Available online at www.sciencedirect.com
ScienceDirect
Journal of Chromatography A, 1158 (2007) 15–24
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Review

Measurement uncertainty

Veronika R. Meyer *

EMPA St. Gallen, Swiss Federal Laboratories for Materials Testing and Research, Lerchenfeldstrasse 5, CH-9014 St. Gallen, Switzerland

Available online 1 March 2007

Abstract

Measurement uncertainty is a statistical parameter which describes the possible fluctuations of the result of a measurement. It is not a mere repeatability but it is at least as high as the intra-laboratory reproducibility. If it is an attribute of a general analytical test procedure it is at least as high as the inter-laboratory reproducibility. Measurement uncertainty can be determined by the addition of the variances of the individual steps of the test procedure or by an approach which starts with one of the above-mentioned reproducibilities. Any measurement uncertainty should be kept low but it is objectionable to state too low a value, e.g. by falsely reporting mere repeatability data instead of properly determined uncertainty data. Some good working principles can help to obtain low measurement uncertainties.

Keywords: Measurement uncertainty; Uncertainty of chromatographic analyses; Standard uncertainty; Influence parameters of measurement uncertainty; Calculation rules for measurement uncertainty

Contents

* Tel.: +41 71 274 77 87; fax: +41 71 274 77 88.
E-mail address: veronika.meyer@empa.ch.

0021-9673/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2007.02.082

1. Introduction

Measurement results must be obtained under well-defined conditions. Valid results, used as a base for decisions, trade, legal actions or publications, must not be a stroke of luck but they should get their authenticity by a carefully determined figure of merit. The characteristic needed for all kinds of measurements is the so-called measurement uncertainty. Its necessity was realized earlier in physics than in chemistry, therefore an important document, the so-called "GUM", was published in 1993 (corrected and reprinted in 1995) with a number of detailed examples taken from physical measurement problems [1]. In paragraph 2.2.3 it defines the term "measurement uncertainty" as follows:

> "Parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand."

Based on the principles described in the GUM, a similar document for analytical chemistry came out in 1995, called the "QUAM"; the now valid second edition is from 2000 [2]. It can be downloaded from the Internet free of charge and all persons in the company which are involved in quality management should have a personal copy. In addition, the "GUM" should at least be found in the library because it gives more mathematical background and is the relevant document in difficult or questionable cases.

The measurement uncertainty datum is the spouse of a result because no industrial, commercial or governmental laboratory will produce results just for the fun of it. A result will be judged or compared with others and will lead to a decision, e.g. pass versus fail, accept versus reject, implement an action or not. Two possible scenarios are shown in Fig. 1. One of them illustrates the situation with a legal limit of the concentration of a pesticide in food. In order to protect the consumers the commodity must be rejected even if the mean analytical result is below the limit in a case where the sum of result and uncertainty is higher than the limit. The other illustration shows the comparability of results found by different laboratories (such as producer and customer, pharmaceutical industry and drug control etc.). The two parties involved will not have an argument if their results plus uncertainties have some overlap.

The norm ISO 17025, "General requirements for the competence of testing and calibration laboratories", demands the determination of the measurement uncertainty of analytical results in paragraph 5.4.6 [3].

2. Top-down and bottom-up approach for the determination of the measurement uncertainty

It is well known that repeatability, described as a standard deviation, a relative standard deviation or a coefficient of variation, covers only a limited part of the possible fluctuation of an analytical result. Moreover, one will not get the same standard deviation if the same sample solution is injected consecutively several times into the chromatograph or if the whole sample preparation is repeated several times, resulting in n different sample solutions (n being at least 3 but preferably higher). The latter standard deviation will be higher because now also all the little fluctuations of weighing, dilution, recovery etc. will add to the standard deviation of the chromatographic instrument and process. The highest intra-laboratory repeatability, the so-called intermediate precision [4], is obtained by repeating the analysis on different days, with different instrumentation and so on. The standard deviation is even higher if several laboratories at different locations are involved; this value is called the (inter-laboratory) reproducibility, see Fig. 2. It is high because the different instruments involved (not only chromatographs but also pipets, etc.) differ in their calibration and repeatability and because the different people have their individual working style. The reproducibility can be identical with the measurement uncertainty of the test procedure if the purity of the reference compound is not an important parameter of the uncertainty budget (see paragraph 4.3 below). This approach which obtains a measurement uncertainty from reproducibility is called the top-down method.

The opposite is the bottom-up approach, calculating the uncertainty by the addition of variances:

$$u_c(M) = M\sqrt{\left(\frac{u(a)}{a}\right)^2 + \left(\frac{u(b)}{b}\right)^2 + \left(\frac{u(c)}{c}\right)^2 + \left(\frac{u(d)}{d}\right)^2} \quad (1)$$

for a measurand M (i.e. an analytical result) if the equation for the calculation of M is based only on multiplications and divisions such as $M = (a \times b \times c)/d$ (see paragraph 3.5 below). $u(x)$ is the standard uncertainty of factor x and $u_c(M)$ is the combined uncertainty.

The bottom-up approach can be tedious if the calculations are done with a pocket calculator. On the other hand, it is illustrative and forces the analyst to understand the details of the test procedure in question. Therefore, the bottom-up approach is a tool which identifies the parameters or working steps with an overly contribution to the total uncertainty, thus offering the chance to improve them.

3. Some tools for the determination of measurement uncertainty

3.1. Flow diagram and Ishikawa diagram

For the correct determination of measurement uncertainty, and especially for the bottom-up approach, it is essential that the analyst understands the analytical test procedure. It should be described in detail in a standard operating procedure (SOP). If this was done correctly it is not difficult to draw a flow diagram of the individual working steps of the method. (If this task is difficult the SOP is not complete or not clear, resulting in frustrated personnel.) Drawing flow diagrams is not mandatory but very helpful. Fig. 3 shows the diagram of a simple analysis with one-point calibration.

An Ishikawa diagram or cause-and-effect diagram (sometimes also termed fishbone diagram) is a useful tool to identify the influence parameters, i.e. the sources of uncertainty, of a whole test procedure or of a single working step [5–7]. By

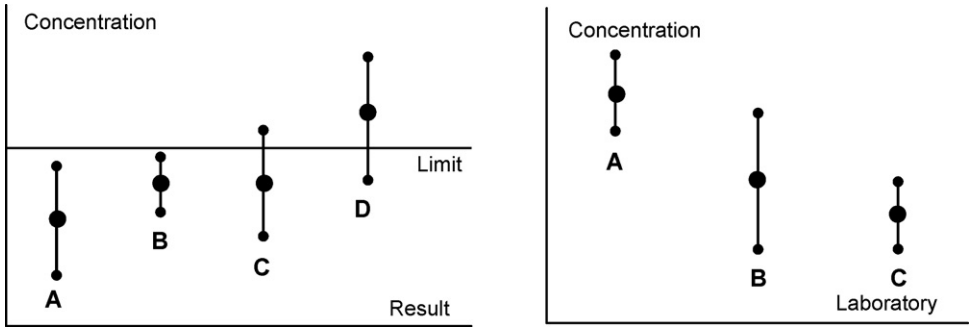

Fig. 1. A result without measurement uncertainty can lead to problems. Left: results A and B are below the maximum allowed pesticide concentration even when their measurement uncertainty is taken into consideration. The mean of result C is below the limit but there is a certain possibility that the maximum allowed concentration is transgressed because of its uncertainty; the food sample will be rejected. Result D has the same consequence although there is some possibility that the limit is not transgressed. Right: A certain sample is analysed by three different laboratories. A and B will accept the results of each other, and so will do laboratories B and C. However, laboratories A and C will run into an argument.

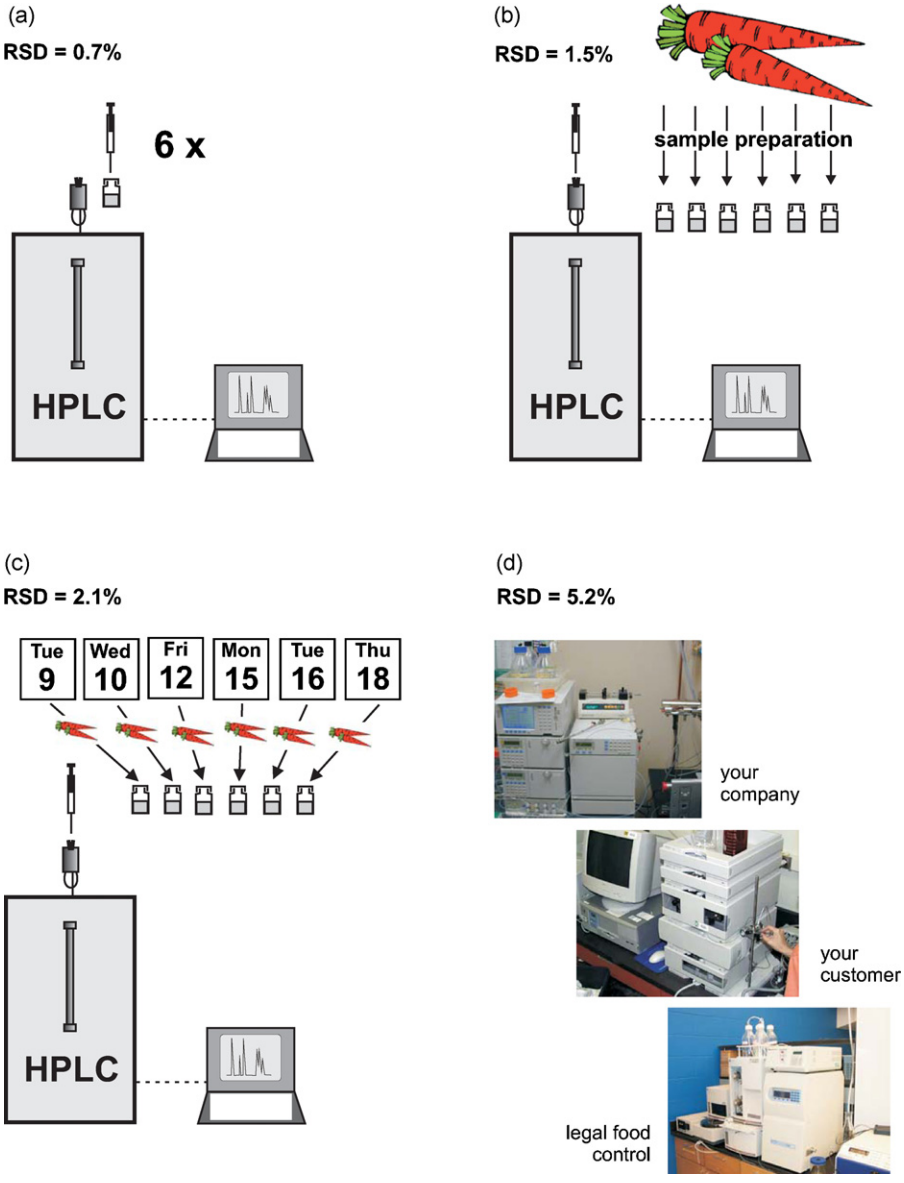

Fig. 2. The six-fold injection of the same sample (a) will yield a lower repeatability (measured as standard deviation) than the six-fold preparation of the same sample (b). If the six preparations are performed on different days (c) the standard deviation will be even higher and the number is called “intermediate precision” (or “day-to-day precision” in this case). If the same sample is investigated in different laboratories (d) the resulting standard deviation is highest and is called “reproducibility”.

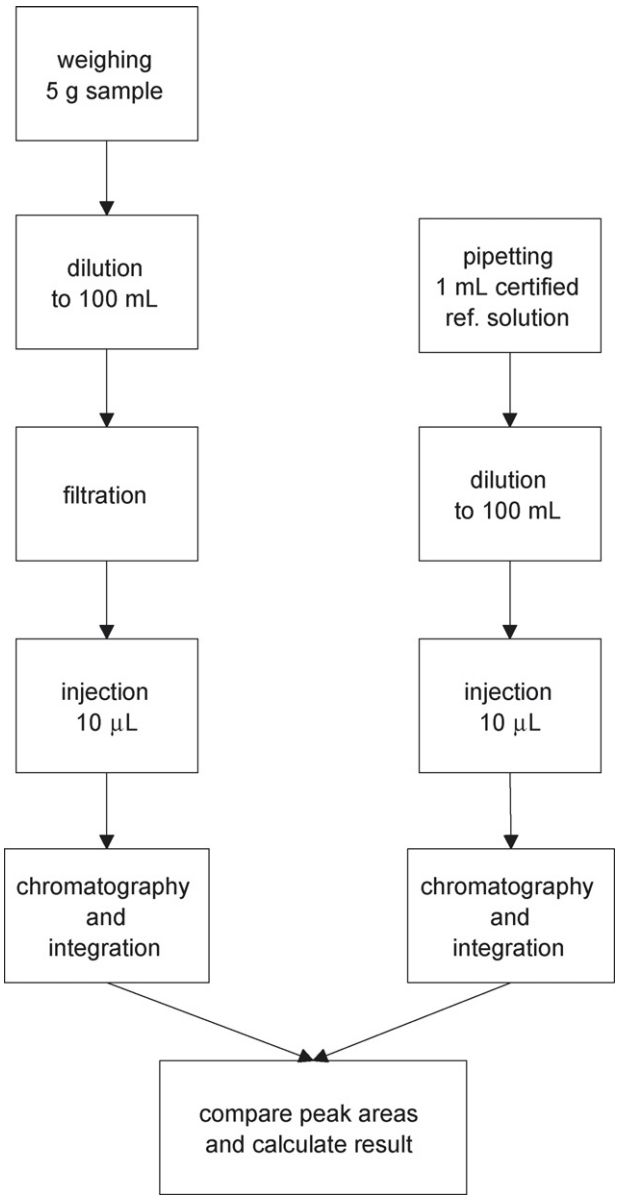

Fig. 3. The flow diagram of a simple chromatographic analysis.

drawing such a structure one can identify, sort and discuss these parameters. Arrows of first, second and if necessary also higher order point from causes (e.g. the operating repeatability of a pipet) to effects (e.g. a volume or, finally, to the analytical result). Fig. 4 is the Ishikawa diagram of the process shown in Fig. 3.

3.2. Equation of the measurand

The measurand, i.e. the equation for the calculation of the analytical result, must be noted in detail. Usually a calculation is not done with the operation (peak area of sample/peak area of reference) × (concentration of reference solution). The equation for the analysis shown in Fig. 3 is as follows:

$$c_S = \frac{V_{\mathrm{Pip}R} \times c_R \times P_R \times V_{\mathrm{Flask}S} \times A_S}{V_{\mathrm{Flask}R} \times m_S \times A_R} \quad (2)$$

with c = concentration, V_{Pip} = volume of a pipet, V_{Flask} = volume of a measuring flask, P = purity, A = peak area, m = mass; the index S refers to the sample and the index R to the reference.

The detailed equation of the measurand must be noted in the SOP. In fact, it is the starting point of the Ishikawa diagram. Although such a diagram can be drawn by free association it should afterwards be compared with the equation. Every factor of the equation must appear in the diagram. (The Ishikawa diagram may show more arrows than factors.)

3.3. Standard deviations

Preliminary remark: the standard uncertainties $u(x)$ in Eq. (1) can be standard deviations but they can also be of another type of distribution. Standard deviations are determined by the n-fold repeat of a measurement; they are a characteristic of normally (Gaussian) distributed data, the typical feature of results which show some accidental fluctuations when the experiment is repeated. If the standard uncertainties cannot be obtained by experiments they are assigned another type of distribution, see paragraph 3.4 below.

By the repeated determination of an influence parameter of measurement uncertainty one gets its standard deviation $s(x)$ which describes the scatter of the results: approx. 68% will lie

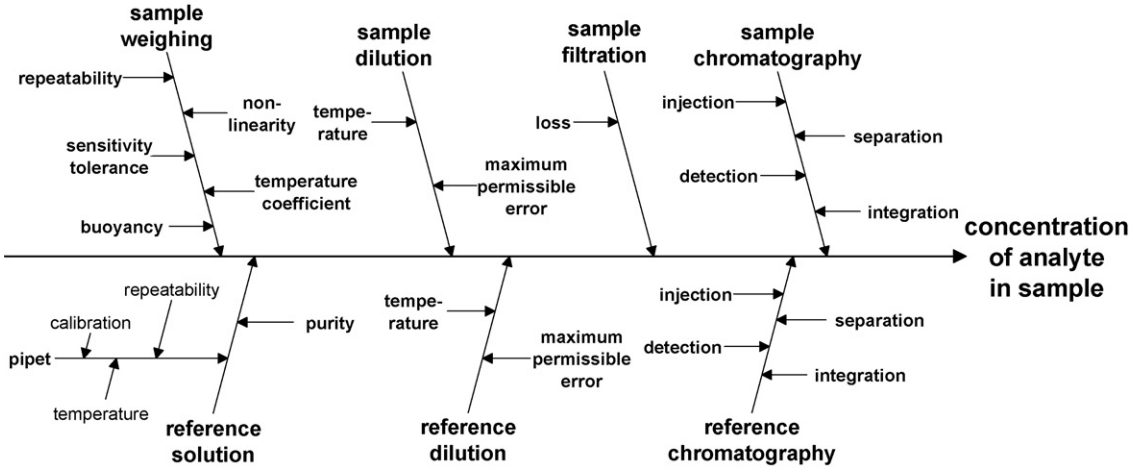

Fig. 4. The Ishikawa diagram with the uncertainty sources of the measurement uncertainty of the analysis shown in Fig. 3.

between $x \pm 1s$ and 95% between $x \pm 2s$ if the number of measurements n is large (whereby "large" is not defined, personally I propose to set $n \geq 10$ for a rather reliable standard deviation).

$$s(x) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (3)$$

A cognate number of merit is the standard deviation of the mean s($\bar{x}$), obtained by dividing $s(x)$ by $\sqrt{n}$:

$$s(\bar{x}) = \frac{s(x)}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (4)$$

$s(\bar{x})$ is a smaller number than $s(x)$, therefore one can be tempted to use it for the calculation of the measurement uncertainty of repeated measurements. However, $s(\bar{x})$ must only be used if a certain well-defined object or a certain well-defined and finished process is described. As an example, if a single coin is weighed n times we will ascribe the standard deviation of the mean $s(\bar{m})$ to the mean mass $\bar{m}$. On the other hand, if a recovery is determined n times in order to know its expected mean value for the next analysis or for many future analyses we must use the standard deviation of the recovery s(Rec) according to Eq. (3). This assignment follows from the fact that the next analysis will provide its own (but unknown to us) recovery, which can be close to the mean or in a region, which is as far apart from the mean as 2s. Similarly, any parameter characterized with a repeatability will contribute its standard deviation $s(x)$ to the uncertainty calculation, i.e. a $u(x)$ of Eq. (1) will be an $s(x)$ in these cases.

Fig. 5a shows the bell-shaped function of the normal distribution with the areas covered by ± 1s and ± 2s.

3.4. Other standard uncertainties

In many cases it is not possible or it would be too time-consuming to determine the variability of an uncertainty source by repeated experiments. Therefore, its standard uncertainty is not a standard deviation but is characterized by another type of distribution. Although distributions can have many different shapes, including asymmetric ones, the QUAM uses only two of them, namely the rectangular and the triangular distribution. In contrast to the normal distribution with "open ends" they cover a limited span of data and have clear boundaries.

A rectangular distribution (Fig. 5b) is chosen if no detailed knowledge of the function is available. A typical and important example is the temperature in a laboratory without air-conditioning over a certain time-span, e.g. a day or a year. The known or assumed temperature extremes, e.g. 18–25 °C, are the boundaries of the distribution. Its width is $2a$ (which means 7 °C in the example given above), its standard uncertainty $u(x)$ is $a/\sqrt{3}$ or $0.6a$ (giving $0.6 \times 3.5\,°C = 2.1\,°C$). The span of $\bar{x} \pm u(x)$ covers 58% of the data.

The triangular distribution (Fig. 5c) is chosen if there is a higher probability for the data to lie in the middle of the interval of $2a$ than at the boundaries. Such a distribution can be assumed for the volume of a measuring flask. The manufacturing process is optimized in such a way that the desired volume of, e.g. 250.0 mL is hit more often than one of 249.9 mL or 250.1 mL. The standard uncertainty of the triangular distribution $u(x)$ is $a/\sqrt{6}$ or $0.4a$, giving $u(V) = 0.4 \times 0.1\,mL = 0.04\,mL$. The span of $\bar{x} \pm u(x)$ covers 65% of the data.

3.5. Calculation rules

The standard uncertainties $u(x_i)$ of the individual uncertainty sources x_i are added to the combined uncertainty $u_c(M)$ of the measurand. The general rule is as follows: set up the equation of the measurand M. Determine the partial differentials of M with respect to all x_i, square them, multiply them with the respective $u^2(x_i)$, add all the summands, and extract the root:

$$u_c(M) = \sqrt{\sum_{i=1}^{n}\left(\frac{\partial M}{\partial x_i}\right)^2 u^2(x_i)} \quad (5)$$

If the equation of M is not simple, e.g. if it includes summands and multiplicands (such as $M = (a+b)/(c-d)$ or the like), the partial differentials can have a complicated structure and the result $u_c(M)$ is even more intricate. However, for many relevant

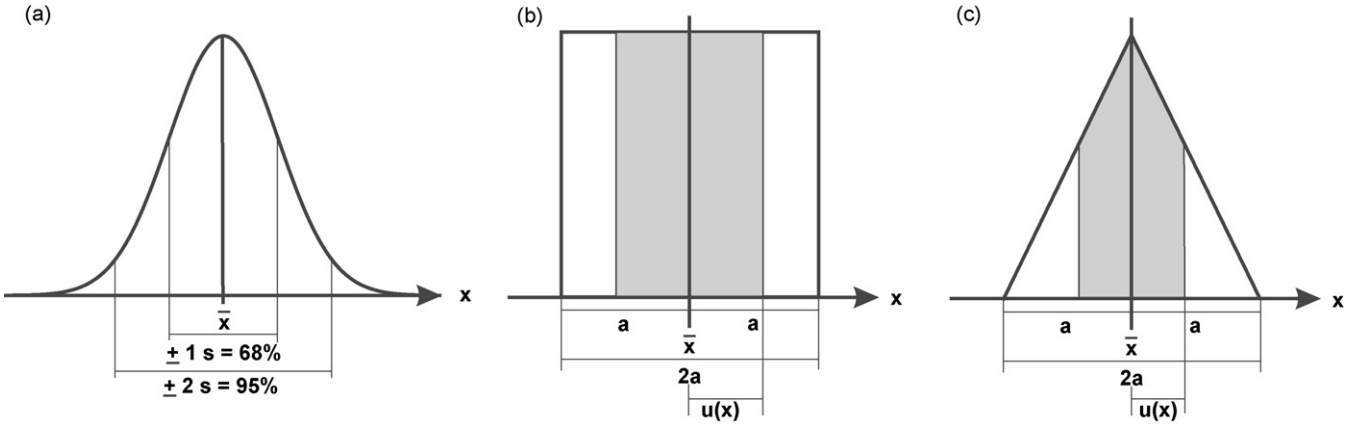

Fig. 5. Distribution functions. (a) Normal distribution with standard uncertainty $u(x)$ = standard deviation $s(x)$. The area of $\bar{x} \pm s(x)$ covers 68% of the data. (b) Rectangular distribution with standard uncertainty $u(x) = 0.6a$. The shaded area of $\bar{x} \pm u(x)$ covers 58% of the data. (c) Triangular distribution with standard uncertainty $u(x) = 0.4a$. The shaded area of $\bar{x} \pm u(x)$ covers 65% of the data.

equations the instruction of Eq. (5) leads to simpler expressions. If the equation for M consists only of multiplicands and divisors, such as $M = (a \times b)/(c \times d)$, Eq. (5) simplifies to:

$$\frac{u_c(M)}{M} = \sqrt{\sum_{i=1}^{n}\left(\frac{u(x_i)}{x_i}\right)^2} \tag{6a}$$

which is identical with:

$$u_c(M) = M\sqrt{\sum_{i=1}^{n}\left(\frac{u(x_i)}{x_i}\right)^2} \tag{6b}$$

Eq. (6a) gives the relative uncertainty (the percentage relative uncertainty is obtained by multiplication with 100) and Eq. (6b) gives the absolute uncertainty with the same unit as the measurand itself, e.g. mg/kg.

This type of equation is very common in analytical chemistry. Another simple class of equations, less frequently seen, are those with only summands (positive or negative ones). They occur, e.g. in the calculation of molecular weights or in the determination of the loss on drying. A measurand of the type $M = a + b - c$ leads to:

$$u_c(M) = \sqrt{\sum_{i=1}^{n} u^2(x_i)} \tag{7}$$

3.6. Monte Carlo method

The calculation rules given above are typical for the bottom-up approach (although some calculations may also be needed in the top-down approach). These rules can lead to a complicated and tedious calculation of the combined measurement uncertainty. An elegant way out is the determination of $u_c(M)$ by the Monte Carlo method [8]. It is based on the n-fold calculation of M with randomly generated values of the influence parameters. These numbers lie somewhere within their uncertainty boundaries around the specified value with a probability given by their relevant distribution function. n must be very high, 10,000 calculations are a minimum. Therefore, a Monte Carlo simulation will not run on an old or cheap personal computer but it can be performed on a PC with reasonable configuration within some minutes.

Monte Carlo has the advantage that the equation of the measurand can be of any complexity. It is not necessary to determine the partial differentials. In addition, multiple-point calibrations or correlated influence parameters are no problem and the resulting combined standard uncertainty is correct in all cases.

3.7. Expanded uncertainty

It was explained in paragraphs 3.3 and 3.4 that the standard uncertainties of the described distribution functions cover a range of 68%, 58% or 65% of the expected data. However, what is usually needed is a 95% level of confidence (only 5% of the experiments will yield a result which is outside of this range). Therefore, a standard uncertainty needs to be expanded with its appropriate coverage factor k. The resulting expanded uncertainty has the symbol U without index: $U(M) = u_c(M) \times k$.

The coverage factor of a normal distribution is 2 (strictly speaking, it is 1.96 to cover 95.0%). For a rectangular distribution it is 1.65 and for a triangular distribution it is 1.93.

In the case of a combined standard uncertainty $u_c(M)$ the expansion can be somewhat more intricate. It depends on the type of distribution, which dominates the combined uncertainty, i.e. which parameter is largest in an equation such as Eq. (1). In many cases this is the normal distribution; then the expansion is 2 if the number of experiments was high, namely 20 or higher. With fewer experiments the coverage factor is defined by Student's t-distribution [9]. With $n = 3$ it is necessary to multiply the ± 1s uncertainty with $k = 4.3$, with $n = 4$ the factor is still $k = 3.2$ and drops only slowly to $k = 2.3$ with $n = 10$. It is obvious that a small number of experiments is "punished" with a large coverage factor.

If a term with rectangular or triangular distribution function is the dominating one the coverage factor is independent of the number of experiments and is 1.65 or 1.93, respectively.

If several terms with different distribution functions are of similar magnitude the central limit theorem states that the overall distribution converges rapidly to a normal distribution [10]. For details see the GUM, paragraph H.2 [1].

4. Some basic uncertainty sources

4.1. Volumetric operations

The uncertainty of the volume of a measuring flask or a pipet has three uncertainty sources, namely the calibration, the repeatability, and the temperature influence. (Note: the QUAM does not list the possible ageing of a glass volumetric instrument as an additional uncertainty source. To our experience it can only occur if a strongly basic solution is stored in a burette over a long time period; even then, the effect is low and needs to be taken into consideration on a metrological level only.) The calibration uncertainty is subject to a triangular distribution because the manufacturer strives for the nominal volume. It can be found on the instrument itself or in the data sheet. The repeatability has a normal distribution (the standard deviation, not the standard deviation of the mean). If no data are available it must be determined experimentally. Note that published data are valid for water only; the repeatability is poorer with organic solvents, viscous solvents or water with detergent, and if untrained personnel is doing the work. The temperature influence has a rectangular (if no detailed temperature data are available) or a triangular distribution (if the laboratory is air-conditioned).

Temperature fluctuations result in a volume fluctuation according to:

$$u(V) = V\gamma u(T) \tag{8}$$

with γ being the cubic coefficient of volume expansion; $\gamma = 2.1 \times 10^{-4}\,\mathrm{K}^{-1}$ for water (in the vicinity of 20 °C) and $\gamma \approx 1 \times 10^{-3}\,\mathrm{K}^{-1}$ for organic solvents. The coefficient of expansion of glass instruments is smaller, therefore it is not of

importance. On the other hand, plastics used for flasks or pipets (polypropylene or polymethylpentene) have a coefficient similar to liquids, which leads to the consequence that the temperature term is negligible if such instruments are used.

In glass flasks and glass pipets the calibration and repeatability terms are combined to the so-called "maximum permissible error". This is the number engraved or printed on such instruments close to the nominal volume, e.g. 10 mL ± 0.025 mL. It is only valid for aqueous solutions and is probably larger for organic solvents. The maximum permissible error has a triangular distribution.

4.2. Weighing

Mass determinations are prone to a large number of uncertainty parameters: repeatability, non-linearity of the characteristic curve (or response curve), sensitivity tolerance (or slope tolerance of the characteristic curve), temperature coefficient, and buoyancy [11]. The relationships are not obvious and intuitive but rather complicated. In many cases the uncertainty of a weighing operation is so small that it is not necessary to consider it for the calculation of the combined standard uncertainty of an analytical method. However, the situation is worse if the weighing good is critical, i.e. if it is volatile or prone to become electrically charged (powders). In such cases the repeatability may be poor and must be determined experimentally.

Sometimes it is not the technical data of the balance, which dominate the total uncertainty but it is the buoyancy term [12]. Although it is possible to calculate this parameter it can be assumed that the densities of many weighing goods are not known better than with a relative uncertainty of 10–40%. This uncertainty has a direct influence on the combined uncertainty of the mass determination. Nevertheless the combined uncertainty is less (or much less) than 10^{-3} (1‰) even in cases of a high density uncertainty if the repeatability is good.

4.3. Purity of standards and reference materials

The quantitative chromatographic methods are not primary methods but they need the comparison with a well-known amount of reference compounds. The purity of a reference needs to be taken into consideration, i.e. in many cases it is necessary to correct the result by the decimal fraction of purity of the reference (depending on the degree of purity and on the tolerated measurement uncertainty). Therefore the purity P_R is noted in Eq. (2). In fact, the purity is an influence parameter which must be considered in all uncertainty determinations, even in the top-down approach with inter-laboratory reproducibility (see Section 6 below).

Certified reference materials have a well-defined degree of purity or of analyte content, including information about the uncertainty. (It may, however, be a problem to find out if the ± sign noted on the label means a standard deviation or the boundaries of a rectangular distribution.) Elaborate analytical examples with certified reference materials can be found in the QUAM [2].

Other chemicals give rather poor informations about their purity with regard to measurement uncertainty. A common specification type is of the kind ≥9*X.X*%, e.g. ≥99.5%. Does such a declaration mean that the compound has a purity between 99.5% and 99.6% or is it highly probable that the purity is close to 99.9%? One can take the interval between the lower specified purity and 100% as the boundaries of a rectangular distribution. In the case of a "≥99.5%" purity the interval has a width 2*a* of 0.5% (see Fig. 5b). This distribution gives a standard uncertainty of $0.6a = 0.6 \times 0.25\% = 0.15\%$ or $u(P) = 0.0015$.

A study with 40 commercially available compounds with a declared purity of either ≥97.0% or ≥99.0% showed that their content is closer to 100% than to the lower limit and that an asymmetric ramp function is a reasonable description of both the expected content and its uncertainty [13].

4.4. Atomic and molecular weights

In chromatographic analyses the weights of atoms or molecules do usually not appear in the equation of the measurand, therefore this paragraph is included here only for completeness of the discussion. In other analytical techniques, e.g. in titrimetry, these weights are a prerequisite for the calculation of the result and need to be considered for the uncertainty determination. For an example see Appendix A3 in the QUAM [2].

Atomic weights and their uncertainties are published regularly by the IUPAC. The most recent list, the "Atomic Weights of the Elements 2005" was published in 2006 [14]. The relative uncertainties range from approx. 1×10^{-9} for sodium to 7×10^{-4} for boron. It is an interesting fact that many elements have a rather high uncertainty of their atomic weight not due to problems with mass determination by mass spectrometry but due to a poorly defined and non-uniform isotopic composition. The IUPAC uncertainty data need to be taken as rectangular distributions.

Molecular weights are sums of atomic weights, therefore the calculation rule according to Eq. (7) is to be used for the determination of their uncertainty. The uncertainty of the atomic mass fraction in a molecule must be calculated with partial differentials according to Eq. (5), thus leading to a complicated expression [15].

4.5. Multiple-point calibration (linear regression)

The uncertainty of a calibration function can only be determined rather reliably if the fluctuations of the y values (the peak areas or peak heights) are considerably higher than the fluctuations of the x values (the concentrations of the calibration solutions). The appendix E.3 of the QUAM presents some approximative equations which allow to calculate this uncertainty [2]. However, such proposals are unsatisfactory. Three other strategies are better:

(a) Use the Monte Carlo method. The possible variations of the *x*/*y* data points must be known. For an *x* point, i.e. a concen-

tration, the uncertainty can be calculated as a combination of the weighing and diluting uncertainties. The variation of a y point, i.e. a peak size, is determined by the n-fold injection of the same solution.

(b) Use a software, see Section 7 below.

(c) Determine the uncertainty by experiments. Repeat n times the whole steps of preparing a stock solution, preparing the standard solutions, setting up the calibration function, injecting a sample (always the same solution), and calculating the analyte concentration in the sample. The resulting standard deviation is the standard uncertainty of the calibration.

4.6. Recovery

The recovery is a parameter which is studied during the validation, therefore its uncertainty is known as a standard deviation. For reasons outlined in paragraph 3.3 it is not allowed to use the standard deviation of the mean as the relevant standard uncertainty.

If the recovery is not determined but estimated it is necessary to define a rather wide span with its possible lower and upper limit, set by experience, and to treat this interval as a rectangular distribution.

5. A simple example

In Fig. 3 and in paragraph 3.2 a simple analytical test procedure with one-point calibration was presented. The equation of the measurand is as follows:

$$c_S = \frac{V_{\mathrm{Pip}R} \times c_R \times P_R \times V_{\mathrm{Flask}S} \times A_S}{V_{\mathrm{Flask}R} \times m_S \times A_R} \tag{2}$$

This is an equation with multiplicands and divisors only. For the calculation of the combined standard uncertainty $u_c(c_S)$ by the bottom-up approach we can therefore use the calculation rule noted in Eq. (6a); we get:

$$\frac{u_c(c_S)}{c_S} = \sqrt{\left(\frac{u(V_{\mathrm{Pip}R})}{V_{\mathrm{Pip}R}}\right)^2 + \left(\frac{u(c_R)}{c_R}\right)^2 + \left(\frac{u(P_R)}{P_R}\right)^2 + \left(\frac{u(V_{\mathrm{Flask}S})}{V_{\mathrm{Flask}S}}\right)^2 + \left(\frac{u(A_S)}{A_S}\right)^2 + \left(\frac{u(V_{\mathrm{Flask}R})}{V_{\mathrm{Flask}R}}\right)^2 + \left(\frac{u(m_S)}{m_S}\right)^2 + \left(\frac{u(A_R)}{A_R}\right)^2} \tag{9}$$

$$\frac{u_c(c_S)}{c_S} = \sqrt{\left(\frac{0.0044\,\mathrm{mL}}{1\,\mathrm{mL}}\right)^2 + \left(\frac{0.06\,\mathrm{mg/L}}{5.0\,\mathrm{mg/L}}\right)^2 + \left(\frac{0.047\,\mathrm{mL}}{100\,\mathrm{mL}}\right)^2 + 0.009^2 + \left(\frac{0.047\,\mathrm{mL}}{100\,\mathrm{mL}}\right)^2 + 0.0001^2 + 0.006^2} = 0.0168 \text{ or } 1.7\%. \tag{12}$$

We need knowledge of eight data and their standard uncertainties.

Piston-driven pipet used for the volume of the aqueous reference solution: $V_{\mathrm{Pip}R} = 1$ mL. Its uncertainty is an additive combination of calibration uncertainty (8 μL), repeatability (3 μL, both data according to the relevant norm [16]), and the temperature effect. The standard uncertainties are: $u(V_{\mathrm{Cal}}) = 8\,\mu\mathrm{L} \times 0.4 = 3.2\,\mu\mathrm{L}$ (triangular distribution); $u(V_{\mathrm{Rep}}) = 3\,\mu\mathrm{L}$ (the published value is a standard deviation); $u(V_T) = 1\,\mathrm{mL} \times 2 \times 10^{-4}\,\mathrm{K}^{-1} \times 1.2\,\mathrm{K} = 2.4 \times 10^{-4}$ mL (Eq. 8) with a rectangular temperature span in the laboratory of ±2 °C, giving $u(T) = 2\,\mathrm{K} \times 0.6 = 1.2$ K. The combined volume standard uncertainty is:

$$u(V_{\mathrm{Pip}R}) = \sqrt{3.2^2 + 3^2 + 0.24^2}\ \ \mu\mathrm{L} = 4.4\,\mu\mathrm{L} \tag{10}$$

Concentration of the reference solution: The manufacturer guarantees a concentration of the analyte of 5.0 mg L^{-1} with an uncertainty of ±0.1 mg L^{-1}. We treat this information as a rectangular distribution and get a standard uncertainty of $u(c_R) = 0.1\,\mathrm{mg\,L^{-1}} \times 0.6 = 0.06\,\mathrm{mg\,L^{-1}}$.

Purity of the reference solution: in our case with a certified reference solution this term is already covered with the above-mentioned guarantee of the manufacturer. The purity term will therefore not appear in our calculation of $u(c_S)$.

Measuring flasks used for the diluted sample and reference solutions: $V_{\mathrm{Flask}} = 100$ mL. Our considerations are similar as in the case of the pipet but we need only the maximum permissible error (MPE) of ±0.100 mL [17] and the temperature effect. The standard uncertainties are: $u(V_{\mathrm{MPE}}) = 0.100\,\mathrm{mL} \times 0.4 = 0.04$ mL (triangular distribution); $u(V_T) = 100\,\mathrm{mL} \times 2 \times 10^{-4}\,\mathrm{K}^{-1} \times 1.2\,\mathrm{K} = 0.024$ mL. The combined volume standard uncertainty is:

$$u(V_{\mathrm{Flask}}) = \sqrt{0.04^2 + 0.024^2}\ \ \mathrm{mL} = 0.047\,\mathrm{mL} \tag{11}$$

Peak areas: we determined a relative repeatability of the reference peak area of 0.6% and of the sample peak (in a poorer chromatogram) of 0.9%. Therefore, $u(A_R)/A_R = 0.006$ and $u(A_S)/A_S = 0.009$.

Mass of the sample: the sample is a coarse-grained solid with a well-known density of 1.2 kg dm^{-3}. The relative standard uncertainty of the weighing operation can be assumed to be not higher than 1×10^{-4} (100 ppm).

These data allow the calculation of the combined standard uncertainty of the analyte concentration in the sample with Eq. (9):

The largest influence parameter is the concentration of the reference solution with a relative standard uncertainty of 1.2%. In order to decrease the combined measurement uncertainty of the analysis it would be best to look for a better reference material. In addition, it is necessary to take such measures to keep the long-term repeatability of the chromatography below 1%, a task which can be demanding [18].

Since the concentration of the reference dominates the combined measurement uncertainty its distribution function dictates the expansion to the 95% level of confidence. It is

a rectangular distribution with a coverage factor $k_{rect} = 1.65$. Therefore the expanded relative measurement uncertainty $U(c_S)$ is $1.7\% \times 1.65 = 2.8\%$.

This example was presented as a "simple" one. For the beginner the necessary calculations and considerations look rather complicated. However, it is a fact that one gets rapidly familiar with them once some examples have been worked through. Self-made Excel sheets for the calculation of the volume and weighing uncertainties and of the combined uncertainty are a great help. In addition, with some experience one will get a feeling of which influence parameters are of importance; in many cases the largest contribution to the combined uncertainty comes from the repeatability of the recovery. Then many other contributions are negligible.

6. The top-down approach

The top-down method for the calculation of the measurement uncertainty does not look at every detail as described in Section 5. It starts from a repeatability or a reproducibility. The analyst needs a good knowledge of how the standard deviation of an analytical test procedure was determined in order to obtain a reliable uncertainty.

The simplest repeatability is obtained by the n-fold injection of the same solution. However, this approach has nothing to do with the combined measurement uncertainty but is only an element of the bottom-up approach as described in the example above. The minimum requirement for a reliable uncertainty of a test procedure is the n-fold repetition of the whole procedure, including all weighing operations and the preparation of all reference and sample solutions. The obtained standard deviation must be amended with those influence parameters which were not altered but kept constant.

If, e.g. only one single piston-operated pipet is available in the laboratory the standard deviation of the test procedure, i.e. its repeatability, should be expanded by the calibration uncertainty of this type of pipet:

$$\frac{u_c(M)}{M} = \sqrt{\left(\frac{Rep(M)}{M}\right)^2 + \left(\frac{u(V_{Cal})}{V}\right)^2} \quad (13)$$

This example is a simple and probably not important one. But similar expansions should be done if the repeatability was determined by a single person or with a limited type of samples. The most important parameter usually not included in a repeatability or a reproducibility is the purity P of the reference material, and it is often not a small one. Therefore the standard deviation of the test procedure needs to be corrected as follows:

$$\frac{u_c(M)}{M} = \sqrt{\left(\frac{Rep(M)}{M}\right)^2 + \left(\frac{u(P)}{P}\right)^2} \quad (14)$$

7. Hints and tools

The bottom-up approach is an excellent tool to identify the "weak points" of a test procedure with regard to its combined measurement uncertainty. Working through the whole process makes also clear if the standard operating procedure is written clearly, completely and unambiguously. Sometimes a method can be simplified markedly, resulting in a lower uncertainty [19].

Sample preparation steps should be kept simple and limited in number [20]. Small volumes (pipets or flasks) have a higher relative uncertainty than large ones. The parallel work-up of sample and reference keeps low the inevitable temperature drift effects which need to be considered. Using an internal standard is one of the best techniques to keep low (or even to cancel) a number of disturbing phenomena such as analyte recovery effects or injection problems. Well-qualified personnel will obtain lower uncertainties than non-skilled people.

The tedious calculation of measurement uncertainties can be avoided by using dedicated software [21]. It allows the determination of the measurement uncertainty of analytical test procedures, including the sample preparation steps, by either the bottom-up or the Monte Carlo approach.

8. Conclusions

The determination of the measurement uncertainty of all validated analytical test procedures should not be looked at as an additional burden but as a worthwhile completion bringing added value. Measurement uncertainty allows to evaluate a result and to compare it with other results, especially those coming from other laboratories. The positive meaning of "measurement uncertainty" must be communicated to the superiors of the laboratory and to the customers.

A realistic measurement uncertainty datum is not too optimistic, i.e. too low, by using some repeatabilities which do not cover the whole analytical process or by using too narrow boundaries of a distribution. Such a number is not honest and, in addition, it can complicate in an unnecessary way the communication with other laboratories, customers or authorities. On the other hand, it makes no sense to declare too high an uncertainty by just guessing it.

The serious determination of measurement uncertainties is less complicated and demanding than it may seem at first glance. The time and money needed to establish it in a laboratory is paid back by reliable analytical data.

References

[1] Guide to the Expression of Uncertainty in Measurement, ISO, Geneva, 1995, ISBN 92-67-10188-9.
[2] S.L.R. Ellison, M. Rösslein, A. Williams, (Eds.), EURACHEM/CITAC Guide, Quantifying Uncertainty in Analytical Measurement, second ed., 2000. http://www.measurementuncertainty.org/mu/guide.
[3] EN ISO/IEC 17025, General requirements for the competence of testing and calibration laboratories, CEN Management Centre, Brussels, 2005.
[4] ICH Q2(R1), Validation of Analytical Procedures: Text and Methodology, ICH, Geneva, 2005.
[5] K. Ishikawa, Introduction to Quality Control, Kluwer, Dordrecht, 1991 (Chapter 4.7.4).
[6] S.L.R. Ellison, V.J. Barwick, J. Accred. Qual. Assur. 3 (1998) 101.
[7] V.R. Meyer, J. Chromatogr. Sci. 41 (2003) 439.

[8] Joint Committee for Guides in Metrology, Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method, Final draft, BIPM, Sèvres, 2006.
[9] Student (William S. Gosset), Biometrika 6 (1) (1908) 1.
[10] W. Feller, Bull. Am. Math. Soc. 51 (1945) 800.
[11] A. Reichmuth, S. Wunderli, M. Weber, V.R. Meyer, Microchim. Acta 148 (2004) 133.
[12] M. Pozivil, W. Winiger, S. Wunderli, V.R. Meyer, Microchim. Acta 154 (2006) 55.
[13] G. van Look, V.R. Meyer, Analyst 127 (2002) 825.
[14] M.E. Wieser, Pure Appl. Chem. 78 (2006) 2051.
[15] V.R. Meyer, Anal. Bioanal. Chem. 377 (2003) 775.
[16] EN ISO 8655-2, Piston-operated volumetric apparatus, Part 2: Piston pipettes, CEN, Brussels, 2002.
[17] ISO 1042, Laboratory glassware—one-mark volumetric flasks, fourth ed., ISO, Geneva, 1998.
[18] S. Küppers, B. Renger, V.R. Meyer, LC-GC Eur. 13 (2000) 114.
[19] R. Brix, S.H. Hansen, V. Barwick, J. Tjørnelund, Analyst 127 (2002) 140.
[20] V.R. Meyer, LC-GC North Am. 20 (2002) 106; LC-GC Eur. 15 (2002) 398.
[21] UncertaintyManager® by EMPA, St. Gallen, Version 2.0, 2006. See http://www.uncertaintymanager.com.