eq. (10)): $U(\text{bias}) = 0.594 + 0.316 = 0.91\,\text{ng ml}^{-1}$
The second method [RSSu] (Eq. (11)): $U(\text{RSS}u) = 2 \times (0.297^2 + 0.316^2)^{1/2} = 0.87\,\text{ng ml}^{-1}$
The third method [RSSU] (Eq. (12)): $U(\text{RSS}U) = (0.594^2 + 0.316^2)^{1/2} = 0.67\,\text{ng ml}^{-1}$
The fourth method [SUMU] (Eq. (13)): $U_+ = 0.594-(-0.316) = 0.91\,\text{ng ml}^{-1}$ and $U_- = 0.594 + (-0.316) = 0.27\,\text{ng ml}^{-1}$
The fifth method [U_e(95%)] (Eq. (15)): $b/u_c = 1.06$; $E \approx 0.70$ (see Table 1)
$U_e(95\%) = 0.594 + 0.70 \times 0.316 = 0.81\,\text{ng ml}^{-1}$
Since $b/u_c > 0.8$, Eq. (16) is also applicable: $U_e(95\%) = 1.65 \times 0.297 + 0.316 = 0.81\,\text{ng ml}^{-1}$

The characteristics of the above computed intervals can be found in Figs. 1–3 for $b/u_c = 1.06$. Since the bias is negative, the right image in Fig. 2 would be the mirror one.

References

[1] ISO, Guide to the Expression of Uncertainty in Measurement, International Organisation for Standardisation, Geneva, 1993.
[2] J.K. Taylor, Anal. Chem. 53 (1981) 1588A.
[3] M. Grabe, Metrologia 23 (1986–1987) 213.
[4] K. Eckschlager, I. Horsák, Z. Kodejs, Z. Ksandr, M. Matherny, I. Obrusník, S. Wicar, Applications of computers in analytical chemistry, in: G. Svehla (Ed.), Wilson and Wilson's Comprehensive Analytical Chemistry, vol. 18, Elsevier, Amsterdam, 1983, p. 271.
[5] S.D. Phillips, K.R. Eberhardt, J. Res. Natl. Inst. Technol. 102 (1997) 577.
[6] R. Wood, Recovery, second Eurachem Workshop on Measurement Uncertainty in Chemical Analysis, Berlin, 29–30 September 1997.
[7] S.L.R. Ellison, Accred. Qual. Assur. 3 (1998) 95.
[8] U. Kurfűst, Accred. Qual. Assur. 3 (1998) 406.
[9] W. Hässelbarth, Accred. Qual. Assur. 3 (1998) 418.
[10] V.J. Barwick, S.L.R. Ellison, Analyst 124 (1999) 981.
[11] V.J. Barwick, S.L.R. Ellison, Accred. Qual. Assur. 5 (2000) 47.
[12] A. Maroto, R. Boqué, J. Riu, F.X. Rius, Accred. Qual. Assur. 7 (2002) 90.
[13] L.R. Ellison, M. Rősslein, A. Williams (Ed.), EURACHEM/CITAC Guide, Quantifying Uncertainty in Analytical Measurement, second ed., 2000.
[14] A. Maroto, R. Boqué, J. Riu, F.X. Rius, Anal. Chim. Acta 440 (2001) 171.
[15] A. Maroto, R. Boqué, J. Riu, F.X. Rius, Anal. Chim. Acta 446 (2001) 133.
[16] K. Eckschlager, Chemometrie 2, Prírodovedecká fakulta Univerzity Karlovy, Prague, 1994, p. 167.
[17] H.J. Bartch, Matematické vzorce, second ed., SNTL, Prague, 1987, p. 743.
[18] M. Meloun, J. Militky, M. Forina, Chemometrics for analytical chemistry, vol. 1, in: PC-Aided Statistical Data Analysis, Ellis Horwood, Chichester, 1992.