

Expert System for the Evaluation of Measurement Uncertainty

Making Use of the Software Tool uncertaintyMANAGER®

Matthias Rösslein* and Bruno Wampfler^a

Abstract: An expert system for the evaluation of the measurement uncertainty is presented. It follows the four-step process established by the Eurachem/CITAC guide QUAM. The expert system provides a considerably better estimate of the overall measurement uncertainty than certain summary approaches used nowadays in most private and public laboratories and in industry. This is demonstrated by an example from the production control in a pharmaceutical lab. In addition, the expert system allows performing the entire process to evaluate the measurement uncertainty much faster than the summary approaches used in industry.

Keywords: Expert system · GUM · GUM supplement · Measurement uncertainty · Software · QUAM

1. Introduction

In the recent past numerous local accreditation bodies started to demand from analytical laboratories a more stringent implementation of the most recent ISO/IEC/EN 17025 standard.^[1] This ISO document demands that ‘Testing laboratories shall have and shall apply procedures for estimating uncertainty of measurement...’ (Clause: 5.4.6.2). Therefore, different analytical laboratories in the private and public sector have started their own efforts to implement procedures to estimate the uncertainty of their measuring values in some way or other.^[2] Very often these efforts have to be accomplished at minimal costs in a commercial environment.

Note 3 of clause 5.4.6.3 ISO 17025(2005) refers the reader to ISO 5725 and to the Guide to the Expression of Uncertainty in Measurement (GUM) for detailed information.^[3,4] Furthermore, the International Laboratory Accreditation Cooperation (ILAC) recommends analytical laboratories to utilise also the Eurachem/CITAC guide ‘Quantifying Uncertainty

in Analytical Measurement 2nd (QUAM) in order to evaluate measurement uncertainty.^[5,6] In particular, the latter guideline puts more emphasis on an approach that is adjusted to the needs of the bench chemist by using overall performance parameters determined during the method validation. However, even then, the procedure is rather tedious and time consuming and is still very demanding for the bench chemist. Therefore, there is a considerable tendency and financial pressure to simplify the general procedure to evaluate the measurement uncertainty. Simple approaches do not lead to a thorough calculation based on the specification and scope of the analytical procedure, but rather to a simple estimation of its value. Such an approach might be adequate for the environmental sector where considerable inhomogeneity of the samples already causes a large variation of the results. However, misleading interpretations and wrong decisions are inevitable within many other sectors of analytical chemistry like production monitoring, doping control or forensic evidence, if evaluation of measurement uncertainty is based on such a simple methodology.^[7]

These considerations lead to the conclusion that only an expert system, that supports the bench chemists comprehensively in estimating the measurement uncertainty, can facilitate the successful introduction of the concepts detailed in the various guidelines. Therefore, Empa started a project to create such an expert system together with partners from industry, from public and private laboratories as well as from instrument manufacturers. In the meantime, this project has been successfully completed with the release of the software tool uncertaintyMANAGER® and with the subsequent foundation of the spin off company ValiTrace GmbH.^[8] In the

present publication some of the capabilities of the expert system are described by means of an example from a typical analytical procedure in production control.

2. Some Basics about the Evaluation of Measurement Uncertainty

The introduction lists a number of guidelines that outline the procedures to evaluate the measurement uncertainty. The primary and most generally valid document is GUM^[4] that subdivides the evaluation of the measurement uncertainty into two major phases as shown in Fig.1: the formulation phase and the calculation phase.

The Eurachem/CITAC guideline QUAM on the other hand suggests a four-step process: ‘Specify, Identify, Quantify and Calculate’ (Fig. 2) with the first three steps reflecting the formulation phase of GUM.^[6] The formulation phase aims at building an appropriate measurement model. The metrologist performs this task, whereas computers execute the calculation phase. The expert system uncertaintyMANAGER greatly facilitates the tasks of the formulation phase, so that also bench chemists can perform the whole evaluation of the measurement uncertainty without being experts in the field of setting up measurement models that represent all the different influence quantities affecting the measurement result.

The information for the first step ‘Specify’ is taken from the description of the analytical procedure (*e.g.* standard operation procedure (SOP)). For our purpose, important parts of this description are the equation of the measurand and a flow chart depicting the operating se-

*Correspondence: Dr. M. Rösslein
Laboratory for Materials-Biology Interactions
Empa, Swiss Federal Laboratories for Materials
Testing and Research
Lerchenfeldstrasse 5
CH-9014 St. Gallen
Tel.: +41 71 274 7784
Fax: +41 71 274 7694
E-mail: matthias.roesslein@empa.ch
^aLaboratory for Biomaterials, emp

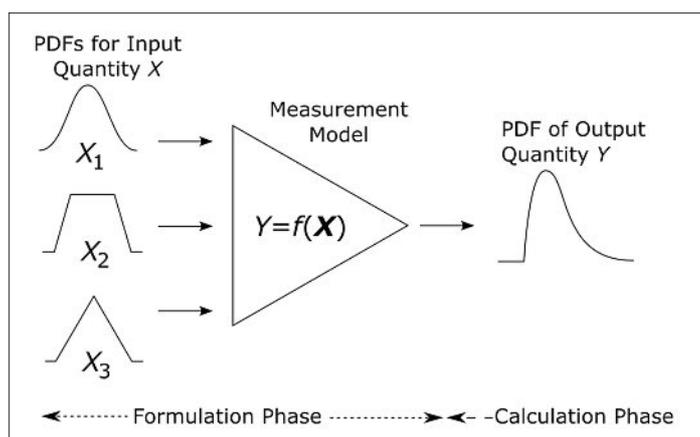


Fig. 1. Two major phases to evaluate the measurement uncertainty according to GUM (PDF = Probability Density Function).

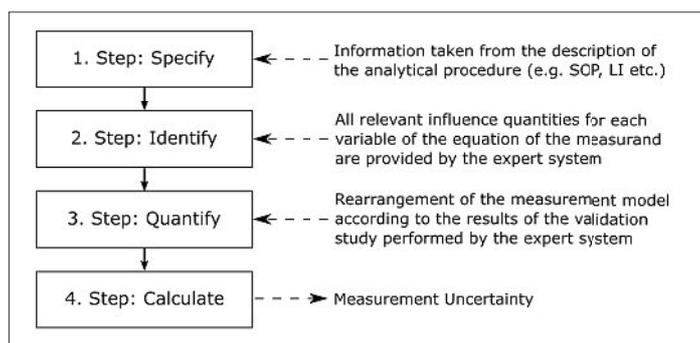


Fig. 2. Four steps to evaluate the measurement uncertainty according to QUAM.

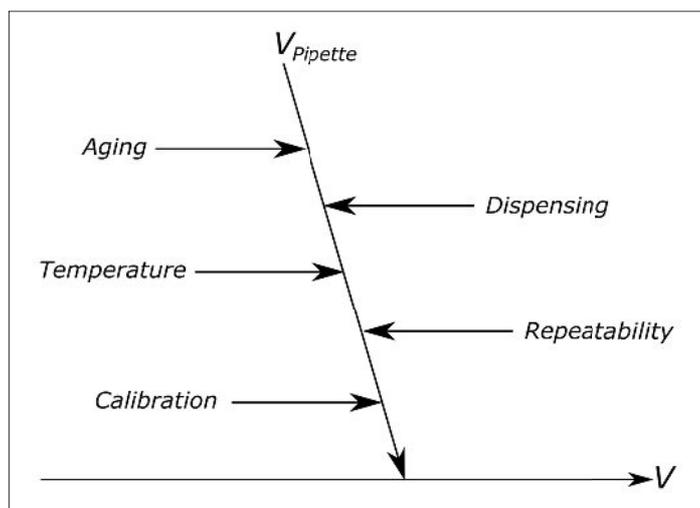


Fig. 3. Cause and effect diagram of a volumetric measurement using a pipette.

quence of the analytical procedure. However, we know from several workshops and seminars on the subject of 'calculating the measurement uncertainty' that these parts are quite often missing in the relevant documents. The equation of the measurand is the building block to start with the next step 'Identify'. The goal of the second step is a structured list of all the influence quantities affecting the procedure of measurement. Initially, Ellison and Barwick have proposed the 'cause and effect diagram' as such a structured list, and they worked out a procedure to construct the diagram.^[9] At this stage the expert system comes into play with a slightly modified procedure,

which was adopted for the needs of a software application:

- i) The variables of the equation of the measurand form the main branches of the diagram.
- ii) Consideration of each step of the analytical procedure and addition of any further factors by introducing a corresponding branch to the diagram and a corresponding term to the equation of the measurand (e.g. soxhlet extraction, derivatisation, etc.).
- iii) For each branch addition of contributory factors for all relevant effects associated with each major step and again addition of corresponding terms to the measurement equation. This means that

the effect of all remaining insignificant factors on the result can be neglected.

We will illustrate the assembly of a cause and effect diagram with the volumetric measurement device 'glass pipette'. The equation of the measurand is $V = V_{\text{pipette}}$ with the left-hand side of the equation as measurand and the right-hand side as a volumetric measurement using a pipette.^[6] The two sides also build the stem and main branch of the cause and effect diagram. For this introductory example there are no additional factors like steps of sample preparation that must be added to the diagram and to the equation of the measurand. In the next step, all influence quantities that might have an effect on the result of the measurement are evaluated. The influence quantities are shown in Fig. 3 and are described as follows:

- The calibration of the pipette performed by the manufacturer.
- The repeatability of the volumetric measurement with the pipette.
- The temperature at which the measurement will be made; the pipette has been calibrated by the manufacture at 20 °C with water. However, the bench chemist is usually using the pipette within a temperature range of 18 °C and 27 °C. The difference between the actual and the calibration temperature must be considered, because the expansion coefficient of the solvent is considerably larger than that of the material of the pipette (glass).
- Dispensing: Pipettes are calibrated for dispensing their nominal volume. For this purpose, it is essential that the whole content of the pipette is dispensed in a well-controlled way over a given amount of time. This influence allows the effect of an incomplete transfer of the solution into the receiver to be quantified due to degrading surface properties of the inner glass walls of the pipette.
- Aging: The aging of the glassware is caused by the amorphous structure of the material. The volumetric glassware might change its form slightly at a temperature notably below its melting point. According to the manufacturer's information the cleaning and drying cycle with temperatures up to 80 °C and higher result in small changes in the form of any volumetric measuring device made up of glass. These changes might add up because the pipettes are normally washed and dried in the same way. Therefore, we also have to allow for the aging of our volumetric measurement device.

Using the expert system, all relevant influence quantities rather than just the significant ones are included, which is in contrast to the original publication by El-

lison and Barwick.^[9] When assembling the cause and effect diagram, the relevant information for the quantification is not yet available, as only a preliminary calculation can reveal which influence quantities have a significant effect. This means that the full measurement model must be available to perform the overall measurement uncertainty calculation. Therefore, the expert system does not triage the influence quantities, when the measurement models for a given analytical procedure are worked out.

Before the overall model of the analytical procedure can be built up, additional considerations about some basic principles of measurements are needed. Nearly all measurements are relative and they follow the same principle as shown in Fig. 4. There is always a method to compare the measurement value of a sample, e.g. signal $A(s)$, with that of a reference $A(r)$. The calibration of a pipette with a reference measurement of the volume takes place within the manufacturing process. It corresponds to the calibration in the cause and effect diagram. All the remaining influence quantities are needed to account for effects that might influence the pipette during its later use to measure a solution ('sample') in any given laboratory. The situation is different if the reference and sample measurements are performed with the same instrument within a relatively short period of time. Many systematic effects are cancelled out in such a situation, because they affect the measured values for the reference and the sample in the same way.

These principles of different types of relative measurements and their effect on the overall model of an analytical procedure are best illustrated in conjunction with step 3 and 4 of the evaluation of the measurement uncertainty with the following example.

3. Example – Control of By-Product Content in Chemical Production

The amount of two different by-products after a synthesis step in chemical production is to be determined. One by-product is at the mass content level of 0.5% and the other one at the level of 0.2%. The content of both by-products is measured in the same solution that is used to determine the content of the key-product. For this purpose 100 mg of the product is dissolved in 100 ml of the mobile phase. The reference stock solution is produced from 50 mg of the reference standard of the by-product, which is dissolved in 100 ml of the mobile phase. The actual measuring solution of the reference is produced by diluting 1 ml of the stock solution with the mobile phase to a volume of 100 ml. For the other by-product the concentration

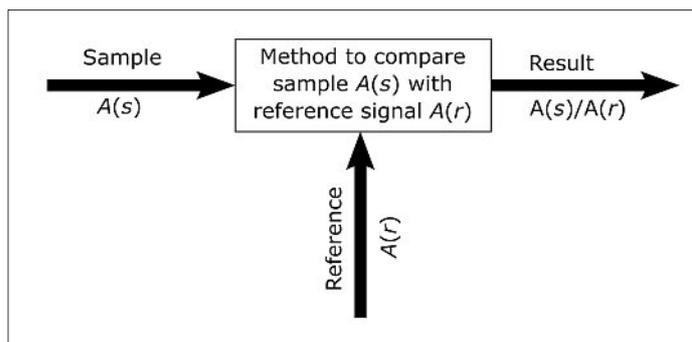


Fig. 4. Principles for HPLC measurements: 'Method to compare the sample signal $A(s)$ with a reference signal $A(r)$ '.

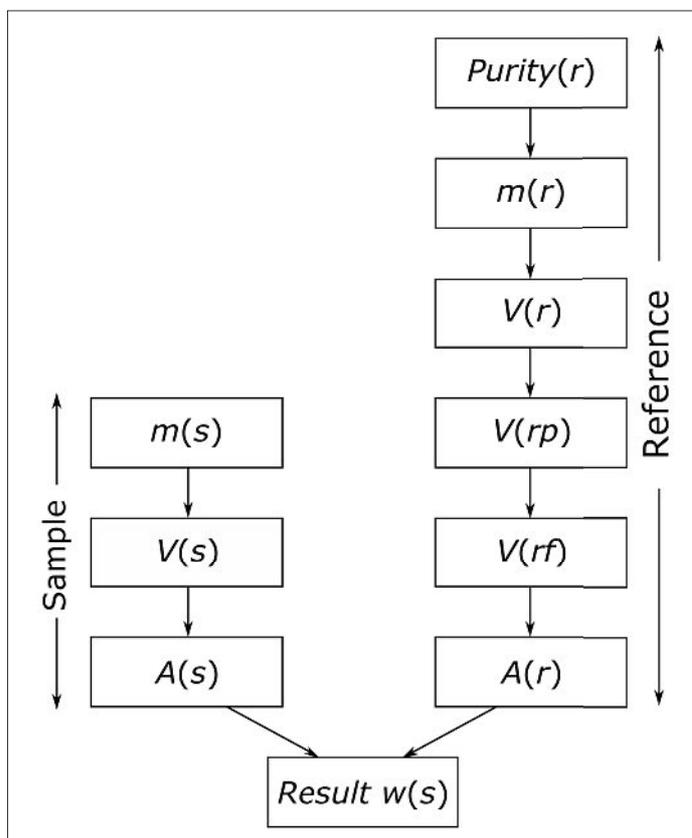


Fig. 5. Flow chart of the operating sequence to determine the content of by-products after a synthesis step in chemical production.

level of the measuring solution of the reference is adjusted by decreasing the amount of weighted reference substance to 20 mg. All readings of the weight measurements are rounded to 0.1 mg.

The Eqn. (1) of the measurand w is given for both analytes as

$$w_s = \frac{A_s}{A_r} \cdot \frac{m_r \cdot Purity_r}{V_r} \cdot \frac{V_{rp}}{V_{rf}} \cdot \frac{V_s}{m_s} \quad (1)$$

With w_s = content of by-product after synthesis step ($\text{g} \cdot \text{g}^{-1}$), A_s = absorption signal of the sample measured using HPLC, A_r = absorption signal of the reference measured using HPLC, m_r = mass of the reference substance (g), $Purity_r$ = purity of the reference substance (by-product) ($\text{g} \cdot \text{g}^{-1}$), V_r = volume of the reference stock solution (ml), V_{rp} = aliquot taken to dilute the reference stock solution (ml), V_{rf} = volume of flask containing the measuring solution of

reference (ml), V_s = initial volume of the measuring solution of the sample (ml), m_s = mass of the sample (g).

Fig. 5 depicts the flow chart of the operating sequence for the analytical procedure. The left branch summarises the operations with the samples, whereas the right branch describes the operations with the reference substance.

On the basis of the equation of the measurand and the operating sequence of the analytical procedure the expert system automatically constructs the proper cause and effect diagram (Fig. 6) following the rules outlined before.

The cause and effect diagram in Fig. 6 shows eight main branches representing the nine variables of the equation of the measurand. The HPLC signals $A(s)$ and $A(r)$ are described as a ratio and thus represented as only one branch, because most of the systematic effects for these types of measurement compensate each other

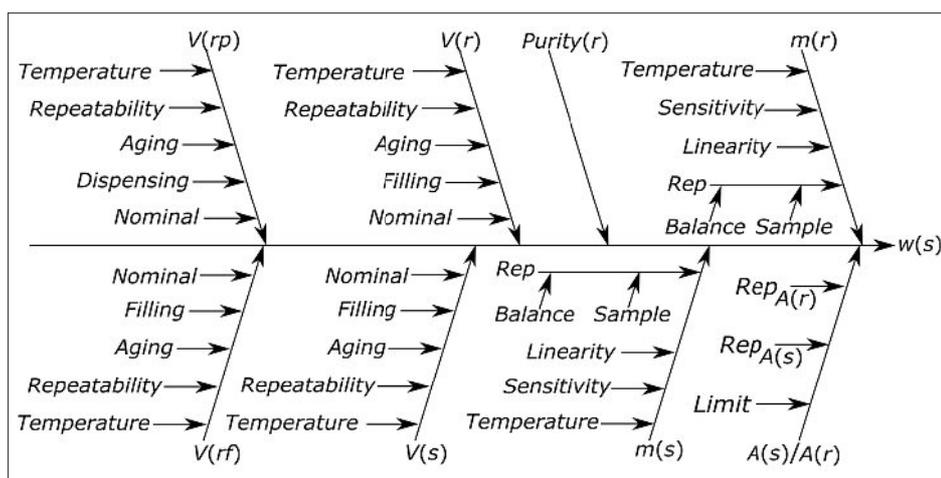


Fig. 6. Cause and effect diagram of the analytical procedure to determine the content of a by-product after a synthesis step.

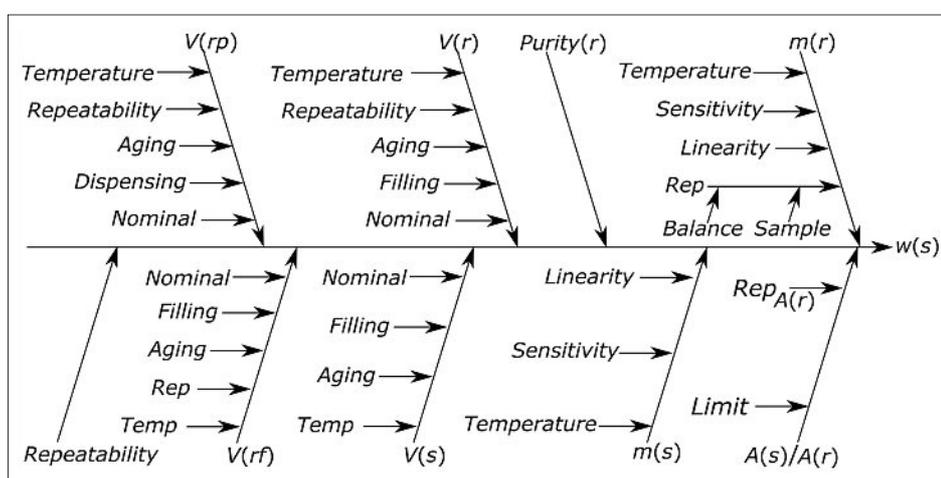


Fig. 7. Combination of the different influence quantities 'repeatability' of the samples to a new main branch.

as detailed in the previous section. The diagram itself was constructed by the expert system using preset models stored in a comprehensive database. These models for different measurement equipments and reference substances have been developed and validated by an international group of experts.

According to QUAM's rules summarised in Fig. 2 all influence quantities that have been identified and listed in the cause and effect diagram are now to be quantified (third step), using as much information and results from the validation study as possible. Hence, a closer look at the repeatability studies of the validation is needed.

In a typical measurement sequence a reference sample is measured first, followed by double measurements of the given samples. In order to evaluate the measurement uncertainty, laboratories often calculate the repeatability from the results of double measurements performed with sample solutions.^[2] This practice has some

direct consequences on the measurement model shown in Fig. 6. Only the repeatability influence quantities of the different variables are included in the determined repeatability, which could vary during the measurements. Their combination results in a new branch 'repeatability', which represents the overall variation of the analytical procedure. Fig. 7 depicts the modified cause and effect diagram for such a measurement sequence. The new main branch 'repeatability' only contains the individual repeatability influence quantities related to the sample. Again the expert system performs the rearrangement of the influence quantities within the cause and effect diagram based on detailed models elaborated and established by the same panel of international experts as mentioned above. In this way the expert system releases the bench chemist from a very complicated and tedious task, which is quite often heavily disputed even within expert bodies.

Another tedious and time-consuming task within the process to evaluate the measurement uncertainty is the actual quantification of all the different influence quantities in the cause and effect diagram shown in Fig. 7. This means for the bench chemist a search through many handbooks and standards to extract the needed values. This task is much better accomplished by using a large database, which contains all these values previously evaluated and where new measurement equipment is added on a regular basis. The expert system uncertaintyMANAGER is built on such a database and therefore directly relieves the bench chemist from one of the larger sticking points in the whole process to calculate the measurement uncertainty.

The final step in the evaluation process is the actual calculation of the measurement uncertainty. There are two established ways to perform this calculation. The first one is based on the uncertainty propagation and it has been detailed in the GUM.^[4] However, there have been considerable criticisms at various points (e.g. confidence interval) since its publication.^[10,11] This prompted the working group 1 of the joint committee for guides in metrology (JCGM) to draft a first supplement to GUM, which was finally published in 2008.^[12] This new guideline states that the Monte Carlo method is the method of choice to perform any measurement uncertainty calculation.

Fig. 8 and Fig. 9 show the combined standard uncertainty for the by-product of 0.5% and 0.2%, respectively and they illustrate the different contributions of the main influence quantities corresponding to the main branches in the cause and effect diagram. These two Figs illustrate that summary approaches, which are based more or less on the variability of the analytical procedure ('Repeatability'), are not suitable for estimating the measurement uncertainty; they fail by and large. A number of private, public and industrial laboratories have started using these summary approaches to reduce their costs for the entire evaluation procedure. Such approaches are even tolerated by some local accreditation bodies.

The contributions of the main branches of the cause and effect diagram show that at least five of them have a considerable effect on the overall measurement uncertainty. The most striking finding is the large increase of the effect from the mass of the reference substance going from the by-product with 0.5% to the one with 0.2%, which is caused by rounding the reading of the measured value.

4. Summary and Conclusion

The GUM framework and the derived Eurachem/CITAC guideline 'QUAM' have

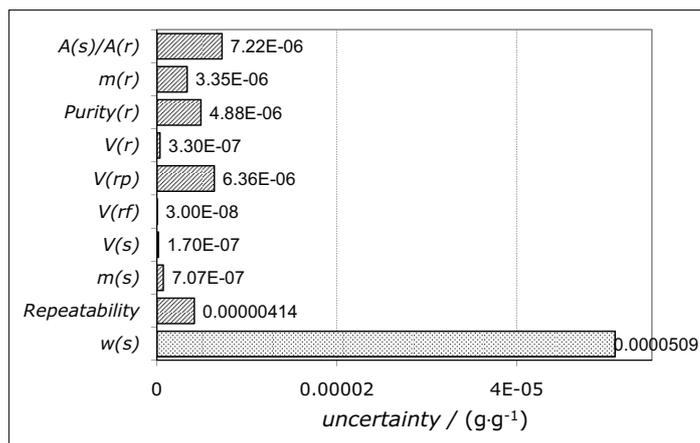


Fig. 8. Contribution of the main branches in the cause and effect diagram to the combined standard uncertainty, $c(s)$ for the by-product of 0.5%.

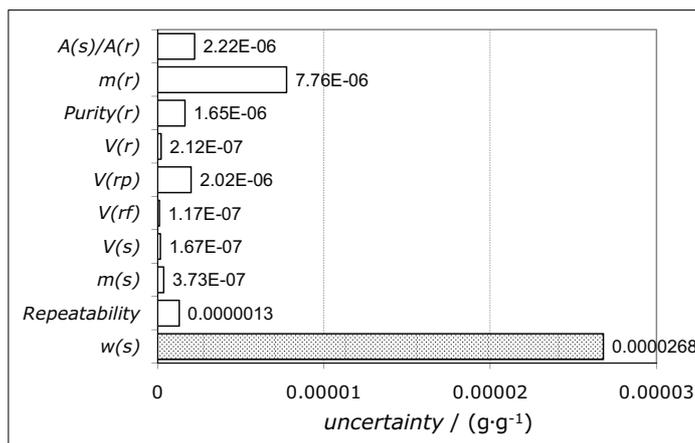


Fig. 9. Contribution of the main branches in the cause and effect diagram to the combined standard uncertainty, $c(s)$ for the by-product of 0.2%.

set up a number of rules for the evaluation of the measurement uncertainty. This procedure is still too tedious, time-consuming and expensive for many bench chemists in the public and private sector. This is why they started to establish individual approaches based mainly on the variation of analytical procedure. According to the example demonstrated in the present publication, these approaches might have considerable shortcomings compared with the calculated overall measurement uncertainties based on the GUM framework. However, the expert system uncertainty-MANAGER guides the users step by step through the entire evaluation process proposed by the Eurachem/CITAC guideline 'QUAM'. Predefined models for the different types of measurements and sample preparation steps, together with a large database and the two methods to calculate the measurement uncertainty make this expert system a very powerful tool in the hand of every bench chemist, saving tremendous time and money and avoiding the risk of wrong decisions, which are based on con-

siderable underestimation of the uncertainty of results from an analytical procedure.

Acknowledgement

The authors thank KTI/CTI (The innovation promotion agency) for co-financing the project.

Received: June 25, 2009

- [1] ISO/IEC/EN 17025: 'General requirement for the competence of testing and calibration laboratories', 2005, Geneva, Switzerland.
- [2] S. Populaire, E. Campos Gimenez, *Accred. Qual. Assur.* **2006**, *10*, 485.
- [3] ISO 5725: 'Accuracy (trueness and precision) of measurement methods and results', Geneva, Switzerland.
- [4] 'Evaluation of measurement data – Guide to the expression of uncertainty in measurement', BIPM – JCGM 100, 2008 (GUM 1995 with minor corrections), <http://www.bipm.org/en/publications/guides/gum.html>.
- [5] ILAC-G17: 'Introducing the concept of uncertainty of measurement in testing in association with the application of the standard ISO/IEC 17025', 2002.
- [6] 'Eurachem/CITAC Guide: Quantifying uncertainty in analytical measurement', Eds S. L. R. Ellison, M. Rösslein, A. Williams, 2nd ed., 2000.
- [7] M. Weber, M. Rösslein, 'Von Toleranzen und Ignoranzen im Dopingnachweis', *Neue Zürcher Zeitung (Forschung und Technik)* **2002**, 27. Feb, Nr. 48, p. 71.
- [8] Software: uncertaintyMANAGER® version 3.0, 2008, ValiTrace GmbH (Empa spin-off company), Speicher (CH); <http://www.uncertaintymanager.com>.
- [9] S. L. R. Ellison, V. J. Barwick, *Accred. Qual. Assur.* **1998**, *3*, 101.
- [10] M. Müller, M. Wolf, M. Rösslein, W. Gander, 'Limits of the uncertainty propagation: examples and solutions using the Monte Carlo method', *Proc. 13th Int. Metrology Congress (Lille)*, 2007.
- [11] M.G. Cox, B. R. L. Siebert, 'The use of a Monte Carlo method for evaluating uncertainty and expanded uncertainty', *Metrologia* **2006**, *43*, 178.
- [12] Evaluation of measurement data – Supplement 1 to the 'Guide to the expression of uncertainty in measurement' – Propagation of distributions using a Monte Carlo method, BIPM – JCGM 101: 2008, <http://www.bipm.org/en/publications/guides/gum.html>.