Estimation of measurement uncertainty in chemical analysis (analytical chemistry) course

This is an introductory course on measurement uncertainty estimation, specifically related to chemical analysis.

Ivo Leito, Lauri Jalukse, Irja Helm

Table of contents

Course introduction

This course will be offered again as a MOOC in spring 2016. The course will take place during March 28 – May 8, 2016. Registration is open.

If you want to know how the course went in spring 2014 and 2015 please see the presentation at the Euroanalysis 2015 conference.

Course introduction

http://www.uttv.ee/naita?id=17710

https://www.youtube.com/watch?v=r34Y-gzf62Y

The course was offered as an online course (MOOC) in Moodle environment during March 03 - April 13, 2014 and March 02 - April 12, 2015. Altogether more than 700 participants from more than 70 countries have participated. Altogether more than 40% of them successfully completed the course and were awarded a certificate from University of Tartu.

The course will be offered again in March 28 - May 08, 2016.

This course has been described in the paper: I. Leito, I. Helm, L. Jalukse. Using MOOCs for teaching analytical chemistry: experience at University of Tartu. *Anal. Bioanal. Chem.* **2015**, DOI: 10.1007/s00216-014-8399-y.

You can have a preview of the MOOC in Moodle environment as a guest. Guest access allows you to view the course contents and read the discussions in forums, posted during the Spring 2014 edition of the MOOC. As a guest you cannot post to forums and take quizzes.

Short description of the course

This is an introductory course on estimation of measurement uncertainty, specifically related to chemical analysis (analytical chemistry). The course gives the main concepts and mathematical apparatus of measurement uncertainty estimation and introduces two principal approaches to measurement uncertainty estimation – the ISO GUM modeling approach (the "bottom-up" or modeling approach) and the single-lab validation approach as implemented by Nordtest (the "top-down" or Nordtest approach). The course contains lectures, practical exercises and numerous tests for self-testing.

In spite of being introductory, the course intends to offer sufficient knowledge and skills for carrying out uncertainty estimation for most of the common chemical analyses in routine laboratory environment. The techniques for which there are examples or exercises include acid-base titration, Kjeldahl nitrogen determination, UV-Vis spectrophotometry, atomic absorption spectroscopy and liquid chromatography mass spectrometry (LC-MS). It is important to stress, however, that for successful measurement uncertainty estimation experience (both in analytical chemistry as such and also in uncertainty estimation) is crucial and this can be acquired only through practice.

The materials of this course can also be useful for people who do not intend to follow the full course but only want to find answers to some specific questions.

Required preliminary knowledge

Introductory level knowledge of analytical chemistry is required. More advanced knowledge of analytical chemistry and introductory knowledge of mathematical statistics is an advantage. Fluency with and access to a spreadsheet software package (MS Excel, OpeOffice, etc) is highly recommended.

Why is measurement uncertainty important

http://www.uttv.ee/naita?id=17711

https://www.youtube.com/watch?v=tn2DLYA72Dk

Study outcomes

The student who has successfully passed the course knows:

- the main concepts related to measurement results and measurement uncertainty, including their application to chemical analysis;
- the main mathematical concepts and tools in uncertainty estimation;
- the main measurement uncertainty sources in chemical analysis;
- the main approaches for measurement uncertainty estimation.

The student who has successfully passed the course is able to:

- decide what data are needed for uncertainty estimation, understand the meaning of the available data and decide whether the available data are sufficient;
- select the uncertainty estimation approach suitable for the available data;
- quantify the uncertainty contributions of the relevant uncertainty sources using the available data;
- carry out estimation of uncertainty using the main approaches of uncertainty estimation.

Organization of the course material

The course (overall volume 1 ECTS) is organized in 12 sections, of which some are in turn split into smaller subsections. The following parts are found in the sections:

1. The sections (and also many subsection) start with a **brief introduction** stating the main topic(s) and study outcomes of the section.
2. The main topic of the respective section is explained in a short **video lecture**.
3. The lecture is followed by a textual part. This text is in most cases meant to complement, not substitute the lecture (although in some cases the contents of the lecture are also repeated in some extent). It rather gives additional explanations and addresses some additional topics that were not covered by the lecture.
4. Most sections end with a self-test, which enables to test the acquired knowledge and skills. The tests contain questions, as well as calculation problems. The self-tests are on one hand meant for the students to monitor his/her progress. On the other hand, however, they also promote thinking and provide (by the feedback of the questions) additional knowledge about measurement uncertainty estimation in different practical situations. So, the self-tests are an intrinsic component of the course and it is strongly recommended to take all of them.

The printout of the current version of the course materials (including lecture slides) can be downloaded from **here**. The text surrounded with grey dashed line is footnote text and appears also at the end of the page.

If you consistently get a message "Server not found" when attempting to watch videos then with high probability the reason is the firewall of your local network. The local network administrators should enable outcoming connections from your network via port 1935. More specifically it is necessary to access the server rtmp://flash.ut.ee:1935.

An additional possibility is to watch the videos in YouTube via channel "ESTIMATION OF MEASUREMENT UNCERTAINTY IN CHEMICAL ANALYSIS".
Direct link: https://www.youtube.com/channel/UCeNhxB_WuTDNcbNHFxsBjUw

Self-testing

Throughout the course there are numerous self-tests for enabling the student to test his/her knowledge and skills in specific topics. Each test is graded as a percentage (100% corresponding to correctly answering all questions and correctly solving all problems).

Feedback is given as:

Correct answer, correctly recognised and marked by the student.

Correct answer, not recognised and not marked by the student.

Incorrect answer, however, considered correct by the student.

Explanatory feedback is displayed when wrong answer is selected. All tests can be taken as many times as needed and the success of taking these tests will not influence the final grade. We recommend that you take all the tests and work with them until you achieve score 100% and only then move to next topic.

Terminology and definitions

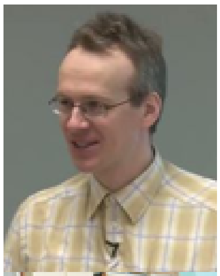
Wherever possible, the used terminology adheres to the 3rd edition of the *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*, JCGM 200:2008, International vocabulary of metrology — Basic and general concepts and associated terms (VIM), 3rd edition. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, 2008. Available for downloading free of charge at http://www.bipm.org/en/publications/guides/vim.html. referred to as "VIM" throughout the course. However, in the interest of better understanding and in order to stress the most important aspects of concepts, in many cases concepts are introduced by definitions that are somewhat simplified compared to the VIM. More deeply interested students are encouraged to consult the VIM.

Main literature sources

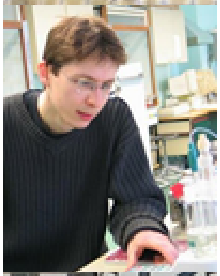
This list of literature references is selecive, not exhauxtive. The references were selected based on the following criteria: (1) Widely used and cited; (2) useful under practical lab conditions (i.e. not too deeply scientific); (3) a fairly recent version is available and (4) the document is preferably available free of charge on the Internet. These references are referred to in the course via superscript numbers in round brackets, e.g.: VIM(1).

(1) JCGM 200:2008, *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*, 3rd edition. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, **2008**. Available on-line from http://www.bipm.org/en/publications/guides/vim.html
(2) JCGM 100:2008 *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. JCGM, **2008**. Available on-line from http://www.bipm.org/en/publications/guides/gum.html
(3) *Quantifying Uncertainty in Analytical Measurement*, 2nd ed.; Ellison, S. L. R.; Williams, A., Eds.; EURACHEM/CITAC, **2012**. Available on-line from http://eurachem.org/index.php/publications/guides
(4) *Measurement Uncertainty Revisited*. Eurolab Technical Report No 1/2007. Eurolab, **2007**. Available on-line from http://www.eurolab.org/documents/1-2007.pdf
(5) *Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories*. B. Magnusson, T. Näykki, H. Hovind, M. Krysell. Nordtest technical report 537, ed. 3. Nordtest, **2011**. Available on-line from http://www.nordtest.info/index.php/technical-reports/item/handbook-for-calculation-of-measurement-uncertainty-in-environmental-laboratories-nt-tr-537-edition-3.html
(6) *Analytical Measurement: Measurement Uncertainty and Statistics*. Eds: N. Majcen, V. Gegevicius. EC-JRC IRMM, **2012**. Available on-line from http://publications.jrc.ec.europa.eu/repository/ bitstream/111111111/29537/1/lana2207enn-web.pdf

Course team



Ivo Leito, professor of analytical chemistry at University of Tartu.
Ivo teaches analytical chemistry and metrology in chemistry at all study levels and organizes short training courses for practitioners on different topics of analytical chemistry and metrology in chemistry. His research work embraces a wide area of topics ranging from studies of superacids and superbases to LC-MS analysis. He is the initiator of the master's programme Applied Measurement Science at University of Tartu.



Lauri Jalukse, research fellow in analytical chemistry at University of Tartu.
Lauri teaches analytical chemistry and metrology in chemistry at all study levels. He is continuously introducing innovative and active learning approaches into teaching. His research work is focused on metrological studies of electrochemical and optical sensors, measurements of dissolved oxygen concentration and moisture content, as well as organization of interlaboratory comparisons.

Irja Helm, research fellow in analytical chemistry at University of Tartu.
Irja teaches practical classes of analytical chemistry. She takes care that metrological concepts and approaches are introduced to students at as early stage of analytical chemistry studies as possible.

Technical design: Educational Technology Centre, University of Tartu.

Contact and feedback

Ivo Leito
University of Tartu
Institute of Chemistry
Ravila 14a, Tartu 50411
ESTONIA
Tel: +372-5 184 176
e-mail: ivo.leito@ut.ee
Skype: leitoivo

Copyright

***

[1] JCGM 200:2008, *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*, 3rd edition. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, **2008**. Available for downloading free of charge at http://www.bipm.org/en/publications/guides/vim.html.

[2] This list of literature references is selecive, not exhauxtive. The references were selected based on the following criteria: (1) Widely used and cited; (2) useful under practical lab conditions (i.e. not too deeply scientific); (3) a fairly recent version is available and (4) the document is preferably available free of charge on the Internet. These references are referred to in the course via superscript numbers in round brackets, e.g.: VIM(1).

u_mooc_pdf_printout_2014d.pdf 1.62 MB

1. The concept of measurement uncertainty (MU)

Brief summary: This section introduces the concepts of **measurand**, **true value**, **measured value**, **error**, **measurement uncertainty** and **probability**.

The concept of measurement uncertainty

http://www.uttv.ee/naita?id=17583

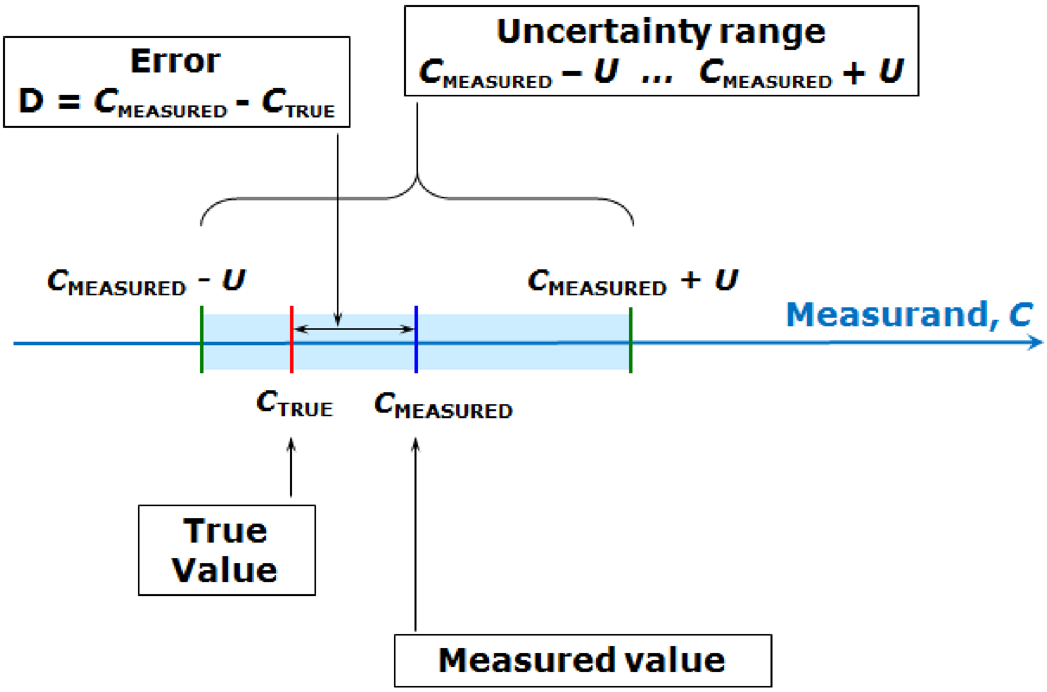
https://www.youtube.com/watch?v=BogGbA0hC3k

Measurement is a process of experimentally obtaining the value of a quantity. The quantity that we intend to measure is called **measurand**. In chemistry the measurand is usually the content (concentration) of some chemical entity (molecule, element, ion, etc) in some object. The chemical entity that is intended be determined is called **analyte**. Measurands in chemistry can be, for example, lead concentration in a water sample, content of pesticide thiabendazole in an orange or fat content in a bottle of milk. In the preceding example lead (element), ascorbic acid (molecule) and fat (group of different molecules) are the analytes. Water, orange and milk are **analysis objects** (or **samples** taken from analysis objects).

In principle, the aim of a measurement is to obtain the **true value** of the measurand. Every effort is made to optimize the **measurement procedure** (in chemistry chemical analysis procedure or analytical procedure Analytical chemists mostly use the term „analytical method". In this course we use the term „procedure" instead of „method", as this usage is supported by the VIM.) in such a way that the measured value is as close as possible to the true value. However, our measurement result will be just an estimate of the true value and the actual true value will (almost) always remain unknown to us. Therefore, we cannot know exactly how near our measured value is to the true value – our estimate always has some uncertainty associated with it.

The difference between the measured value and the true value is called **error.** Error can have either positive or negative sign. Error can be regarded as being composed of two parts – **random error** and **systematic error** – which will be dealt with in more detail in coming lectures. Like the true value, also the error is not known to us. Therefore it cannot be used in practice for characterizing the quality of our measurement result – its agreement with the true value.

The quality of the measurement result, its **accuracy**, is characterized by **measurement uncertainty** (or simply **uncertainty**), which defines an interval around the measured value $C_{MEASURED}$, where the true value C_{TRUE} lies with some probability. The measurement uncertainty U itself is the half-width of that interval and is always non-negative. Here and in the lecture the capital U is used to denote a generic uncertainty estimate. The symbol U is picked on purpose, because expanded uncertainty (generally denoted by capital U) fits very well with the usage of uncertainty in this section. However, it is not explicitly called expanded uncertainty here, as this term will be introduced in later lectures. The following scheme (similar to the one in the lecture) illustrates this:



Scheme 1.1. Interrelations between the concepts true value, measured value, error and uncertainty.

Measurement uncertainty is always associated with some probability – as will be seen in the next lectures, it is usually not possible to define the uncertainty interval in such a way that the true value lies within it with 100% probability.

Measurement uncertainty, as expressed here, is in some context also called the **absolute measurement uncertainty**. This means that the measurement uncertainty is expressed in the same units as the measurand. As will be seen in subsequent lectures, it is sometimes more useful to express measurement uncertainty as **relative measurement uncertainty**, which is the ratio of the absolute uncertainty U_{abs} and the measured value y:

$$U_{rel} = \frac{U_{abs}}{y} \qquad (1.1)$$

Relative uncertainty is a unitless quantity, which sometimes is also expressed as per cent.

Measurement uncertainty is different from error in that it does not express a difference between two values and it does not have a sign. Therefore it cannot be used for correcting the measurement result and cannot be regarded as an estimate of the error because the error has a sign. Instead measurement uncertainty can be regarded as our estimate, what is the highest probable absolute difference between the measured value and the true value. With high probability the difference between the measured value and the true value is in fact lower than the measurement uncertainty. However, there is a low probability that this difference can be higher than the measurement uncertainty.

Both the true value and error (random and systematic) are abstract concepts. Their exact values cannot be determined. However, these concepts are nevertheless useful, because their **estimates** can be determined and are highly useful. In fact, as said above, our measured value is an estimate of the true value.

***

[1] Analytical chemists mostly use the term „analytical method“. In this course we use the term „procedure“ instead of „method“, as this usage is supported by the VIM.

[2] Here and in the lecture the capital U is used to denote a generic uncertainty estimate. The symbol U is picked on purpose, because expanded uncertainty (generally denoted by capital

U) fits very well with the usage of uncertainty in this section. However, it is not explicitly called expanded uncertainty here, as this term will be introduced in later lectures.

2. The origin of measurement uncertainty

Brief summary: Explanation, on the example of pipetting, where measurement uncertainty comes from. The concept of **uncertainty sources** – effects that cause the deviation of the measured value from the true value – is introduced. The main uncertainty sources of pipetting are introduced and explained: repeatability, calibration, temperature effect. Explanation of random and systematic effects is given. The concept of **repeatability** is introduced.

The first video demonstrates how pipetting with a classical volumetric pipette is done and explains where the uncertainty of the pipetted volume comes from.

Why measurement results have uncertainty? The concept of uncertainty source explained on the example of pipetting

http://www.uttv.ee/naita?id=17577

https://www.youtube.com/watch?v=ufWJB9orWdU

The second video demonstrates pipetting with a modern automatic pipette and explains the uncertainty sources in pipetting with an automatic pipette. This second video explains how to pipet with an automatic pipette if accurate volume is desired. In many routine, high-volume applications (e.g. in biochemistry), however, speed is more impoirtant than accuracy and in such cases some steps shown here, most importantly, rinsing, can be omitted. Also, in some cases there is a very limited volume available of the solution that is pipetted – in such case also rinsing is not possible. On the other hand, if still higher accuracy is desired then the so-called reverse pipetting technique can be used. Reverse pipetting is more accurate than the commonly used forward pipetting, which is the technique demonstrated and explained in this video. .

Measurement uncertainty sources of pipetting with an automatic pipette

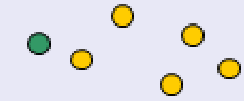
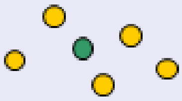
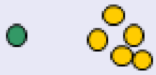
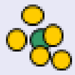
http://www.uttv.ee/naita?id=18164

https://www.youtube.com/watch?v=hicLweJcJWY

Measurement results have uncertainty because there are **uncertainty sources** (effects that cause uncertainty). These are effects that cause deviations of the measured value from the true value. These sources also cause the existence of error and could therefore also be called error sources. If the used measurement procedure is well known then the most important uncertainty sources are usually also known. Efforts should be made to minimize and, if possible, eliminate uncertainty sources by optimizing the measurement procedure (analysis procedure). The uncertainty sources that cannot be eliminated (and it is never possible to eliminate all uncertainty sources) have to be taken into account in uncertainty estimation.

The magnitudes of the deviations caused by uncertainty sources are usually unknown and in many cases cannot be known. Thus, they can only be estimated. If we can estimate the magnitudes of all important uncertainty sources then we can combine them and obtain the estimate of measurement uncertainty, which in this case will be called **combined measurement uncertainty**. How this combining is mathematically done, will be demonstrated in the coming lectures.

If we make a number of repeated measurements of the same measurand then ideally all these repeated measurements should give exactly the same value and this value should be equal to the true value of the measurand. In reality the results of the repeated measurements almost always differ to some extent and their mean value also usually differs from the true value. The uncertainty sources cause this. In a somewhat simplified way the uncertainty sources (or effects) can be divided into **random effects** and **systematic effects**. It is in principle not wrong to call them random and systematic sources of uncertainty, but this is not usually done. This is largely because, as we will see in a coming lecture, the concept of measurement uncertainty stresses that random and systematic effects should be treated the same way. The following scheme illustrates this (green circles denote true values, yellow circles denote measured values):

Situation	Random effects	Systematic effects	Uncertainty
1. 	Strong	Strong	High
2. 	Strong	Weak (or absent)	Medium
3. 	Weak	Strong	Medium
4. 	Weak	Weak (or absent)	Low

Scheme 2.1. The influence of random and systematic effects on measurement uncertainty.

Random effects cause the difference between the repeated measurement results (and thus, obviously, also from the true value). However, if a large number of repeated measurements are made then the mean value will have little influence from the random effects (situation 2 on the scheme). So, the influence of random effects can be decreased by increasing the number of repetitions. Systematic effects cause deviation of all measurements in the series into the same direction by the same magnitude. It is more correct to say „by a predictable magnitude". This means that the magnitude is not necessarily always the same – it can vary, e.g. as the magnitude of the measurand value varies – but it can be predicted, i.e. it is not random. Increasing the number of repetitions does not enable decreasing their influence (situation 3 on the scheme).

In principle it is desirable to determine the magnitude and direction of the systematic effects and correct the measurement results for the systematic effects. However, it can often be so difficult and work-intensive, that it becomes impractical. Therefore in many cases, rather than accurately determining the systematic effects and correcting for them their possible magnitudes is estimated and are taken into account as uncertainty sources. In lectures 5.4 and 6 random and systematic effects are treated more comprehensively.

There are in general four main sources of uncertainty in volumetric measurements, i.e. measurements by pipettes, burettes, measuring cylinders and volumetric flasks

- Uncertainty due to the **non-ideal repeatability** of measurement (often called **repeatability uncertainty)**. In the case of pipetting this means that however carefully we try to fill and empty the pipette, we will nevertheless every time get a slightly different volume. This is sometimes referred to as the "human effect" or the "human factor", but in fact, if a machine would do the pipetting then there would also be difference between the volumes (although probably smaller). Repeatability is a typical random effect and contributes to uncertainty with glass pipettes as well as with automatic pipettes. Its influence on the measurement result can be decreased by making repeated measurements but it can never be entirely eliminated.
- Uncertainty due to calibration of the volumetric equipment (often called **calibration uncertainty**). In the case of volumetric glassware this is the uncertainty in the positions of the marks on the volumetric glassware. In the case of automatic pipettes this uncertainty is caused by the systematically too high or too low displacement of the piston inside the pipette. In the case of a given pipette it is a typical systematic uncertainty source. This uncertainty source can be significantly reduced by recalibrating the pipette in the laboratory by the person who actually works with it. Accurate weighing of water at

controlled temperature is the basis of calibration of volumetric instruments and also the way how usually the repeatability uncertainties of different volumetric instruments are found.

- Uncertainty due to the temperature effect (often called as **temperature uncertainty**). All volumetric ware is usually calibrated at 20 °C and volumes usually refer to volumes at 20 °C. The density of the liquid changes (almost always decreases) with temperature. If pipetting is done at a higher temperature than 20 °C then there amount of liquid (in terms of mass or number of molecules) pipetted is smaller than if it were done at 20 °C. Consequently, the volume of that amount of liquid at 20 °C is also smaller than if the pipetting were done at 20 °C. In the case of volumetric glassware temperature affects the dimensions of the volumetric ware (its volume increases with temperature). The effect of liquid density change is ca 10 times stronger. Therefore the volume change of volumetric glassware is almost always neglected. In the case of automatic pipettes the effect of temperature is more complex. If the air inside the pipette warms then the delivered liquid volume can change to some extent. If the temperature of the laboratory and, importantly, temperature of the pipetted liquid, is constant during repeated measurements then the temperature effect is a systematic effect.
- Application-specific uncertainty sources. These are not caused by the volumetric equipment but by the liquid that is handled or by the system that is investigated. Some examples:
    - If a foaming solution is pipetted, measured by a volumetric flask or a measuring cylinder then it is not clear where exactly the solution "ends", i.e., there is no well-defined meniscus. This will cause an additional uncertainty. Depending on situation this effect can be random or systematic or include both random and systematic part.
    - If titration is carried out using visual indicator then the end-point of titration, i.e. the moment when the indicator changes color is assumed to match the equivalence point (the stoichiometry point). However, depending on the titration reaction and on the actual analyte that is titrated, the end-point may come earlier or later than the equivalence point. In the case of the given titration this will be a systematic effect. This effect can be minimized by some other means of end-point detection, e.g. potentiometric titration. In fact, even if the end-point is determined potentiometrically, it can still have some systematic deviation from the equivalence point. However, this effect is usually so small that it can be neglected.

There are some other uncertainty sources that usually turn out to be less important, because they can be minimized or eliminated by correct working practices (however, they can be important if these correctly practices are not applied). The remaining effects will usually influence the repeatability of pipetting or its calibration uncertainty and can be taken into account within those uncertainty sources.

- If the pipette is not kept vertically (both glass and automatic pipettes), waiting is not long enough after the end of drainage of solution (glass pipette) then the pipetted volume will be lower than the one obtained with correct pipetting. No waiting is needed in the case of automatic pipettes because no liquid film remains (and must not remain) on pipette to inner walls.
- When using a glass pipette then there is always some possibility that small residues of the previous solution are still in the pipette. It is therefore a good idea to rinse the pipette before pipetting (e.g. two times) with the solution that will be pipetted (and discarding the rinse solution it into waste, not into the vessel from where the solution is taken). In the case of automatic pipettes it is a good idea to use a new tip every time e new solution is pipetted. In that case such contamination is usually negligible. Also, when pipetting the same solution numerous times with the same pipetteit is a good idea to monitor it for the absence of droplets on the inner walls and replace the tip when the droplets appear.
- If the walls of a glass pipette are not clean then droplets may remain on the walls after the pipetted solution has been drained. This leads to a different volume from the case when no droplets are left on pipette walls after draining the solution. The obvious thing to do is to clean the pipette.
- If the pipetted liquid is very different from water (e.g. some highly viscous liquid, such as vegetable oil) then the pipetted volume may be systematically different from the nominal volume of the pipette. This effect exists both with glass pipettes and with automatic

pipettes. In such a case the pipette should either be recalibrated using the liquid under question or weighing should be used instead of volumetry In the case of automatic pipettes also reverese pipetting instead of the more common forward pipetting can be used to decrease the uncertainty when pipetting viscous liquids. .

In section 4 the uncertainty sources of pipetting (the same pipetting experiment that was performed in the video) will be quantified and combined into the measurement uncertainty estimate of pipetted volume. Sections 4.1 to 4.5 present the uncertainty calculation using a factory-calibrated pipette. Section 4.6 presents an example of measurement uncertainty calculation of pipetted volume using a self-calibrated pipette. In section 5 an overview of the majority of uncertainty sources that are encountered in chemical analysis will be given.

***

[1] This second video explains how to pipet with an automatic pipette if accurate volume is desired. In many routine, high-volume applications (e.g. in biochemistry), however, speed is more impoirtant than accuracy and in such cases some steps shown here, most importantly, rinsing, can be omitted. Also, in some cases there is a very limited volume available of the solution that is pipetted – in such case also rinsing is not possible. On the other hand, if still higher accuracy is desired then the so-called reverse pipetting technique can be used. Reverse pipetting is more accurate than the commonly used forward pipetting, which is the technique demonstrated and explained in this video.

[2] These sources also cause the existence of error and could therefore also be called error sources.

[3] It is in principle not wrong to call them random and systematic sources of uncertainty, but this is not usually done. This is largely because, as we will see in a coming lecture, the concept of measurement uncertainty stresses that random and systematic effects should be treated the same way.

[4] It is more correct to say „by a predictable magnitude". This means that the magnitude is not necessarily always the same – it can vary, e.g. as the magnitude of the measurand value varies – but it can be predicted, i.e. it is not random.

[5] In fact, even if the end-point is determined potentiometrically, it can still have some systematic deviation from the equivalence point. However, this effect is usually so small that it can be neglected.

[6] In the case of automatic pipettes also reverese pipetting instead of the more common forward pipetting can be used to decrease the uncertainty when pipetting viscous liquids.

Brief summary: This section presents the most basic concepts and tools for practical estimation of measurement uncertainty. First, the concepts of random quantities and distribution functions are introduced. Then the Normal distribution – the most important distribution function in measurement science – is explained and its two main parameters – the mean value and standard deviation – are introduced (3.1). Based on standard deviation the concept of standard uncertainty is introduced (3.1, 3.2). Thereafter the A type and B type uncertainty estimation is introduced (3.3). The mean value of random quantities is also a random quantity and its reliability can be described by the standard deviation of the mean (3.4). Besides the normal distribution three more distribution functions are introduced: rectangular and triangular distribution (3.5) as well as the Student distribution (3.6).

3. The basic concepts and tools

3.1. The Normal distribution

Brief summary: This lecture starts by generalizing that all measured values are **random quantities** from the point of view of mathematical statistics. The most important distribution in measurement science – the **Normal distribution** – is then explained: its importance, the parameters of the Normal distribution (**mean** and **standard deviation**). The initial definitions of **standard uncertainty** (*u*), **expanded uncertainty** (*U*) and **coverage factor** (*k*) are given. A link between these concepts and the Normal distribution is created.

The Normal distribution

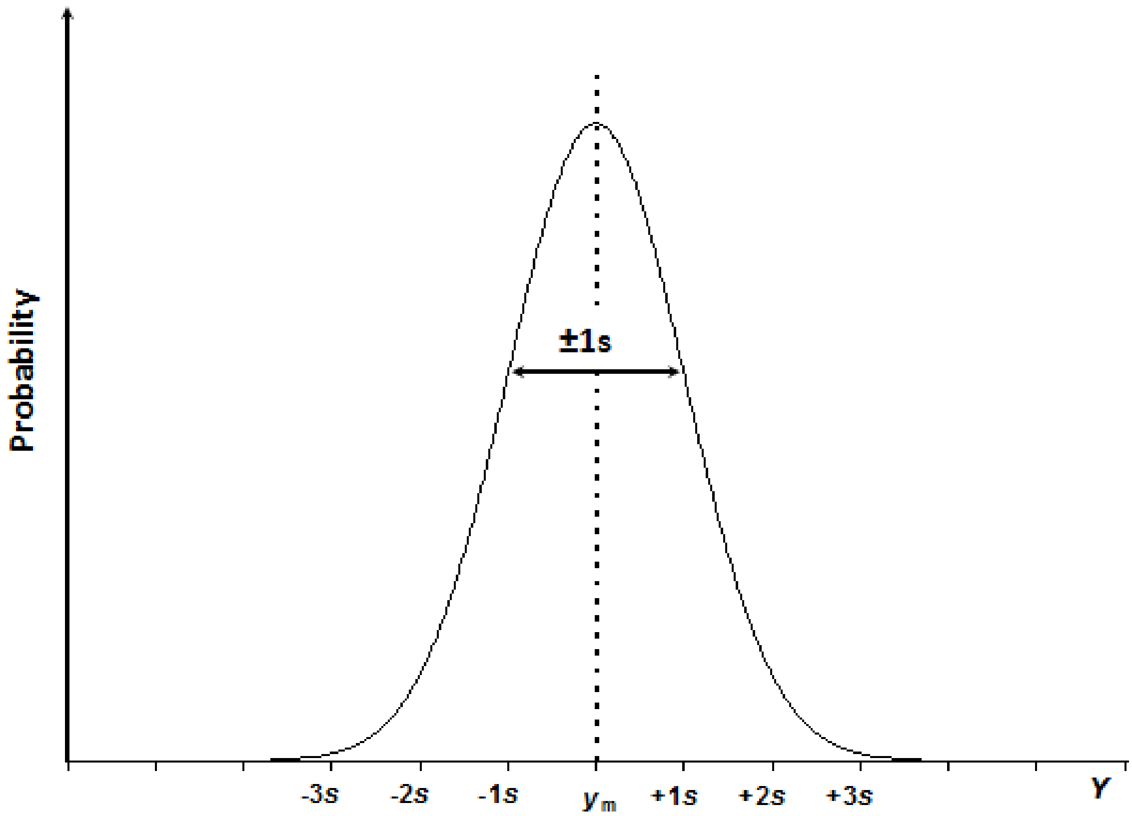
http://www.uttv.ee/naita?id=17589

https://www.youtube.com/watch?v=N-F6leWyNZk

All measured quantities ~~(measurands)~~ are from the point of view of mathematical statistics random quantities. Random quantities can have different values. This was demonstrated in the lecture on the example of pipetting. If pipetting with the same pipette with nominal volume 10 ml is repeated multiple times then all the pipetted volumes are around 10 ml, but are still slightly different. If a sufficiently large number of repeated measurements are carried out and if the pipetted volumes It is fair to ask, how do we know the individual pipetted volumes if the pipette always „tells" us just that the volume is 10 ml? In fact, if we have only the pipette and no other (more accurate) measurement possibility of volume then we cannot know how much the volumes differ form each other or from the nominal volume. However, if a more accurate method is available then this is possible. In the case of pipettig a very suitable and often used more accurate method is weighing. It is possible to find the volume of the pipetted water, which is more accurate than that obtained by pipetting, by weighing the pipetted solution (most often water) and divided the obtained mass by the density of water at the temperature of water. Water is used in such experiments because the densities of water at different temperatures are very accurately known (see e.g. http://en.wikipedia.org/wiki/Properties_of_water#Density_of_water_and_ice). are plotted according to how frequently they are encountered then it becomes evident that although random, the values are still governed by some underlying relationship between the volume and frequency: the maximum probability of a volume is somewhere in the range of 10.005 and

10.007 ml and the probability gradually decreases towards smaller and larger volumes. This relationship is called **distribution function**.

There are numerous distribution functions known to mathematicians and many of them are encountered in the nature, i.e. they describe certain processes in the nature. In measurement science the most important distribution function is the **normal distribution** (also known as the Gaussian distribution). Its importance stems from the so-called **Central limit theorem**. In a simplified way it can be worded for measurements as follows: if a measurement result is simultaneously influenced by many uncertainty sources then if the number of the uncertainty sources approaches infinity then the distribution function of the measurement result approaches the normal distribution, irrespective of what are the distribution functions of the factors/parameters describing the uncertainty sources. In reality the distribution function of the result becomes indistinguishable from the normal distribution already if there are 3-5 (depending on situation) significantly contributing Significantly cointributing uncertainty sources are the important uncertainty sources. We have already qualitatively seen in section 2 that different uncertainty sources have different „importance". In the coming lectures we will also see how the „importance" of an uncertainty source (its uncertainty contribution) can be quantitatively expressed. uncertainty sources. This explains, why in so many cases measured quantities have normal distribution and why most of the mathematical basis of measurement science and measurement uncertainty estimation is based on the normal distribution.



Scheme 3.1. The normal distribution curve of quantity *Y* with mean value y_m and standard deviation *s*.

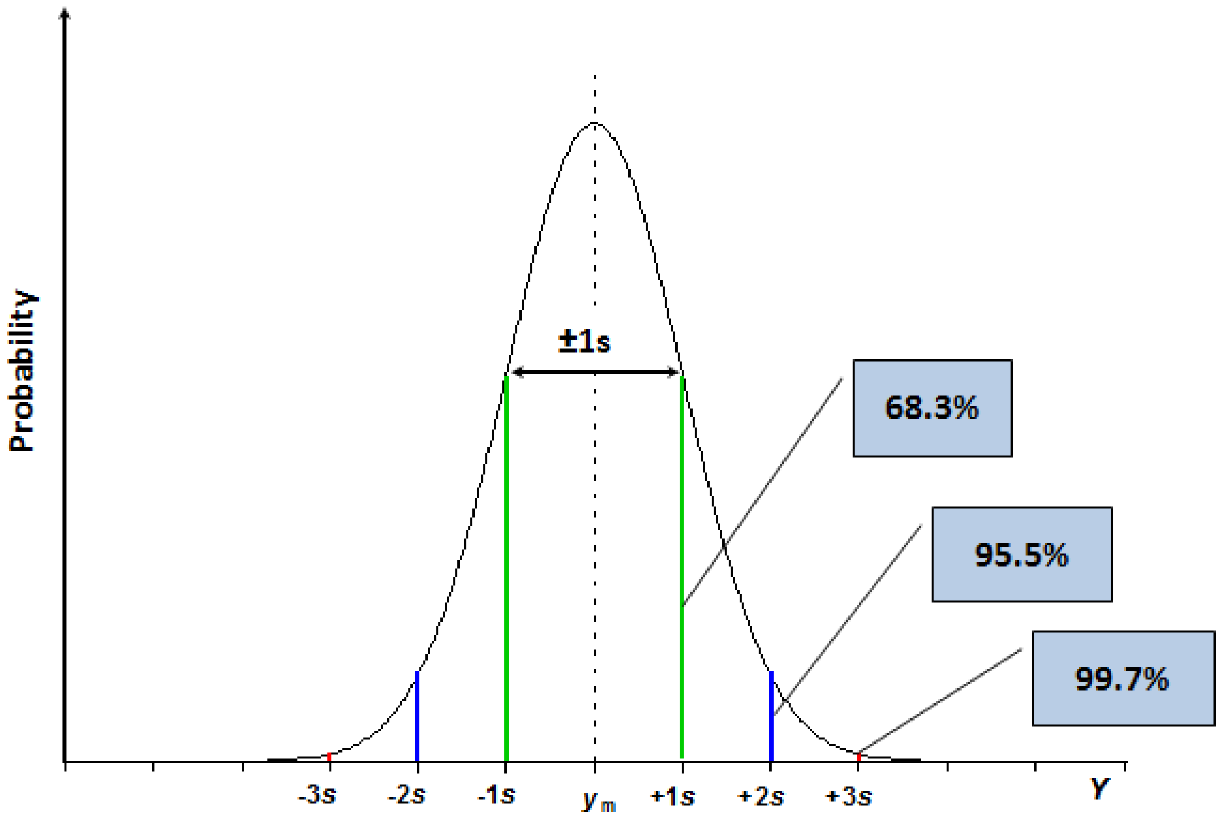
The normal distribution curve has the bell-shaped appearance (Scheme 3.1), and is expressed by equation 3.1:

$$f(y) = \frac{1}{s\sqrt{2\pi}} e^{-\frac{(y-y_m)^2}{2s^2}} \qquad (3.1)$$

In this equation $f(y)$ is the probability that the measurand Y has value y. y_m is the **mean** value of the **population** and s is the **standard deviation** of the population. y_m characterizes the position of the normal distribution on the Y axis, s characterizes the width (spread) of the distribution function, which is determined by the scatter of the data points. The mean and standard deviation are the two parameters that fully determine the shape of the normal distribution curve of a particular random quantity. The constants 2 and p are normalization factors, which are present in order to make the overall area under the curve equal to 1.

The word "population" here means that we would need to do an infinite number of measurements in order to obtain the true y_m and *s* values. In reality we always operate with a limited number of measurements, so that the mean value and standard deviation that we have from our experiments are in fact **estimates** of the true mean and true standard deviation. The larger is the number of repeated measurements the more reliable are the estimates. The number of parallel measurements is therefore very important and we will return to it in different other parts of this course.

The normal distribution and the standard deviation are the basis for definition of **standard uncertainty**. Standard uncertainty, denoted by *u*, is the uncertainty expressed at standard deviation level, i.e., uncertainty with roughly 68.3% coverage probability (i.e. the probability of the true value falling within the uncertainty range is roughly 68.3%). The probability of 68.3% is often too low for practical applications. Therefore uncertainty of measurement results is in most cases not reported as standard uncertainty but as **expanded uncertainty**. Expanded uncertainty, denoted by *U*, is obtained by multiplying standard uncertainty with a **coverage factor**, This definition of expanded uncertainty is simplified. A more rigorous definition goes via the combined standard uncertainty and is introduced in section 4.4. denoted by *k*, which is a positive number, larger than 1. If the coverage factor is e.g. 2 (which is the most commonly used value for coverage factor) then in the case of normally distributed measurement result the coverage probability is roughly 95.5%. These probabilities can be regarded as fractions of areas of the respective segments from the total area under the curve as illustrated by the following scheme:



Scheme 3.2. The same normal distribution curve as in Scheme 3.1 with 2*s* and 3*s* segments indicated.

Since the exponent function can never return a value of zero, the value of $f(y)$ (eq 3.1) is higher than zero with any value of *y*. This is the reason why uncertainty with 100% coverage is (almost) never possible.

It is important to stress that these percentages hold only if the measurement result is normally distributed. As said above, very often it is. There are, however, important cases when measurement result is not normally distributed. In most of those cases the distribution function has “heavier tails”, meaning, that the expanded uncertainty at e.g. $k = 2$ level will not correspond to coverage probability of 95.5%, but less (e.g. 92%). The issue of distribution of the measurement result will be addressed later in this course.

***

[1] It is fair to ask, how do we know the individual pipetted volumes if the pipette always „tells“ us just that the volume is 10 ml? In fact, if we have only the pipette and no other (more accurate) measurement possibility of volume then we cannot know how much the volumes differ form each other or from the nominal volume. However, if a more accurate method is available then this is possible. In the case of pipettig a very suitable and often used more accurate method is weighing. It is possible to find the volume of the pipetted water, which is more accurate than that obtained by pipetting, by weighing the pipetted solution (most often water) and divided the obtained mass by the density of water at the temperature of water. Water is used in such experiments because the densities of water at different temperatures are very accurately known (see e.g. http://en.wikipedia.org/wiki/Properties_of_water#Density_of_water_and_ice).

[2] Significantly cointributing uncertainty sources are the important uncertainty sources. We have already qualitatively seen in section 2 that different uncertainty sources have different „importance“. In the coming lectures we will also see how the „importance“ of an uncertainty source (its uncertainty contribution) can be quantitatively expressed.

[3] This definition of expanded uncertainty is simplified. A more rigorous definition goes via the *combined* standard uncertainty and is introduced in section 4.4.

3.2. Mean, standard deviation and standard uncertainty

Brief summary: the lecture explains calculation of **mean** (V_m) and **standard deviation** (s). Illustrates again the 68% probability of s. Explains how the **standard uncertainty** of repeatability u (V, REP) can be estimated as standard deviation of parallel measurement results. Stresses the importance of standard uncertainty as the key parameter in carrying out uncertainty calculations: uncertainties corresponding to different sources (not only to repeatability) and to different distribution functions are converted to standard uncertainties when uncertainty calculations are performed.

Mean, standard deviation and standard uncertainty

http://www.uttv.ee/naita?id=17554

https://www.youtube.com/watch?v=ND3iryaVQ68

One of the most common approaches for improving the reliability of measurements is making replicate measurements of the same quantity. In such a case very often the measurement result is presented as the **mean value** of the replicate measurements. In the case of pipetting n times with the same pipette volumes V_1, V_2, …, V_n are obtained and the mean value V_m is calculated as follows:

$$V_{\mathrm{m}} = \frac{V_1 + V_2 + \ldots + V_n}{n} = \frac{\sum_{i=1}^{n} V_i}{n} \qquad (3.2)$$

As explained in section 3.1, the mean value calculated this way is an estimate of the true mean value (which could be obtained if it were possible to make an infinite number of measurements).

The scatter of values obtained from repeated measurements is characterized by **standard deviation** of pipetted volumes, which for the same case of pipetting is calculated as follows:

$$s(V) = \sqrt{\frac{\sum_{i=1}^{n} (V_i - V_{\mathrm{m}})^2}{n-1}} \qquad (3.3)$$

The n − 1 in the denominator is often called **number of degrees of freedom**. We will see later that this is an important characteristic of a set or repeated measurements. The higher it is the more reliable mean and standard deviation can be from the set.

Two important interpretations of the standard deviation:

1. If V_m and s (V) have been found from a sufficiently large number of measurements (usually 10-15 is enough) then the probability of every next measurement (performed under the same conditions) falling within the range V_m ± s (V) is roughly 68.3%.
2. If we make a number of repeated measurements under the same conditions then the standard deviation of the obtained values characterized the uncertainty due to non-ideal repeatability (often called as repeatability standard uncertainty) of the measurement: u (V, REP) = s(V). Non-ideal repeatability is one of the uncertainty sources in all measurements. We will see later that standard deviation of measurements repeated under conditions that change in predefined way (i.e. it is not repeatability) is also extremely

useful in uncertainty calculation, as it enables taking a number of uncertainty sources into account simultaneously.

Standard deviation is the basis of defining **standard uncertainty** – uncertainty at standard deviation level, denoted by small *u*. Three important aspects of standard uncertainty are worth stressing here:

1. Standard deviation can be calculated also for quantities that are not normally distributed. This enables to obtain for them standard uncertainty estimates.
2. Furthermore, also uncertainty sources that are systematic by their nature and cannot be evaluated by repeating measurements can still be expressed numerically as standard uncertainty estimates.
3. Converting different types of uncertainty estimates to standard uncertainty is very important, because as we will see in section 4, most of the calculations in uncertainty evaluation, especially combining the uncertainties corresponding to different uncertainty sources, are carried out using standard uncertainties.

Standard uncertainty of a quantity (in our case volume *V*) expressed in the units of that quantity is sometimes also called absolute standard uncertainty. Standard uncertainty of a quantity divided by the value of that quantity is called **relative standard uncertainty**, u_{rel} (similarly to eq 1.1). In the case of volume *V*:

$$u_{\mathrm{rel}}(V) = \frac{u(V)}{V} \quad (3.4)$$

***

[1] We will see later that standard deviation of measurements repeated under conditions that changer in predefined way (i.e. it is not repeatability) is also extremely useful in uncertainty calculation, as it enables taking a number of uncertainty sources into account simultaneously.

3.3. A and B type uncertainty estimates

Carrying out the same measurement operation many times and calculating the standard deviation of the obtained values is one of the most common practices in measurement uncertainty estimation. Either the full measurement or only some parts of it can be repeated. In both cases useful information can be obtained. The obtained standard deviation (or the standard deviation of the mean, explained in section 3.4) is then the standard uncertainty estimate. Uncertainty estimates obtained as standard deviations of repeated measurement results are called **A type** uncertainty estimates. If uncertainty is estimated using some means other than statistical treatment of repeated measurement results then the obtained estimates are called **B type** uncertainty estimates. The other means can be e.g. certificates of reference materials, specifications or manuals of instruments, estimates based on long-term experience, etc.

Uncertainty estimates of A- and B-type

http://www.uttv.ee/naita?id=18165

https://www.youtube.com/watch?v=jdbx5UMQD9k

3.4. Standard deviation of the mean

Brief summary: Like the individual values, the mean value calculated from them is also a random quantity and for it also a standard deviation can be calculated. It is possible to calculate it from the standard deviation of the individual value. It is explained when to use the standard deviation of the individual value and when to use the standard deviation of the mean: whenever the individual result is used in further calculation the standard deviation of the individual result has to be used; whenever the mean value is used in further calculations, the standard deviation of the mean has to be used.

Standard deviation of the mean

http://www.uttv.ee/naita?id= 17580

https://www.youtube.com/watch?v=GLsHHIW1yjo

The standard deviation $s(V)$ calculated using the formula 3.3 is the standard deviation of an *individual* pipetting result (value). When the mean value is calculated from a set of individual values which are randomly distributed then the mean value will also be a random quantity. As for any random quantity, it is also possible to calculate standard deviation for the mean $s(V_m)$. One possible way to do that would be carrying out numerous measurement series, find the mean for every series and then calculate the standard deviation of all the obtained mean values. This is, however, too work-intensive. However, there is a very much simpler approach for calculating $s(V_m)$, simply divide the $s(V)$ by square root of the number of repeated measurements made:

$$s(V_\mathrm{m}) = \frac{s(V)}{\sqrt{n}} \quad (3.5)$$

So, for a set of repeated pipetting values we have in fact two standard deviations: standard deviation of the single value $s(V)$ and standard deviation of the mean $s(V_m)$. It is important to ask: when we use one and when another of them?

The general rule of thumb is the following: when the measured value reported or used in subsequent calculations is a single value then we use standard deviation of the single value; when it is the mean value then we use the standard deviation of the mean.

Let us illustrate this by two examples:

1. **Pipetting**. When we deliver a certain volume by a pipette then pipetting is a one-time operation: we cannot repeat the pipetting with the same liquid amount. So we use the standard deviation of single pipetting as pipetting repeatability uncertainty.
2. **Weighing**. When we weigh a certain amount of a material then we can weigh it repeatedly. So, if we need to minimize the influence of weighing repeatability in our measurement then we can weigh the material repeatedly and use in our calculations the mean mass. In this case the repeatability standard deviation of this mean mass is the standard deviation of the mean. If, on the other hand, it is not very important to have the lowest possible repeatability uncertainty of mass then we weigh only once and use the mass value from the single weighing and as its repeatability uncertainty we will use the standard deviation of a single value. As we will see later, modern balances are highly accurate instruments and uncertainty due to weighing is seldom among the important uncertainty sources. So, unless some disturbing effects interfere with weighing, it is usually not necessary to weigh materials with many repetitions.

In the case of single pipetting or single weighing the repeatability uncertainty of course cannot be estimated from this single operation. In these cases repeatability is determined separately and then used for the concrete measurements.

***

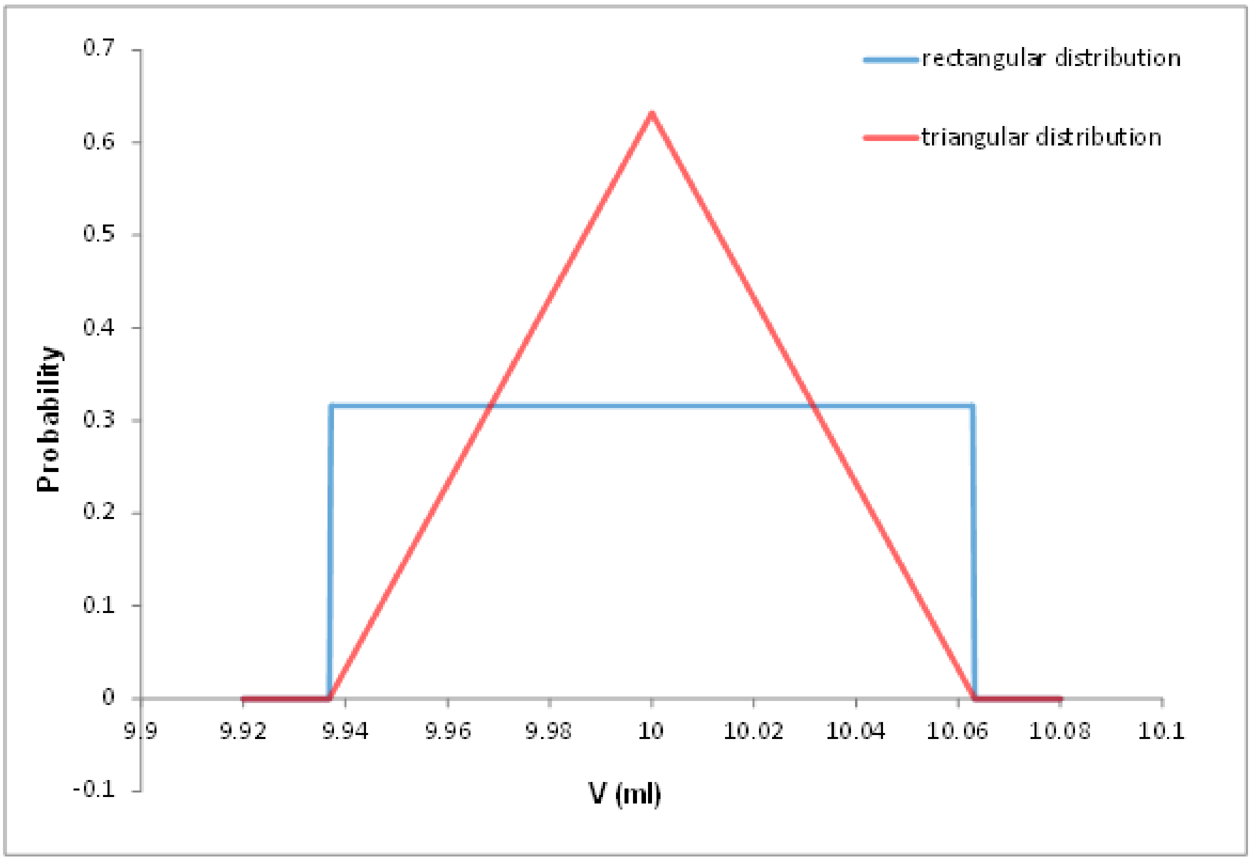
[1] As we will see later, modern balances are highly accurate instruments and uncertainty due to weighing is seldom among the important uncertainty sources. So, unless some disturbing effects interfere with weighing, it is usually not necessary to weigh materials with many repetitions.

Brief summary: Rectangular distribution and **triangular distribution** are explained, as well as how the uncertainties corresponding to rectangular or triangular distribution can be converted to standard uncertainties. Often the information on distribution function is missing and then usually some distribution function is assumed or postulated. Rectangular and triangular distributions are among the most common postulated distribution functions. Recommendations are given, which of these distributions to assume.

Other distribution functions: rectangular and triangular distribution

http://www.uttv.ee/naita?id=17584

https://www.youtube.com/watch?v=g_PefybO2Ao



Scheme 3.3. Rectangular and triangular distributions. Both of them correspond to the situation (10.000 ± 0.063) ml.

In measurement uncertainty estimation situations often occur where it is necessary to make choice between two alternatives of which one may possibly lead to somewhat overestimated uncertainty and the other one to somewhat underestimated uncertainty. In such situation it is usually reasonable to rather somewhat overestimate than underestimate the uncertainty.

3.5. Rectangular and triangular distribution

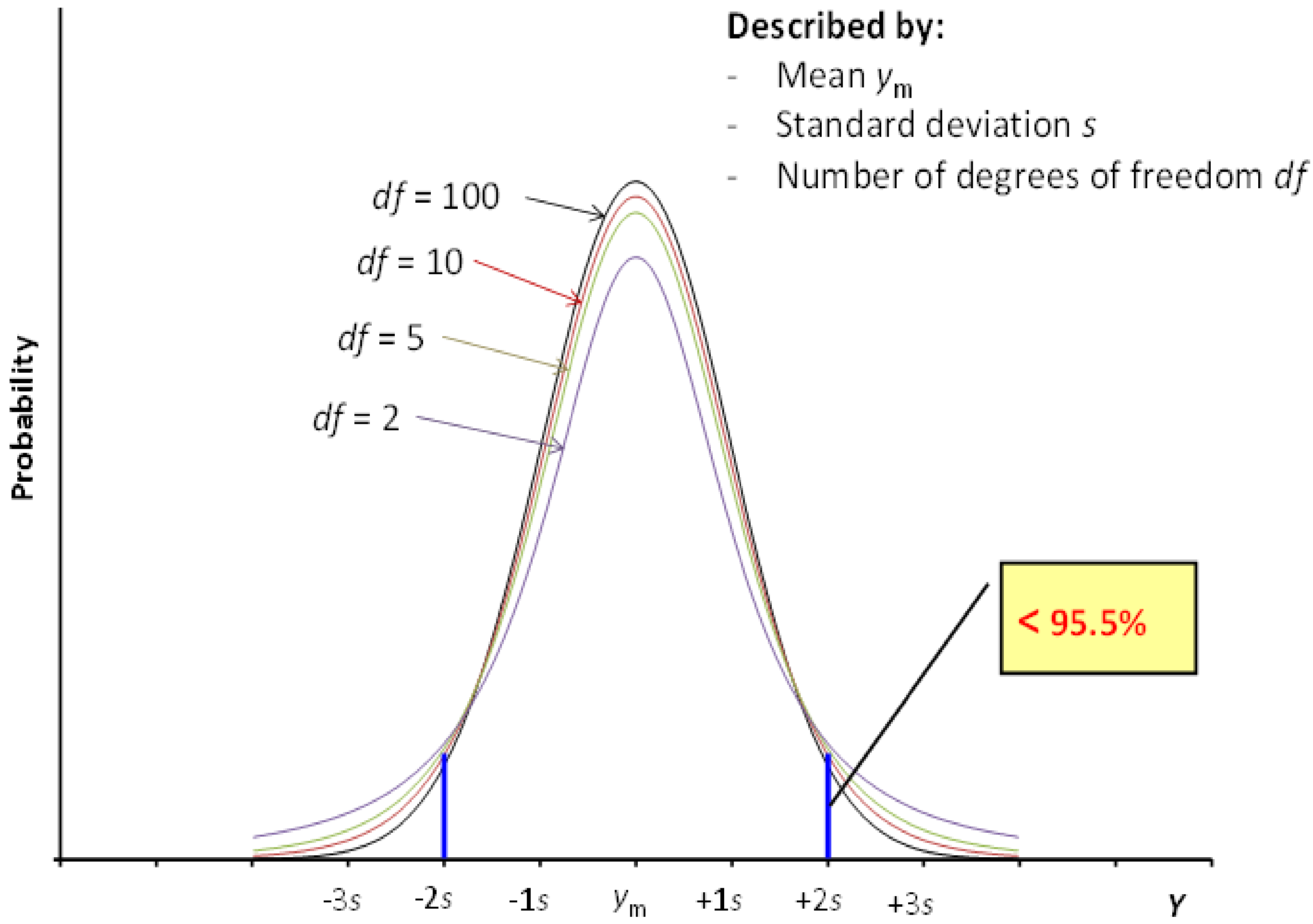
Brief summary: Like the individual values, the mean value calculated from them is also a random quantity. If the individual values are distributed according to the Normal distribution then the mean value calculated from them is distributed according to the **Student distribution** (also called as ***t*-distribution**). The properties of the *t*-distribution compared to the Normal distribution are explained. Importantly, the shape of the t-distribution curve depends on the number of **degrees of freedom**. If the number of individual values approaches infinity then the shape of the t-distribution curve approaches the Normal distribution curve.

Other distribution functions: the Student distribution

http://www.uttv.ee/naita?id=17708

https://www.youtube.com/watch?v=CWU8KM2z59I

If a measurement is repeated and the mean is calculated from the results of the single individual measurements then, just as the individual results, their mean will also be a random quantity. If the individual results are normally distributed then their mean is distributed according to the **Student distribution** (also known as the **t-distribution**). Student distribution is presented in Scheme 3.5.



Scheme 3.5. the Student distribution.

Similarly to the normal distribution the Student distribution also has mean value y_m and standard deviation s. y_m is the mean value itself, It can look strange at first sight that while the mean value ym is the only mean value we have, we immediately take it as the mean value of the distribution of mean values. However, if we had more mean values, we would anyway pool them into a single mean value (with a much higher *df!*) and use this value. and standard deviation is the standard deviation of the mean, calculated as explained in section 3.4. But differently from the normal distribution there is in addition a third characteristic – the number of **degrees of freedom** *df*. This number is equal to the number of repeated measurements minus one. So, the four Student distribution graphs in Scheme 3.5 correspond to 101, 11, 6 and 3 repeated measurements, respectively.

If *df* approaches infinity then the t-distribution approaches the normal distribution. In reality 30-50 degrees of freedom is sufficient for handling the t-distribution as the normal distribution. So, the curve with *df* = 100 in Scheme 3.5 can be regarded as the normal distribution curve.

The lower is the number of degrees of freedom the "heavier" are the tails of the Student distribution curve and the more different is the distribution from the normal distribution. This means that more probability resides in the tails of the distribution curve and less in the middle part. Importantly, the probabilities pictured in Scheme 3.2 for the ±1*s*, ±2*s* and ±3*s* ranges around the mean do not hold any more, but are all lower.

So, if a measurement result is distributed according to the t-distribution and if expanded uncertainty with predefined coverage probability is desired then instead of the usual coverage factors 2 and 3 the respective **Student coefficients** Student coefficients (i.e. *t*-distribution values) for a given set of coverage probability and number of degrees of freedom can be easily obtained from special tables in statistical handbooks (use two-sided values!), from calculation or data treatment software, such as MS Excel or Openoffice Calc or from the Internet, e.g. from the address https://en.wikipedia.org/wiki/Student_distribution should be used. Measurement result can be distributed according to the Student distribution if there is a heavily dominating The contributions of different uncertainty sources can be expressed numerically. This is explained in section 9.9 and the respective calculations are shown in 9.7. In this context the phrase „heavily dominating" means that the contribution (uncertainty index) of the respective input quantity is above 75%. A type uncertainty source that has been evaluated as a mean value from a limited number of repeated measurements. More common, however, is the situation that there is an influential but not heavily dominating A-type uncertainty source. In such a case the distribution of the measurement result is a convolution Convolution of two distribution functions in mathematical statistics means a combined of distribution function, which has a shape inbetween the two distribution functions that are convoluted. of the normal distribution and the t-distribution. What to do in this case is explained in section 9.8.

***

[1] It can look strange at first sight that while the mean value y_m is the only mean value we have, we immediately take it as the mean value of the distribution of mean values. However, if we had more mean values, we would antway pool them into a single mean value (with a much highe *df*!) and use this value.

[2] Student coefficients (i.e. *t*-distribution values) for a given set of coverage probability and number of degrees of freedom can be easily obtained from special tables in statistical handbooks (use two-sided values!), from calculation or data treatment software, such as MS Excel or Openoffice Calc or from the Internet, e.g. from the address https://en.wikipedia.org/wiki/Student_distribution

[3] The contributions of different uncertainty sources can be expressed numerically. This is explained in section 9.9 and the respective calculations are shown in 9.7. In this context the phrase „heavily dominating" means that the contribution (uncertainty index) of the respective input quantity is above 75%.

[4] Convolution of two distribution functions in mathematical statistics means a combined of distribution function, which has a shape inbetween the two distribution functions that are convoluted.

3.6. The Student distribution

4. The first uncertainty quantification

Brief summary: In this section the basic concepts and tools of the previous sections are put into practice on the example of a simple analytical chemistry operation – pipetting. The uncertainty sources identified in section 2 are now quantified (4.1), the obtained individual uncertainty estimates are (when needed) converted to standard uncertainties and are then combined into the combined standard uncertainty (4.2). The uncertainty components making up the combined standard uncertainty are compared and some conclusions are made (4.3). The combined standard uncertainty is converted into expanded uncertainty (4.4) and the result is presented (4.5). This approach is then practiced one calculation example (calibration of pipette, 4.6).

4.1. Quantifying uncertainty components

Brief summary: The same pipetting as in lecture 2 is now examined from the point of view of quantifying the uncertainty sources. All the important sources of uncertainty are quantitatively expressed as **uncertainty components** – uncertainty estimates quantitatively describing the respective uncertainty source. The uncertainty components are quantified. An example is given on converting an uncertainty estimate with (assumedly) rectangular distribution into a standard uncertainty estimate.

Introduction to quantifying measurement uncertainty

http://www.uttv.ee/naita?id=17555

https://www.youtube.com/watch?v=CaYJGxtQBzo

The main uncertainty sources are the same as explained in lecture 2 and here the uncertainty components corresponding to them are quantified.

- Uncertainty due to the **non-ideal repeatability**, which in the case of pipetting means that however carefully to fill and empty the pipette, we will nevertheless every time get a slightly different pipetted volume. Repeatability is a typical random effect. The standard uncertainty due to repeatability $u(V, \text{REP})$ can be calculated as standard deviation [1] Since pipetting for delivering a certain liquid volume is done only once and cannot be averaged (i.e. it is not possible to pipet several times and then "average" the volumes) the suitable estimate of repeatability uncertainty is the standard deviation of a single measurement, not standard deviation of the mean. of accurate volumes delivered by the pipette. [2] The accurate volumes can be measured by weighing the water delivered by the pipette and converting it into volume by using accurate density data. . According to the data for the used pipette is

$$u(V, \text{REP}) = 0.006\ \text{ml} \quad (4.1)$$

- Uncertainty due to calibration of the volumetric equipment (often called **calibration uncertainty** or uncertainty of the nominal volume). This is the uncertainty in the positions of the marks on the volumetric ware. In the case of a given pipette it is a typical systematic effect. Calibration uncertainty of the pipette used in this example is specified by the producer as ± 0.03 ml. [3] This uncertainty can be significantly reduced if the pipette is recalibrated in laboratory by weighing the delivered water. There is no information on the distribution or coverage of this uncertainty estimate. This is very common if uncertainty estimates are obtained from instrument documentation. In such case it is the safest to assume that the uncertainty estimate corresponds to rectangular distribution. In order to carry out uncertainty calculation we have to convert this uncertainty to standard uncertainty. For doing this, as explained in section 3.5, we have to divide it by square root of 3:

$$u(V, \text{CAL}) = \frac{0.03\ \text{ml}}{\sqrt{3}} = 0.017\ \text{ml} \quad (4.2)$$

- Uncertainty due to the temperature effect (often called as **temperature uncertainty**). All volumetric ware is usually calibrated at 20 °C and volumes usually refer to volumes at 20 °C [4] If volumetric glassware is calibrated in the same laboratory then a different temperature can be used. . Temperature change affects first of all the density of the liquid (the effect of expansion/contraction of glass is significantly smaller). If pipetting is done at a higher temperature than 20 °C then there amount of liquid (in terms of mass or number of molecules) pipetted is smaller than corresponds to the volume at 20 °C. Consequently, the volume of that amount of liquid at 20 °C is also smaller than if the pipetting were done at 20 °C. If the temperature of the laboratory and, importantly, temperature of the pipetted liquid, is constant during repeated measurements then the temperature effect is a systematic effect. The following video explains calculating the standard uncertainty of liquid volume due to the temperature $u(V, \text{TEMP})$:

Quantifying the uncertainty due to temperature effect in volumetric measurement

http://www.uttv.ee/naita?id= 17825

https://www.youtube.com/watch?v=CDjX8K_Vsds

$u(V, \text{TEMP})$ is dependent on the volume V of liquid delivered, [5] In the case of a volumetric pipette the nominal volume is the same as the delivered volume but in the case of a burette it is usually not. So, if from a 25 ml burette 12.63 ml of solution is delivered then the volume that has to be used for temperature effect calculation is 12.63 ml, not 25 ml. the maximum possible temperature difference from 20°C (Δt) and the thermal expansion coefficient of water γ. It is calculated as follows:

$$u(V, \mathrm{TEMP}) = \frac{V \cdot \Delta t \cdot \gamma}{\sqrt{3}} = \frac{10.00\,\mathrm{ml} \cdot 4^{\circ}\mathrm{C} \cdot 0.00021^{\circ}\mathrm{C}^{-1}}{\sqrt{3}} 0.0049\,\mathrm{ml} \quad (4.3)$$

Dividing by square root of 3 is for transforming the uncertainty estimate into standard uncertainty (assuming rectangular distribution of Δt). It is important to note that the V in eq 4.3 always refers to the actual measured volume, not the full capacity of the volumetric device. For example, if 21.2 ml of solution was measured with a 50 ml burette then the volume to use is 21.2 ml, not 50 ml.

***

[1] Since pipetting for delivering a certain liquid volume is done only once and cannot be averaged (i.e. it is not possible to pipet several times and then “average” the volumes) the suitable estimate of repeatability uncertainty is the standard deviation of a single measurement, not standard deviation of the mean.

[2] The accurate volumes can be measured by weighing the water delivered by the pipette and converting it into volume by using accurate density data.

[3] This uncertainty can be significantly reduced if the pipette is recalibrated in laboratory by weighing the delivered water.

[4] If volumetric glassware is calibrated in the same laboratory then a different temperature can be used.

[5] In the case of a volumetric pipette the nominal volume is the same as the delivered volume but in the case of a burette it is usually not. So, if from a 25 ml burette 12.63 ml of solution is delivered then the volume that has to be used for temperature effect calculation is 12.63 ml, not 25 ml.

4.2. Calculating the combined standard uncertainty

The uncertainty components that were quantified in the previous lecture are now combined into the **combined standard uncertainty** (u_c) – standard uncertainty that takes into account contributions from all important uncertainty sources by combining the respective uncertainty components. The concept of **indirect measurement** – whereby the value of the **output quantity** (measurement result) is found by some function (model) from several **input quantities** – is introduced and explained. The majority of chemical measurements are indirect measurements. The general case of combining the uncertainty components into combined standard uncertainty as well as several specific cases are presented and explained.

The first video lecture explains in a simple way how the uncertainty components are combined in the particular example of pipetting. The second video lecture presents the general overview of combining the uncertainty components.

Combining the uncertainty components into the combined standard uncertainty in the case of pipetting

http://www.uttv.ee/naita?id=17556

https://www.youtube.com/watch?v=S5v58VQ4zSg

In all cases where combined standard uncertainty is calculated from uncertainty components all the uncertainty components have to be converted to standard uncertainties.

In the example of pipetting the combined standard uncertainty is calculated from the uncertainty components found in the previous section as follows:

$$u_c(V) = \sqrt{u(V,\mathrm{REP})^2 + u(V,\mathrm{CAL})^2 + u(V,\mathrm{TEMP})^2} = \sqrt{0.006^2 + 0.017^2 + 0.005^2} = 0.019\ \mathrm{ml} \quad (4.4)$$

This is the typical way of calculating combined standard uncertainty if all the uncertainty components refer to the same quantity and are expressed in the same units. It is often used in the case of **direct measurements** – measurements whereby the measurement instrument (pipette in this case) gives immediately the value of the result, without further calculations needed.

Combining the uncertainty components into the combined standard uncertainty: simple cases and the general case

http://www.uttv.ee/naita?id=17826

https://www.youtube.com/watch?v=FJ4hn9LgGmw

An indirect measurement is one where the **output quantity** (result) is found ba calculation (using a **model equation**) from several input quantities. A typical example is titration. In case of titration with 1:1 mole ratio the analyte concentration in the sample solution C_S (the output quantity) is expressed by the input quantities – volume of sample solution taken for titration (V_S), titrant concentration (C_T) and titrant volume consumed for titration (V_T) – as follows:

$$C_S = \frac{V_T \cdot C_T}{V_S} \quad (4.5)$$

In the general case if the output quantity Y is found from input quantities X_1, X_2, … X_n according to some function F as follows

$$Y = F(X_1, X_2, \dots, X_n) \quad (4.6)$$

then the combined standard uncertainty of the output quantity $u_c(y)$ can be expressed via the standard uncertainties of the input quantities $u(x_i)$ as follows:

$$u_c(y) = \sqrt{\left[\frac{\partial Y}{\partial X_1} u(x_1)\right]^2 + \left[\frac{\partial Y}{\partial X_2} u(x_2)\right]^2 + \dots\dots + \left[\frac{\partial Y}{\partial X_n} u(x_n)\right]^2} \quad (4.7)$$

The terms $\left[\frac{\partial Y}{\partial X_i} u(x_i)\right]$ are the uncertainty components. The terms $\frac{\partial Y}{\partial X_i}$ are partial derivatives. At first sight the eq 4.7 may seem very complex but it is in fact not too difficult to use – the uncertainty components can be calculated numerically using the Kragten's spreadsheet method (as is demonstrated in section 9.7).

In specific cases simpler equations hold. If the output quantity is expressed via the input quantities as follows

$$Y = X_1 - X_2 + \dots + X_n \quad (4.8)$$

then $$u_c(y) = \sqrt{u(x_1)^2 + u(x_2)^2 + \dots + u(x_n)^2} \quad (4.9)$$

Importantly, irrespective of whether the input quantities are added or subtracted, the squared standard uncertainties under the square root are always added.

This way of combining uncertainty components is in principle the same as used above for the case of pipetting.

If the measurement model is

$$Y = \frac{X_1 \cdot X_2}{X_3 \cdot X_4} \quad (4.10)$$

then $$\frac{u_c(y)}{y} = \sqrt{\left(\frac{u(x_1)}{x_1}\right)^2 + \left(\frac{u(x_2)}{x_2}\right)^2 + \left(\frac{u(x_3)}{x_3}\right)^2 + \left(\frac{u(x_4)}{x_4}\right)^2} \quad (4.11)$$

As can be seen, here it is the relative standard uncertainties that are combined and the squared summing gives us the relative combined standard uncertainty of the output quantity. The absolute combined standard uncertainty of the output quantity is found as follows:

$$u_c(y) = y \cdot \sqrt{\left(\frac{u(x_1)}{x_1}\right)^2 + \left(\frac{u(x_2)}{x_2}\right)^2 + \left(\frac{u(x_3)}{x_3}\right)^2 + \left(\frac{u(x_4)}{x_4}\right)^2} \quad (4.12)$$

The file used in second video can be downloaded from here.

Combined Standard Uncertainty

Ivo Leito
University of Tartu
Institute of Chemistry

2013

1

Combined Standard Uncertainty

- When estimating the standard uncertainty of an **output quantity** then the standard uncertainties of all **input quantities** are taken into account
- The standard uncertainty of the output quantity obtained in this way is called **combined standard uncertainty**
- Is denoted by $u_c(y)$

2

Indirect Measurement

- If the result is obtained as an **output quantity** by calculations from one or several **input quantities** then we speak about **indirect measurement**
- Example: Titration
  - Titrant concenration
  - Volume of sample solution taken for titration
  - Volume of titrant consumed for titartion

3

Measurement model

- Expression that enables calculating the output quantity value (result value) from the input quantity values
- Titration with 1:1 stoichiometry:

Titrant volume

Titrant concentration

$$C_S = \frac{V_T \cdot C_T}{V_S}$$

Analyte concentration in sample solution

Volume of sample solution taken for titration

4

Finding u_c

- If the model is:

$$Y = X_1 - X_2 + ... + X_n$$

- then:

$$u_c(y) = \sqrt{u(x_1)^2 + u(x_2)^2 + ... + u(x_n)^2}$$

- All uncertainties have to be converted to **standard uncertainties** before calculations!

5

Finding u_c

- If the model is:

$$Y = \frac{X_1 \cdot X_2}{X_3 \cdot X_4}$$

- Then:

$$u_c(y) = y \cdot \sqrt{\left(\frac{u(x_1)}{x_1}\right)^2 + \left(\frac{u(x_2)}{x_2}\right)^2 + \left(\frac{u(x_3)}{x_3}\right)^2 + \left(\frac{u(x_4)}{x_4}\right)^2}$$

$$\frac{u_c(y)}{y} = \sqrt{\left(\frac{u(x_1)}{x_1}\right)^2 + \left(\frac{u(x_2)}{x_2}\right)^2 + \left(\frac{u(x_3)}{x_3}\right)^2 + \left(\frac{u(x_4)}{x_4}\right)^2}$$

6

Finding u_c

- The general case:

$$Y = F(X_1, X_2, \dots, X_n)$$

- then:

$$u_C(y) = \sqrt{\left[\frac{\partial Y}{\partial X_1} u(x_1)\right]^2 + \left[\frac{\partial Y}{\partial X_2} u(x_2)\right]^2 + \dots + \left[\frac{\partial Y}{\partial X_n} u(x_n)\right]^2}$$

Uncertainty component

7

Finding u_c

Important:

The equations on the previous slides are applicable only in the case of non-correlating input quantities!

8

Brief summary: The uncertainty components of the previous lecture are compared. The property of squared summing – suppressing the less influential uncertainty components – is explained. The meaning of the obtained combined standard uncertainty estimate is explained in terms of probability (the probability of the true value of pipetted volume being within the calculated uncertainty range).

Comparing the uncertainty components

http://www.uttv.ee/naita?id=17578

https://www.youtube.com/watch?v=Xnq2-7nq_bg

We can see the uncertainty component with the largest magnitude is the calibration uncertainty $u(V, \mathrm{CAL})$ = 0.017 ml. The combined standard uncertainty is in fact quite similar to it: $uc(V)$ = 0.019 ml. If the summing were made not by the squaring and square root approach but by simple arithmetic sum then the value would be 0.028. This is a good illustration of the property of the squared summing: the smaller uncertainty components are suppressed by larger uncertainty components.

The idea of the squared summing of the components is that the different effects causing uncertainty influence the result in different directions (thus partially canceling) and their magnitudes are not necessarily equal to the values of the uncertainty estimates but can also be smaller (see section 1).

Looking at the uncertainty contributions is very useful if one wants to reduce the uncertainty. In order to reduce the uncertainty of a particular measurement it is always necessary to focus on decreasing the uncertainties caused by the largest components. So, in this case it is not very useful to buy a more expensive air conditioner for the room because the resulting uncertainty improvement will be small. It will also not be possible to improve the uncertainty markedly by reducing the repeatability component. Clearly, whatever is done with these two components the combined standard uncertainty $u_c(V)$ (eq 4.4) cannot decrease from 0.019 ml to lower than 0.017 ml, which is not a significant decrease. Thus, if more accurate pipetting is needed, then the way to go is to calibrate the pipette in the laboratory. This way it is realistic to achieve threefold lower calibration uncertainty, which leads to two times lower combined uncertainty of the pipetted volume. See section 4.6 for an example.

4.4. Expanded uncertainty

Brief summary: The probability of roughly 68% that is provided by the standard uncertainty is often too low for the users of measurement uncertainty. Therefore measurement uncertainty is presented to customers mostly as **expanded uncertainty**, U. Expanded uncertainty is calculated from the standard uncertainty by multiplying it with a **coverage factor**, k.

In the case of the pipetting example the k = 2 expanded uncertainty is found as follows:

$$U(V) = u\mathrm{c}(V) \cdot k = 0.019\ \mathrm{ml} \cdot 2 = 0.038\ \mathrm{ml} \quad (4.13)$$

Expanded uncertainty at k = 2 level is the most common way of expressing uncertainty of measurement and analysis results.

Calculating the expanded uncertainty

http://www.uttv.ee/naita?id=17579

https://www.youtube.com/watch?v=tBwkuL0ap14

4.5. Presenting measurement results

Brief summary: The pipetting result – the value and expanded uncertainty – is presented. It is stressed that it is important to clearly say, what was measured. The correct presentation of measurement result includes value, uncertainty and information about the probability of the uncertainty. It is explained that in simplified terms we can assume that k = 2 corresponds to roughly 95% of coverage probability. It is explained how to decide how many decimals to give when presenting a measurement result and the uncertainty.

The correct presentation of the measurement result in this case would look as follows:

The volume of the pipetted liquid is $V = (10.000 \pm 0.038)\ \text{ml},\ k = 2,\ \text{norm.}$ (4.14)

The parentheses (brackets) mean that the unit "ml" is valid both for the value and the uncertainty. "norm." means that the output quantity is expected to be approximately normally distributed. This, together with coverage factor value 2, means that the presented uncertainty is expected to corresponds to approximately 95% coverage probability (see section 3.1 for details).

When can we assume that the output quantity is normally distributed? That is, when can we write "norm." besides the coverage factor? Rigorous answer to this question is not straightforward, but a simple rule of thumb is that when there are at least three main uncertainty sources of comparable influence (i.e. the smallest and largest of the uncertainty components differ by ca 3 times or less) then we can assume that the distribution function of the output is sufficiently similar to the normal distribution. In this particular case there is a dominating quantity with assumedly rectangular distribution, which leads to a distribution function with very „weak tails" (meaning: this in fact not exactly a normal distribution). So, 95% coverage probability is achieved already by expanded uncertainty of 0.0034 ml (as evidenced by Monte Carlo simulations). Thus, the presented uncertainty of 0.0038 ml is a conservative estimate (which is not bad as explained in section 3.5).

Section 9.8 presents a more sophisticated approach of calculating expanded uncertainty that corresponds to a concrete coverage probability.

Presenting of measurement result

http://www.uttv.ee/naita?id=17576

https://www.youtube.com/watch?v=x9l1jIVqT7k

***

[1] In this particular case there is a dominating quantity with assumedly rectangular distribution, which leads to a distribution function with very „weak tails" (meaning: this in fact not exactly a normal distribution). So, 95% coverage probability is achieved already by expanded uncertainty of 0.0034 ml (as evidenced by Monte Carlo simulations). Thus, the presented uncertainty of 0.0038 ml is a conservative estimate (which is not bad as explained in section 3.5).

4.6. Practical example

This is an example of calculating the volume and its uncertainty of liquid delivered from a self-calibrated volumetric pipette

The uncertainty of the pipetted volume u(*V*) has three main uncertainty components: uncertainty due to repeatability, u(*V*,rep); uncertainty due to pipette calibration, u(*V*,cal) and uncertainty due to the temperature difference from 20 °C, u(*V*,temp).

The estimate of the probable maximum difference of the pipette volume from the nominal volume, expressed as ±x is often used as the estimate of calibration uncertainty of the pipette (as was done in section 4.1). It is usually given by the manufacturer without any additional information about its coverage probability or distribution function. In such a case it is the safest to assume that rectangular distribution holds and to convert the uncertainty estimate to standard uncertainty by dividing it with square root of three.

Usually in the case of high-accuracy work the pipette is calibrated in the laboratory in order to obtain lower calibration uncertainty. As was seen in sections 4.2 and 4.3, if the uncertainty due to factory calibration is used, then this calibration uncertainty component is the most influential one. So, reducing it would also reduce the overall uncertainty. It is very important, that calibration and pipetting are performed under the same conditions and preferably by the same person. From the calibration data we can obtain two important pieces of information: (1) the correction term for the pipette volume $V_{correction}$ with uncertainty *u* (*V*,cal) and (2) repeatability of pipetting *u* (*V*,rep). The example presented here explains this.

For calibration of a pipette water is repeatedly pipetted (at controlled temperature in order to know its density), the masses of the pipetted amounts of water are measured and the pipetted volumes of water are calculated (using density of water at the calibration temperature). Here are the calibration data of a 10 ml pipette:

The standard deviation is calculated according to the equation:

$$s = \sqrt{\frac{\sum_i^n (V_i - V_m)^2}{n-1}} = 0.0057\ \text{ml}$$

Uncertainty due to repeatability of pipetting *u* (*V*, REP) is equal to this standard deviation 0.0057 ml. Pipetting is often used in titration analysis. If the solution that is titrated is pipetted then this repeatability contribution is already accounted for in the repeatability of the titration results and is not separately taken into account in the uncertainty of pipette volume.

Uncertainty of the calibration (in fact the uncertainty of the correction) has to be always taken into account and this is expressed (when calibration is done) as the standard deviation of the mean:

$$u(V,\text{CAL}) = s(V_m) = \frac{s(V)}{\sqrt{n}} = \sqrt{\frac{\sum_i^n (V_i - V_m)^2}{n(n-1)}} = 0.0018\ \text{ml}$$

In this example the correction is -0.0080 ml and its standard uncertainty is 0.0018 ml.

When there is a possibility that pipetting is performed at a different temperature from the

temperature of water	23.8	°C
density of pure water	0.9973	g/cm³

	m(water) (g)	V(water) (ml)
1	9.96037	9.98734
2	9.96454	9.99152
3	9.9632	9.99017
4	9.97152	9.99852
5	9.9683	9.99529
6	9.97413	10.00113
7	9.96806	9.99505
8	9.96165	9.98862
9	9.96397	9.99095
10	9.9544	9.98135
average	9.965014	**9.9920**
Correction		**-0.0080**
Experimental standard deviation, $s(V)$		**0.0057**
Relative standard deviation, $\frac{s(V)}{V}$		**0.058%**
Standard deviation of the mean, $s(V_m)$		**0.0018**

calibration (and this possibility exists almost always), then an additional uncertainty source due to temperature change is introduced and it has to be taken into account.

In this case we assume that pipette's using temperature does not differ from the calibration temperature by more than 4 °C (Δt = ±4°C, assuming rectangular distribution). Water's density depends on temperature, therefore we have to consider also the thermal expansion coefficient of water, which is $\gamma_W = 2.1 \cdot 10^{-4}$ 1/°C. So,

$$u(V_{\mathrm{temp}}) = \frac{V_p \cdot \Theta \cdot \gamma_w}{\sqrt{3}} = \frac{9.992\,\mathrm{ml} \cdot 4\,°\mathrm{C} \cdot 2.1 \cdot 10^{-4}\,\frac{1}{°C}}{\sqrt{3}} = 0.0048\,\mathrm{ml}$$

Now, when we will perform a single pipetting, the volume is 9.992 ml and its combined standard uncertainty is

$$u(V) = \sqrt{0.0057^2 + 0.0018^2 + 0.0048^2} = 0.0077\,\mathrm{ml}$$

The k = 2 expanded uncertainty Later in this course (section 9.8) we will see, how to rigorously find, whether we can say that the k = 2 expanded uncertainty in a particular case corresponds to 95% (this depends on the so-called effective number of degrees of freedom). And if not then what k should be used to achieve approximately 95% coverage probability. In the case of this example the effective number of degrees of freedom is 26 and the respective coverage factor (actually the Student coefficient) with the probability of 95% is 2.06, which is only very slightly different from 2 (the expanded uncertainty would increase from 0.015 ml to 0.016 ml). of the pipetted volume can be found as follows:

$$U(V) = u_c(V) \cdot k = 0.0077 \cdot 2 = 0.0154\ \mathrm{ml}$$

As explained in section 4.5 if the first significant digit of the uncertainty is 1... 4 then uncertainty should be presented with two significant digits. Thus we can write the result:

The volume of the pipetted liquid is:

$$V = (9.992 \pm 0.015)\ \mathrm{ml},\ k = 2,\ \mathrm{norm.} \qquad (4.15)$$

It is interesting to compare now this expanded uncertainty with the expanded uncertainty obtained in section 4.5 (eq 4.14). We see that when the pipette is calibrated in our laboratory then the uncertainty of the volume is more than two times lower. We also see that the uncertainty component due to pipette calibration, which back then was the largest uncertainty component, is now the smallest.

***

[1] Later in this course (section 9.8) we will see, how to rigorously find, whether we can say that the k = 2 expanded uncertainty in a particular case corresponds to 95% (this depends on the so-called effective number of degrees of freedom). And if not then what k should be used to achieve approximately 95% coverage probability.
In the case of this example the effective number of degrees of freedom is 26 and the respective coverage factor (actually the Student coefficient) with the probability of 95% is 2.06, which is only very slightly different from 2 (the expanded uncertainty would increase from 0.015 ml to 0.016 ml).

5. Principles of measurement uncertainty estimation

Brief summary: The main principles of measurement uncertainty estimation – the so-called GUM principles – are presented on the example of determination of pesticides in oranges. These principles have been laid down in the ISO GUM ISO GUM originally refers to the Guide To The Expression of Uncertainty in Measurement, ISO, Geneva, Switzerland, 1993 (Reprinted 1995). In 2008 this document was revised and reissued as ISO JCGM 100:2008 Evaluation of measurement data - Guide to the expression of uncertainty in measurement. The latter document is available on-line from http://www.bipm.org/en/publications/guides/gum.html and they are now universally accepted as being the common foundation of all the different uncertainty estimation approaches. These principles are the following:

1. The basis of any measurement (thus obviously also measurement uncertainty evaluation) is the definition of the measurand;
2. The used measurement procedure has to correspond to the measurand definition;
3. All relevant sources of uncertainty have to be carefully considered and those that are important have to be taken into account;
4. The random and systematic effects are treated the same way when estimating measurement uncertainty – both are evaluated as standard uncertainties, which thereafter are combined into the combined standard uncertainty.

***

[1] ISO GUM originally refers to the *Guide To The Expression of Uncertainty in Measurement*, ISO, Geneva, Switzerland, 1993 (Reprinted 1995). In 2008 this document was revised and reissued as ISO JCGM 100:2008 *Evaluation of measurement data - Guide to the expression of uncertainty in measurement*. The latter document is available on-line from http://www.bipm.org/en/publications/guides/gum.html

5.1. Measurand definition

Brief summary: The first principle of measurement uncertainty is: the measurand must be correctly and unambiguously defined. The importance of measurand definition is explained on the example of pesticide determination in oranges.

Defining the measurand

http://www.uttv.ee/naita?id=17585

https://www.youtube.com/watch?v=Bf_0EssHQ60

Defining the measurand in the case of pesticide determination in oranges is not trivial. On one hand it is important to define whether the result is applied to a single orange or few oranges that were taken as the **sample** or whether it is applied to the **whole lot** of oranges (the whole analysis object, also called sampling target). On the other hand, oranges are not homogenous. Pesticides are applied on orange surface, not inside. At the same time pesticide can diffuse from the orange peel to the inside. So, a number of different possibilities exist: whole orange, whole peel, outside part of the peel, only orange flesh.

Combining together we get 8 possibilities, in what exactly we can determine pesticides. Measuring pesticide content according to any of these will lead to different and mutually non-comparable results.

In addition (not explained in the video) instead of defining the measurand via the total analyte content in the sample (or part of the sample) it is often more practical to look at some part of the analyte only.

A good example is phosphorus determination in soil. Although it is possible to determine the **total** phosphorus content in soil it is in fact more interesting to determine only the part that is available to plants – the **bioavailable** phosphorus – because it is this part of the total phosphorus content that contributes to the fertility of the soil and is therefore of interest in agriculture.

Total phosphorus content and bioavailable phosphorus content are different measurands and their values for the same soil differ strongly. This has important implications for the measurement procedure. This is explained in section 5.2.

5.2. Measurement procedure

Brief summary: The main steps of a measurement/analysis procedure are presented on the example of pesticide measurement are presented: Sample preparation (in this case: homogenization, extraction(s), extract purification), instrument calibration, actual analysis. It is stressed that the measurement procedure must correspond to measurand definition.

Measurement procedure

http://www.uttv.ee/naita?id=17586

https://www.youtube.com/watch?v=BKlB_iB4wp4

This scheme of a chemical analysis procedure is very general. In specific cases there can be deviations from this scheme (more steps or less steps). In particular, sampling is not introduced here as a step of chemical analysis procedure. This holds if samples are brought to the laboratory for analysis and the laboratory itself does not do sampling (see section 5.3 for more details).

It is worth stressing the importance of **sample preparation as** a step in analytical procedure. The majority of analytical procedures need that the sample is converted into a solution which contains as large as possible share of the analyte from the sample (ideally all of it) and as little as possible of the other components of the sample **matrix**. In analytical chemistry sample matrix is the term for describing jointly all sample components except the analyte(s). The matrix components often act as **interfering** compounds, which can artificially increase or decrease the result. Therefore it is important to minimize their content in the solution obtained from the sample. If the interfering compounds cannot be fully eliminated and the interference cannot be corrected (which is quite usual in chemical analysis) then their effect has to be taken into account in measurement uncertainty estimation.

Sample preparation is often the most work-intensive part of chemical analysis and in most cases it is also the part, which has the largest uncertainty contribution. Sample preparation usually involves either of the two approaches:

1. Essentially destroying the sample matrix so that a solution containing the analyte(s) and few matrix components is obtained. This is often done by digestion with acids or fusing with alkalies or salts. This approach is suitable for determining elements.
2. Separating the analyte(s) from the sample matrix so that a solution containing the analyte(s) is obtained where the amount of matrix components is as small as possible. This is usually done by a set of extractions. This approach is suitable for organic analytes.

Obviously the choice of sample preparation procedure depends on whether the measurand corresponds to the total analyte content in the sample or some part of it, e.g. the bioavailable analyte content (see the text in section 5.1). In the case of e.g. total phosphorus content determination in soil the analyst can make the choice of the sample preparation procedure. All procedures that lead to determination of the total phosphorus content (often involving complete destruction of the matrix) are suitable. In the case of determining of e.g. bioavailable phosphorus content in soil the sample preparation procedure must mimic the way the plants get phosphorus from soil. So, sample preparation involves leaching at predefined conditions. Such sample preparation procedures are often standardized and whenever the results are meant to be mutually comparable they must be obtained with the same procedure. Thus, in this latter case the sample preparation procedure becomes part of the measurand definition.

5.3. Sources of measurement uncertainty

Brief summary: The overview of possible uncertainty sources, on the example pesticide analysis, is presented. Although the uncertainty sources are presented on the example of pesticide analysis, the same uncertainty sources hold for the majority of other analytical methods. Most of the uncertainty sources are linked to specific steps in the analysis procedure. It is stressed that sample preparation is usually the biggest contributor to measurement uncertainty. When performing chemical analysis then every care should be taken to minimize (preferably eliminate) the influence of the uncertainty sources, as far as possible. And what cannot be eliminated, has to be taken into account. It is not necessary to quantify every uncertainty source individually. Instead, it is often more practical to quantify several uncertainty sources jointly.

Measurement uncertainty sources

http://www.uttv.ee/naita?id=17587

https://www.youtube.com/watch?v=4y2cjJ8Jpsg

5.4. Treatment of random and systematic effects

Brief summary: Although within a measurement series random and systematic effects influence measurement results differently, they are mathematically taken into account the same way – as uncertainty components presented as standard uncertainties.

Treatment of random and systematic effects

http://www.uttv.ee/naita?id=17712

https://www.youtube.com/watch?v=hdh5xVVZTbg

In the case of pipetting (demonstrated and explained in sections 2 and 4.1) there are three main sources of uncertainty: repeatability, calibration uncertainty of the pipette and the temperature effect. These effects influence pipetting in different ways.

1. Repeatability is a typical random effect. Every pipetting operation is influenced by random effects that altogether cause the differences between the volumes that are pipetted under identical conditions;
2. The uncertainty due to calibration of the pipette is a typical systematic effect: If instead of 10.00 ml the mark on the pipette is, say, at 10.01 ml then the pipetted volume will be systematically too high. This means that although individual pipetting results can be lower than 10.01 ml (and in fact even below 10.00 ml), the average volume will be higher than 10.00 ml: approximately 10.01 ml.
3. The temperature effect can be, depending on the situation, either systematic or random effect or (very commonly) mixture of the two. Which way it is depends on the stability of the temperature during repetitions (which is influenced by the overall duration of the experiment).

Although the three uncertainty sources influence pipetting results in different ways they are all taken into account the same way – via uncertainty contributions expressed as standard uncertainties.

In principle, it is possible to investigate the systematic effects, determine their magnitudes and take them into account by correcting the results. When this is practical, this should be done. If this is not done then the results will be biased, i.e. will be systematically shifted from the true value.

An example where systematic effect can be determined and correction introduced with reasonable effort is calibration of pipette, explained in the example in section 4. Two cases were examined: without correcting and with correcting:

1. In subsection 4.1 the calibration uncertainty of ± 0.03 ml as specified by the producer is used. This corresponds to the situation that there is possibly a systematic effect – the *possible* The word possible means here that in fact there may be no systematic effect – the actual pipette volume can be 10.00 ml. We simply do not know. difference of the true pipette volume from its nominal volume, but it is not closely investigated or corrected and the uncertainty ± 0.03 ml is assigned to it, which with very high probability covers this effect. As a result the standard uncertainty of calibration was u (V, CAL) = 0.017 ml (rectangular distribution is assumed).
2. In subsection 4.6 it was explained how to determine the actual volume of pipette by calibration. It was found that the pipette volume was 10.006 ml. The calibration that was carried out in the laboratory still has uncertainty, but this uncertainty now is due to the repeatability during calibration (i.e. random effects) and is by almost 10 times smaller: u (V, CAL) = 0.0018 ml.

Thus, when correcting for systematic effects can be done with reasonable effort then it can lead to significant decrease of measurement uncertainty. However, in many cases accurate determination of a systematic effect (accurate determination of bias) can involve a very large effort and because of this can be impractical. It can also happen that the uncertainty of correction is not much smaller than the uncertainty due to possible bias. In fact, with reasonable (i.e. not very large) effort the outcome of bias determination often is that there may be a systematic effect and may not be. In such cases correction cannot be done and the uncertainty due to the effect has to cover the *possible* systematic effect (*possible* bias).

***

[1] The word *possible* means here that in fact there may be no systematic effect – the actual pipette volume can be 10.00 ml. We simply do not know.

6. Random and systematic effects revisited

Brief summary: This section explains that whether an effect will influence the measurement result as a random or as a systematic effect depends on the conditions. Effects that are random in short term can become systematic in long term. This is the reason why repeatability is by its value smaller than within-lab reproducibility and the latter is in turn smaller than the combined standard uncertainty. This section also explains that the A and B type uncertainty estimates do not correspond one to one to the random and systematic effects.

How a within-day systematic effect can become a long-term random effect?

http://www.uttv.ee/naita?id=17713

https://www.youtube.com/watch?v=qObLSS7mfDo

An effect that within a short time period (e.g. within a day) is systematic can over a longer time period be random. Examples:

1. If a number of pipetting operations are done within a day using the same pipette then the difference of the actual volume of the pipette from its nominal volume (i.e. calibration uncertainty) will be a systematic effect. If pipetting is done on different days and the same pipette is used then it is also a systematic effect. However, if pipetting is done on different days and different pipettes are used then this effect will change into a random effect.
2. An instrument is calibrated daily with calibration solutions made from the same stock solution, which is remade every month. In this case the difference of the actual stock solution concentration and its nominal concentration is a systematic effect within a day and also within few weeks. But over a longer time period, say, half a year, It cannot be strictly defined, how long is „long-term". An approximate guidance could be: one year is good, „several months" (at least 4-5) is minimum. Of course it also depends on the procedure. this effect becomes random, since a number of different sock solutions will have been in use during that time period.

Conclusions:

1. An effect, which is systematic in short term can be random in long term;
2. The longer is the time frame the more effects can change from systematic into random.

As explained in a past lecture if the measurement of the same or identical sample is repeated under identical conditions (usually within the same day) using the same procedure then the standard deviation of the obtained results is called **repeatability standard deviation** and denoted as $\boldsymbol{s_r}$. If the measurement of the same or identical sample is repeated using the same procedure but under changed conditions whereby the changes are those that take place in the laboratory under normal work practices then the standard deviation of the results is called **within-lab reproducibility** or **intermediate precision** and it is denoted as $\boldsymbol{s_{RW}}$. The terms „within-lab reproducibility" and „intermediate precision" are synonyms. The VIM(1) prefers intermediate precision. The Nordtest handbook(5) uses within-lab reproducibility (or reproducibility within laboratory). In order to stress the importance of the „long-term", in this course we often refer to sRW as the within-lab long-term reproducibility

The conclusions expressed above are the reason why s_r is smaller than s_{RW}. Simply, some effects that within day are systematic and are not accounted for by s_r become random over a longer time and s_{RW} takes them into account. The combined standard uncertainty u_c is in turn larger that the intermediate precision, because it has to take into account all significant effects that influence the result, including those that remain systematic also in the long term. This important relation between these three quantities is visualized in Scheme 6.1.

Repeatability < Within-lab reproducibility < Combined uncertainty

s_r < s_{RW} < u_c

Scheme 6.1. Relations between repeatability, within-lab reproducibility and combined uncertainty.

The random and systematic effects cannot be considered to be in one-to-one relation with type A and B uncertainty estimation. These are categorically different things. The effects refer to the intrinsic causal relationships, while type A and B uncertainty estimation refers rather to approaches used for quantifying uncertainty. Table 6 illustrates this further.

Table 6.1. Interrelations between random and systematic effects and A and B types of uncertainty estimates.

Effect	Type A estimation	Type B estimation
Random	The usual way of estimating uncertainties caused by random effects	Type B estimation of the uncertainty caused by random effects is possible if no repeated measurements are carried out and the data on the magnitude of the effect is instead available from different sources.
Systematic	This is only possible if the effect will change into a random effect in the long term	The usual way of estimating uncertainties caused by the systematic effects

When estimating the uncertainty contributions due to random effects, then it is important that a number of repeated measurements are carried out. On the other hand, if, e.g. repeatability of some analytical procedure is estimated then each repetition has to cover all steps in the procedure, including sample preparation. For this reason making extensive repetitions is very work-intensive. In this situation the concept of **pooled standard deviation** becomes very useful. Its essence is pooling standard deviations obtained from a limited number of measurements. The following video explains this:

Pooled standard deviation

http://www.uttv.ee/naita?id=18228

https://www.youtube.com/watch?v=xsltS41PZW0

Depending on how the experiments are planned, the pooled standard deviation can be used for calculating of either repeatability s_r or within-lab reproducibility s_{RW}. The experimental plan and calculations when finding repeatability s_r are explained in the following video:

Pooled standard deviation in practice: estimating repeatability

http://www.uttv.ee/naita?id=18232

https://www.youtube.com/watch?v=DM_zf85PYic

The experimental plan and calculations when finding within-lab reproducibility s_{RW} are explained in the following video:

Pooled standard deviation in practice: estimating within-lab long-term reproducibility

http://www.uttv.ee/naita?id=18234

https://www.youtube.com/watch?v=nPJY8HfPxNs

***

[1] It cannot be strictly defined, how long is „long-term". An approximate guidance could be: one year is good, „several months" (at least 4-5) is minimum. Of course it also depends on the procedure.

[2] The terms „within-lab reproducibility" and „intermediate precision" are synonyms. The VIM(1) prefers intermediate precision. The Nordtest handbook(5) uses within-lab reproducibility (or reproducibility within laboratory). In order to stress the importance of the „long-term", in this course we often refer to s_{RW} as the within-lab long-term reproducibility.

***

The slides of the presentation and the calculation files – with initial data only, as well, as with calculations performed – are available from here:

- pooled_standard_deviation.pdf 18 KB
- pooled_standard_deviation_repeatability_initial.xlsx 16 KB
- pooled_standard_deviation_repeatability_solved.xlsx 16 KB
- pooled_standard_deviation_reproducibility_initial.xlsx 15 KB
- pooled_standard_deviation_reproducibility_solved.xlsx 16 KB

Pooled Standard Deviation

Ivo Leito
University of Tartu
Institute of Chemistry
2013

1

Pooled Standard Deviation

- If it is impossible to make many repeated measurements with the same sample
- Then precision can be estimated during longer time in the form of **pooled standard deviation**
- Pooled standard deviation can be used to calculate:
  - Repeatability
  - Within-lab reproducibility

2

Pooled Standard Deviation

- General formula for the case when experiment is done with different samples, each measured repeatedly:

$$s_{\text{pooled}} = \sqrt{\frac{(n_1 - 1)s_1^{\,2} + (n_2 - 1)s_2^{\,2} + \ldots + (n_k - 1)s_k^{\,2}}{n_1 + n_2 + \ldots + n_k - k}}$$

- Symbols:
  - k number of samples
  - s_1, s_2, etc are within sample standard deviations
  - n_1, n_2, etc are numbers of measurements made for different samples

3

Pooled Standard Deviation

- If the number of measurements made with each sample was the same:

$$s_{\text{pooled}} = \sqrt{\frac{s_1^{\,2} + s_2^{\,2} + \ldots + s_k^{\,2}}{k}}$$

- Symbols:
  - k number of samples
  - s_1, s_2, etc are within sample standard deviations
  - n_1, n_2, etc are numbers of repeated measurements with every sample

4

Pooled standard deviation in practice

- How to set up experiment for **repeatability** s_r evaluation using s_{pooled}?
- How to set up experiment for **within-lab long-term reproducibility** s_{RW} evaluation using s_{pooled}?

5

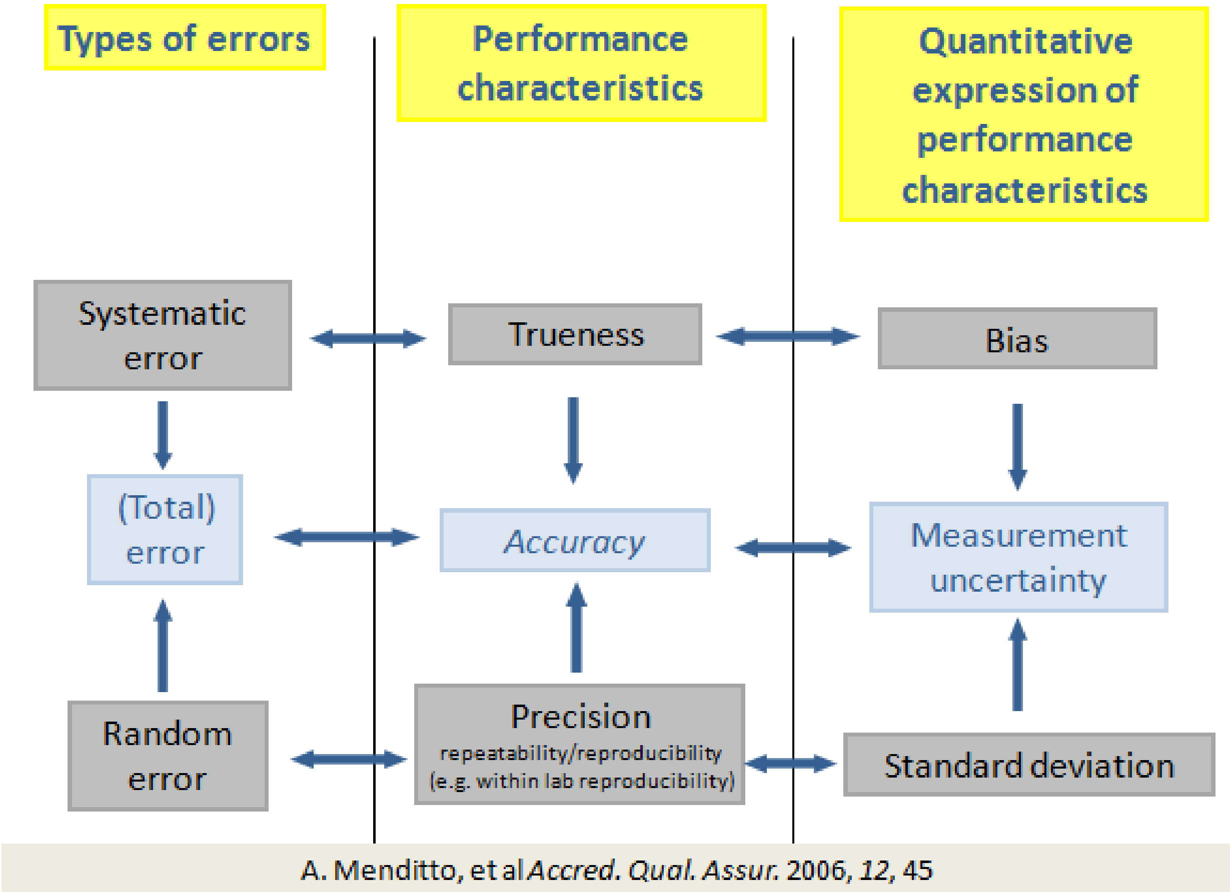
7. Precision, trueness, accuracy

Brief summary: Interrelations between the different error types (**random**, **systematic**, **total**), their corresponding performance characteristics (**precision**, **trueness**, **accuracy**) and the parameters for quantitatively expressing these performance parameters (**standard deviation**, **bias**, **measurement uncertainty**) are explained in this section.

Interrelation between the concepts of precision, trueness, accuracy and measurement uncertainty

http://www.uttv.ee/naita?id=17824

https://www.youtube.com/watch?v=NdDK03f0wew



Scheme 7.1. Interrelations between the different error types, the performance characteristics used to estimate them and the ways of expressing the estimates quantitatively. This type of scheme was originally published in article A. Menditto, et al *Accred. Qual. Assur*. 2006, *12*, 45.

The difference between the measured value and the true value is called error or *total* error (see section 1). This error can be divided into two parts – **random error** (having different magnitude and sign in the case of repeated measurements) and **systematic error** (having the same or systematically changing magnitude and sign in the case of repeated measurements). As seen in section 1 errors cannot be known exactly. Therefore instead of errors themselves we operate with *estimates* of errors – the performance characteristics.

Thus, **trueness** is the estimate of the systematic error. For determining trueness we do not need to know the true value but we need to know a **reference value**. Reference value (differently from the true value) has uncertainty, but usually a small one. Different types of **precision** are estimates of the random error. For obtaining the "true" precision we would need to make an infinite number of repeated measurements. There are different types of precision, depending on the conditions under which precision is determined, e.g. repeatability (section 1) and intermediate precision (section 6). Accuracy embraces both trueness and precision and be considered as describing the total error.

These performance characteristics can be quantitatively expressed. **Bias** – difference between the measured value obtained from multiple repeated measurements with the same sample and the reference value – is the quantitative expression of trueness. **Standard deviation** – again obtained from multiple measurements with the same sample – is the quantitative expression of precision. These two can be combined into a **measurement uncertainty** estimate, which can be regarded as the quantitative expression of accuracy.

8. Overview of measurement uncertainty estimation approaches

Brief summary: In this section an overview is given about the main types of approaches that can be used for estimation of measurement uncertainty.

Overview of the approaches for estimating measurement uncertainty

http://www.uttv.ee/naita?id=17704

https://www.youtube.com/watch?v=Ed8S8KN1GlU

This section presents step by step the modeling approach to measurement uncertainty estimation. This approach is described in detail in the ISO GUM[2] and has been interpreted for chemistry in the Eurachem measurement uncertainty guide[3]. It is often called also the "bottom-up" approach. This means that the uncertainties of the input quantities are found and thereafter combined into the combined standard uncertainty. The uncertainty estimation carried out in section 4 in principle also used this approach. The presentation in this section is based on a practical example – determination of ammonium nitrogen in water.

- 9.1. Step 1– Measurand definition
- 9.2. Step 2 – Model equation
- 9.3. Step 3 – Uncertainty sources
- 9.4. Step 4 – Values of the input quantities
- 9.5. Step 5 – Standard uncertainties of the input quantities
- 9.6. Step 6 – Value of the output quantity
- 9.7. Step 7 – Combined standard uncertainty
- 9.8. Step 8 – Expanded uncertainty
- 9.9. Step 9 – Looking at the obtained uncertainty

The slides used in this section can be downloaded from here:

uncertainty_of_photometric_nh4_determination_iso_gum_modeling.pdf 82 KB

Estimation of measurement uncertainty in chemical analysis (analytical chemistry)

Ivo Leito
University of Tartu
Institute of Chemistry

2013

1

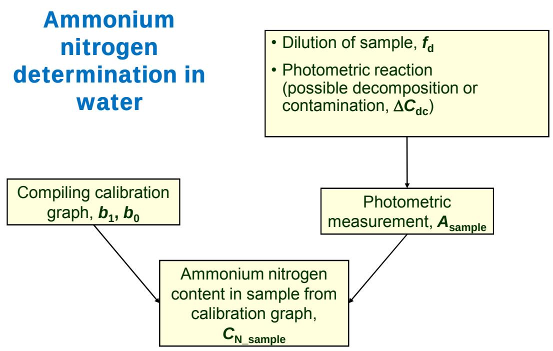
Measurement uncertainty by the modeling approach:

Determination of NH_4^+ in water

- A dye (photometric complex) is formed quantitatively from NH_4^+ and its absorbance is measured at 640-660 nm by a photometer
- The concentration of ammonium nitrogen is found from calibration graph

2

Procedure



3

Step 1 – defining the measurand

Measurand = The quantity intended to be measured

Our measurand:
Concentration of NH_4^+ expressed as ammonium concentration C_{N_sample} mg/l in the water sample

4

Step 2 – Model

Model is the equation which enables calculating the measurand (**output quantity *Y***) value from the values of directly measured quantities (**input quantities X_1 .. X_n**) :

$$Y = f(X_1, X_2, \ldots\ldots, X_n)$$

5

Step 2 – Model

- Model:

$$\boxed{C_{N\_sample} = \frac{(A_{sample} - b_0)}{b_1} \times f_d + \Delta C_{dc}}$$

- A_{sample} – absorbance of the dye solution obtained from the sample
- b_1 and b_0 – slope and intercept of the calibration graph
- f_d – dilution factor
- ΔC_{dc} – component taking into account uncertainty originating from possible decomposition or contamination

6

Step 3 – Uncertainty sources

- **All possible uncertainty sources need to be considered**
  - **The important ones need to be accounted for**
  - **This can be done individually or by grouping**
- For this the source has to be linked with some input quantity in the model
- If an important uncertainty source exists that cannot be linked with any input quantity then the model has to be modified

7

Step 3 – Uncertainty sources

– **Sampling**
- Sample non-representativeness

The result is expressed for sample, sampling is not included

– **Sample preparation**
- Inhomogeneity
- Separation of analyte incomplete
- Analyte adsorbs

The sample is homogenous, the analyte is not separated and does not adsorb

- Analyte or photometric complex decomposes
- Analyte volatilizes
- Incomplete reaction
- Contamination

Analyte or the photometric complex can decompose or get contaminated: ΔC_{dc}

8

Step 3 – Uncertainty sources

– Preparation and dilution of solutions

f_d accounts for this

– Weighing

Is included in b_1 and b_0 uncertainty

– Calibration of instrument
- Standard substance purity
- Solution preparation

b_1 and b_0 uncertainty

– Measurement of sample
- Interferences
- Repeatability of reading
- Drift of reading

A_{sample} accounts for these

- Memory effects

Absent in our example

9

Step 3 – Uncertainty sources

- Interferences
- Repeatability and drift of photometer

- Contamination
- Decomposition
- Volatilization

$$C_{N\_sample} = \frac{(A_{sample} - b_0)}{b_1} \times f_d + \Delta C_{dc}$$

- Standard substance purity
- Preparation of solutions
- Repeatability and drift of photometer

- Preparation of solutions

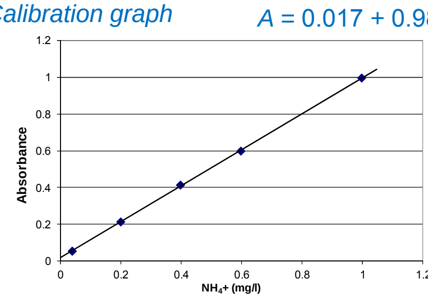
10

Step 4 – Finding values of input quantities

$A = b_0 + b_1 \times C$

Calibration graph

$A = 0.017 + 0.981 \times C$



11

Step 4 – Finding values of input quantities

Quantity	Value	Unit
A_{sample}	0.1860	AU*
b_0	0.0171	AU*
b_1	0.9808	AU×l/mg
f_d	1.2500	–
ΔC_{dc}	0.0000	mg/l

* Absorbance in fact does not have a unit, AU is used for clarity

12

Step 5 – Standard uncertainties of the input quantities: A_{sample}

- Absorbance of the sample solution A_{sample}:

u (A_{sample}, rep) =	0.0010	AU
u (A_{sample}, drift) =	0.0012	AU
u (A_{sample}, chem) =	0.0030	AU

$$u(A_{\text{sample}}) = \sqrt{u(A_{\text{sample}},\text{rep})^2 + u(A_{\text{sample}},\text{drift})^2 + u(A_{\text{sample}},\text{chem})^2} = 0.0034\,\text{AU}$$

13

Step 5 – Standard uncertainties of the input quantities: b_0 and b_1

- Standard deviations of b_0 and b_1 as found from regression statistics are used as standard uncertainty estimates

$u(b_0)$ = 0.0025 AU

$u(b_1)$ = 0.0046 AU×l/mg

- This is an approximation!

14

Step 5 – Standard uncertainties of the input quantities: f_d

- The standard uncertainty of dilution factor is estimated here as 0.5% of the dilution factor value
- This is a safe estimate if volumetric operations are performed correctly

$u(f_d)$ = 1.25 / 200 = 0.0063

15

Step 5 – Standard uncertainties of the input quantities: ΔC_{dc}

- The *possible* contribution of decomposition or contamination at this concentration level is estimated from the experience of the laboratory as follows:

$u(\Delta C_{dc})$ = 0.004 mg/l

16

Step 5 – Standard uncertainties of the input quantities: summary

- **The uncertainties of the input quantities have to be used as standard uncertainties (*u*)**

Quantity	Value	*u*	Unit
A_{sample}	0.1860	0.0034	AU
b_0	0.0171	0.0025	AU
b_1	0.9808	0.0046	AU×l/mg
f_d	1.2500	0.0063	–
ΔC_{dc}	0.0000	0.0040	mg/l

17

Step 6 – Calculating the measurand value

$$\boxed{C_{\text{N\_sample}} = \frac{(A_{\text{sample}} - b_0)}{b_1} \times f_{\text{d}} + \Delta C_{\text{dc}}}$$

$$C_{\text{N\_sample}} = \frac{(0.1860 - 0.0171)}{0.9808} \times 1.25 + 0$$

C_{N_sample} = 0.215 mg/l

18

Step 7 – Finding combined standard uncertainty (1)

- In the case on non-correlating input quantities:

$$u_c(y) = \sqrt{\left[\frac{\partial Y}{\partial X_1}u(x_1)\right]^2 + \left[\frac{\partial Y}{\partial X_2}u(x_2)\right]^2 + \ldots.. + \left[\frac{\partial Y}{\partial X_n}u(x_n)\right]^2}$$

$u_c(y)$ = combined standard uncertainty of the output quantity
$u(x_i)$ = standard uncertainties of the input quantities

19

Step 7 – Finding combined standard uncertainty (2)

$$u_c(C_{\mathrm{N\_sample}}) = \sqrt{\left(\frac{\partial C_{\mathrm{N\_sample}}}{\partial A_{\mathrm{sample}}}u(A_{\mathrm{sample}})\right)^2 + \left(\frac{\partial C_{\mathrm{N\_sample}}}{\partial b_0}u(b_0)\right)^2 + \left(\frac{\partial C_{\mathrm{N\_sample}}}{\partial b_1}u(b_1)\right)^2 + \left(\frac{\partial C_{\mathrm{N\_sample}}}{\partial f_{\mathrm{d}}}u(f_{\mathrm{d}})\right)^2 + \left(\frac{\partial C_{\mathrm{N\_sample}}}{\partial \Delta C_{\mathrm{dc}}}u(\Delta C_{\mathrm{dc}})\right)^2}$$

20

Step 7 – Finding combined standard uncertainty (3)

$$u_c(C_{\mathrm{N\_sample}}) = \sqrt{(0.00429)^2 + (-0.00324)^2 + (-0.0010)^2 + (0.00108)^2 + (0.00400)^2}$$

$u_c(C_{\mathrm{N\_sample}})$ = 0.00686 mg/l

21

Step 8 – Finding expanded uncertainty

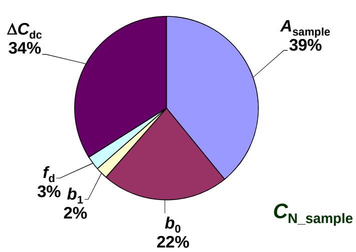
- Expanded uncertainty ***U*** is found by multiplying $\boldsymbol{u_c}$ with coverage factor ***k***
  - Very often k = 2, which in the case of normal distribution corresponds to ca 95% probability

U = 0.00686 x 2 = 0.014 mg/l

Result: **$C_{\mathrm{N\_sample}}$ = (0.215 ± 0.014) mg/l**
k = 2

22

Step 9 – Contributions of uncertainty sources



The most important uncertainty sources:

- Interference
- Contamination
- Instability of the dye

$C_{\mathrm{N\_sample}}$ = (0.215 ± 0.014) mg/l
k = 2, norm.

$$C_{\mathrm{N\_sample}} = \frac{(A_{\mathrm{sample}} - b_0)}{b_1} \times f_{\mathrm{d}} + \Delta C_{\mathrm{dc}}$$

23

Step 9 – Contributions of uncertainty sources

- Uncertainty contributions (indexes) are found as follows:

$$Index(x_1) = \frac{\left[\frac{\partial Y}{\partial X_1}u(x_1)\right]^2}{\left[\frac{\partial Y}{\partial X_1}u(x_1)\right]^2 + \left[\frac{\partial Y}{\partial X_2}u(x_2)\right]^2 + \ldots.. + \left[\frac{\partial Y}{\partial X_n}u(x_n)\right]^2}$$

24

9. The ISO GUM Modeling approach

9.1. Step 1– Measurand definition

Definition of the measurand

http://www.uttv.ee/naita?id=17636

https://www.youtube.com/watch?v=P_i268pDgvM

The measurand definition is the most basic step of any measurement. In this step it is defined what is actually measured and this definition is also the basis for the measurement procedure and model equation.

The measurand in this case is concentration of NH_4^+ expressed as ammonium ion concentration *C*N_sample [mg/l] in the water sample.

9.2. Step 2 – Model equation

Model equation

http://www.uttv.ee/naita?id=17637

https://www.youtube.com/watch?v=N45YNwToyac

The model equation (equation 9.1) enables calculating the **output quantity** value (result value) from the **input quantity** values. Input quantities are the directly measured quantities (or are calculated from directly measured quantities). In addition, the model equation has to enable accounting for all important measurement uncertainty sources. In the case of this analysis the model equation is the following:

$$C_{\mathrm{N\_sample}} = \frac{(A_{\mathrm{sample}} - b_0)}{b_1} \times f_{\mathrm{d}} + \Delta C_{\mathrm{dc}} \quad (9.1)$$

The output quantity is $C_{\mathrm{N\_sample}}$ – ammonium ion concentration in the water sample.

The input quantities are:

A_{sample} – absorbance of the dye solution obtained from the sample;
b_1 and b_0 – slope and intercept of the calibration graph;
f_{d} – dilution factor;
ΔC_{dc} – component taking into account uncertainty originating from possible decomposition or contamination.

The parameters A_{sample}, b_1, b_0 and f_{d} in the equation account for the directly measured input quantities. In contrast, ΔC_{dc} is introduced only for taking into account uncertainty due to possible decomposition of the photometric complex and due to possible contamination. Its value is zero as will be seen in section 9.4, so, it does not contribute to the value of $C_{\mathrm{N\_sample}}$. However, its uncertainty is different from zero and therefore will contribute to the uncertainty of $C_{\mathrm{N\_sample}}$.

9.3. Step 3 – Uncertainty sources

All possible uncertainty sources have to be considered and those that are likely to be influential have to be taken into account.

Uncertainty sources: in general

http://www.uttv.ee/naita?id=17638

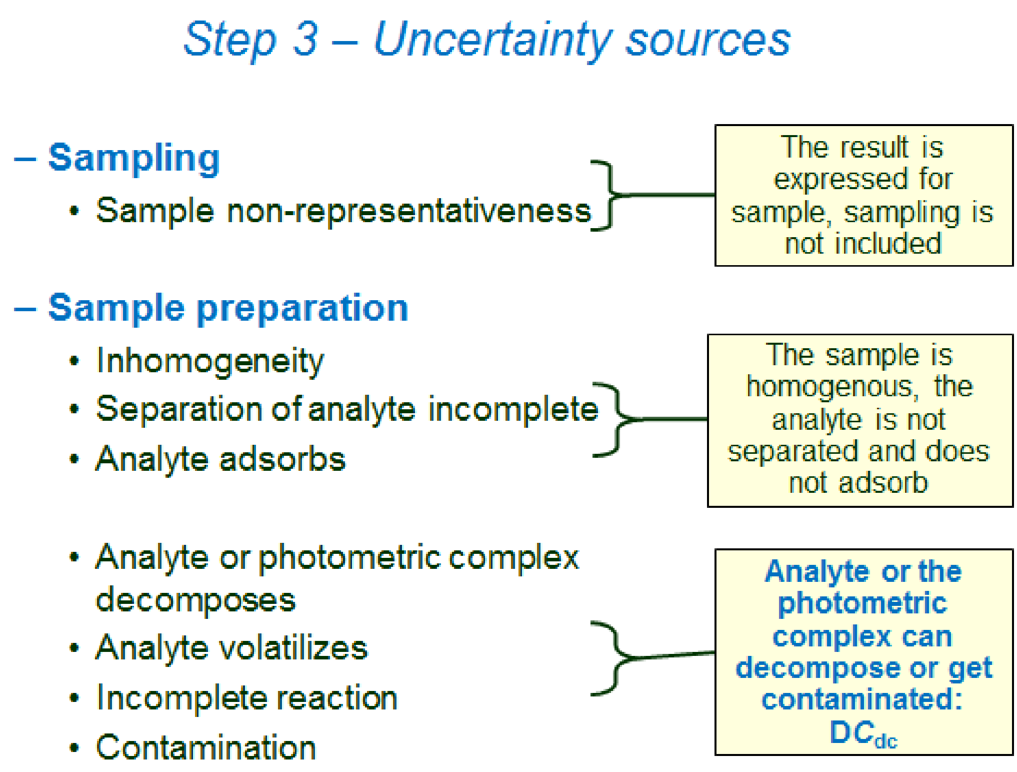
https://www.youtube.com/watch?v=EWYNWtRgvmI

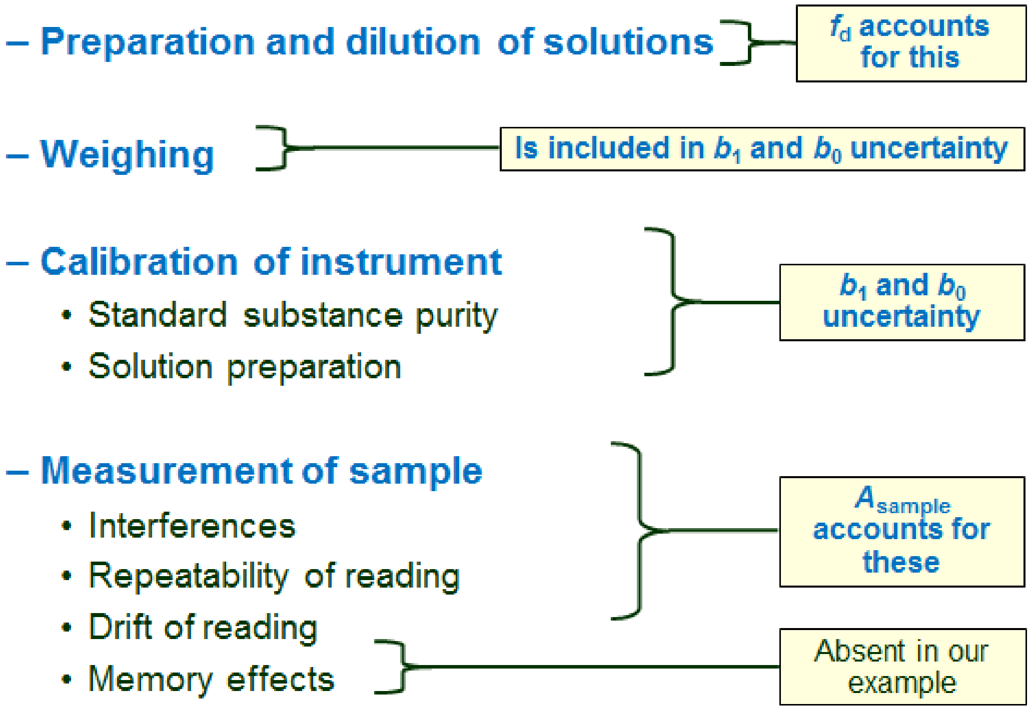
Uncertainty sources: one by one

http://www.uttv.ee/naita?id=17639

https://www.youtube.com/watch?v=kEfhKCj44N0

The following schemes list the main uncertainty sources in chemical analysis and comment on the presence or absence of the respective uncertainty sources in our case. Comment on the memory effect as uncertainty source: Memory effect is problematic first of all in the case of trace analysis and secondly, if the analyte is specifically prone to adsorption on glass or plastic surfaces. In this case, although the concentration of ammonium in the sample is quite low, it is still not yet a true trace analysis. Ammonium ion is by no means a strongly adsorbing species. Therefore we can leave the memory effect as uncertainty source out of consideration.





***

[1] Comment on the memory effect as uncertainty source: Memory effect is problematic first of all in the case of trace analysis and secondly, if the analyte is specifically prone to adsorption on glass or plastic surfaces. In this case, although the concentration of ammonium in the sample is quite low, it is still not yet a true trace analysis. Ammonium ion is by no means a strongly adsorbing species. Therefore we can leave the memory effect as uncertainty source out of consideration.

9.4. Step 4 – Values of the input quantities

Finding the values of the input quantities

http://www.uttv.ee/naita?id=17640

https://www.youtube.com/watch?v=TP8AguK1jSo

The values of Asample, b1, b0 and fd are found from the measured data (b1 and b0 are found from regression analysis). The value of ΔCdc is zero.

9.5. Step 5 – Standard uncertainties of the input quantities

Standard uncertainties of the input quantities

http://www.uttv.ee/naita?id=17641

https://www.youtube.com/watch?v=ZIEhfCoNyxU

The standard uncertainty of A_{sample} is found from the following three uncertainty components:

(1) The uncertainty due to the repeatability of photometric measurement:

$$u\ (A_{sample},\ \text{rep}) = 0.0010\ \text{AU} \quad (9.2)$$

This uncertainty includes the repeatability of the instrument, repeatability of positioning the cell in the instrument and possible disturbances, such as a random dust particle on the optical windows of the cell.

(2) The uncertainty due to the possible drift of the spectrophotometer parameters:

$$u\ (A_{sample},\ \text{drift}) = 0.0012\ \text{AU} \quad (9.3)$$

(3) The uncertainty due to the possible interfering effects:

$$u\ (A_{sample},\ \text{chem}) = 0.0030\ \text{AU} \quad (9.4)$$

These can be due to some compound absorbing (or scattering) light at the same wavelength that is used for measurement (leading to increase of the absorbance value) or due to some disturbance in forming the photometric complex (leading to decrease of the absorbance value).

As a result:

$$u(A_{\text{sample}}) = \sqrt{u(A_{\text{sample}},\text{rep})^2 + u(A_{\text{sample}},\text{drift})^2 + u(A_{\text{sample}},\text{chem})^2} = 0.0034\ \text{AU} \quad (9.5)$$

The standard uncertainties of the slope b_1 and intercept b_0 are found as standard deviations of the respective regression coefficients (see the XLS files in section 9.7). This is an approximate way of taking into account the uncertainty due to linear regression analysis, because it (1) neglects the systematic effects affecting all the points on the regression line and (2) neglects the negative correlation between b_1 and b_0 (which always exists). The first of these effects leads to underestimation of uncertainty and the second one to overestimation of uncertainty. So, this approach should only be used if it is not expected that linear regression analysis will be among the main contributors to uncertainty. The same example solved with full rigor is available from http://www.ut.ee/katsekoda/GUM_examples/. Please look at the example "Ammonium by Photometry" with elaboration level "High (uncertainty estimated at full rigor, suitable for experts)". Comparison of the obtained combined uncertainties: 0.00686 mg/l obtained here (section 9.7) and 0.0065 mg/l obtained with full rigor shows that this approach is acceptable, especially since it leads rather to uncertainty overestimation than underestimation.

The $u(f_d)$ is found on an assumption that the relative combined standard uncertainty of all involved volumetric operations is not higher than 0.5%. If volumetric operations are carried out carefully then this is a safe assumption under usual laboratory conditions. Considering that the value of f_d is 1.25 (unitless) we get the following:

$$u\ (f_d) = 1.25 \cdot 0.5\%\ /\ 100\% = 0.0065\ \text{(unitless)} \quad (9.6)$$

The uncertainty of ΔC_{dc} accounts for *possible* decomposition of the photometric complex and *possible* contamination of the sample. The word "*possible*" is stressed here: it is well possible that actually there is neither decomposition of the photometric complex nor contamination of the sample. However, in order to rigorously establish this, extensive research would be needed. Therefore in this example we use an estimate based on experience from our laboratory:

$$u\ (\Delta C_{dc}) = 0.004 \text{ mg/l}$$

The standard uncertainties of the input quantities are summarized in table:

Quantity	Value	u	Unit
A_{sample}	0.1860	0.0034	AU
b_0	0.0171	0.0025	AU
b_1	0.9808	0.0046	AU×l/mg
f_d	1.2500	0.0063	–
ΔC_{dc}	0.0000	0.0040	mg/l

***

[1] The same example solved with full rigor is available from http://www.ut.ee/katsekoda/GUM_examples/. Please look at the example "Ammonium by Photometry" with elaboration level "High (uncertainty estimated at full rigor, suitable for experts)". Comparison of the obtained combined uncertainties: 0.00686 mg/l obtained here (section 9.7) and 0.0065 mg/l obtained with full rigor shows that this approach is acceptable, especially since it leads rather to uncertainty overestimation than underestimation.

9.6. Step 6 – Value of the output quantity

Calculating the value of the output quantity

http://www.uttv.ee/naita?id=17642

https://www.youtube.com/watch?v=OKzpnlMsjJM

The output quantity value (the measurand value) is calculated from the input quantity values (section 9.4) using the mathematical model (section 9.2). The measurand value in this example is: *C*N_sample = 0.215 mg/l.

9.7. Step 7 – Combined standard uncertainty

Calculating the combined standard uncertainty

http://www.uttv.ee/naita?id=17643

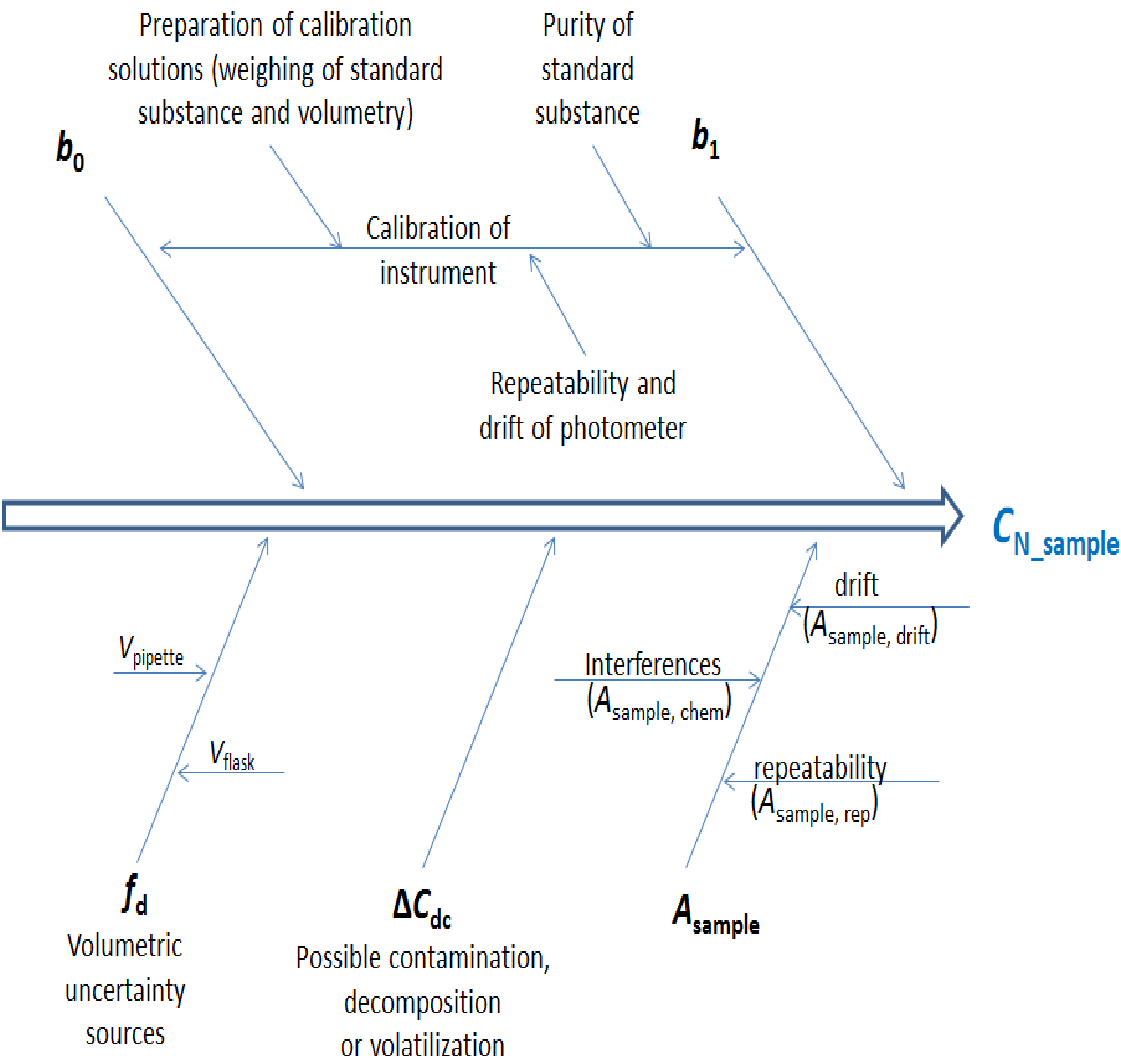
https://www.youtube.com/watch?v=v6JZT2PkE-M

Numerical calculation of the uncertainty components: the Kragten method

http://www.uttv.ee/naita?id=17721

https://www.youtube.com/watch?v=qRx8cFVitgk

The influence of measurement uncertainty sources, grouped according to input quantities, on the result can be schematically presented in the form of the so-called "fish-bone" diagram:



Scheme 9.1. Cause effect diagram: Photometric ammonia determination

The initial XLS file (i.e. containing only the data but not the calculations) used in this example and the XLS file containing the also the combined standard uncertainty (and expanded uncertainty) calculation according to the Kragten's approach can be downloaded from here:

- uncertainty_of_photometric_nh4_determination_kragten_initial.xls 49 KB
- uncertainty_of_photometric_nh4_determination_kragten_solved.xls 52 KB

9.8. Step 8 – Expanded uncertainty

The expanded uncertainty can be found at two different levels of sophistication. The simpler approach uses simply a preset *k* value (most often 2) and the actual coverage probability is not discussed. This approach is presented in the first video lecture.

Finding the expanded uncertainty (simpler approach)

http://www.uttv.ee/naita?id=17644

https://www.youtube.com/watch?v=KomDnLRArDs

The second approach is more sophisticated. It is an approximation approach based on the assumption that the distribution function of the output quantity can be approximated by a Student distribution with the effective number of degrees of freedom found by the so-called Welch-Satterthwaite method. This enables then to use the Student coefficient corresponding to a desired level of confidence (coverage probability) as the coverage factor. This approach is explained in the second video lecture.

Finding the expanded uncertainty (the Welch-Satterthwaite method)

http://www.uttv.ee/naita?id=17916

https://www.youtube.com/watch?v=CyIWJjG_8ck

The XLS file containing the combined standard uncertainty and expanded uncertainty calculation and the XLS file containing the expanded uncertainty calculation using coverage factor found using the effective number of degrees of freedom form the Welch-Satterthwaite approach can be downloaded from here:

- uncertainty_of_photometric_nh4_determination_kragten_solved.xls 52 KB
- uncertainty_of_photometric_nh4_determination_kragten_solved_df.xls 48 KB

9.9. Step 9 – Looking at the obtained uncertainty

Calculation of the uncertainty components (uncertainty indexes) is explained and demonstrated in section 9.7.

Analysis of the uncertainty sources

http://www.uttv.ee/naita?id=17645

https://www.youtube.com/watch?v=zE5yi6NwAvI

Conclusions 1

http://www.uttv.ee/naita?id=17647

https://www.youtube.com/watch?v=7VqqLGLQHgg

Conclusions 2

http://www.uttv.ee/naita?id=18096

https://www.youtube.com/watch?v=ENg2aTyBwKg

Brief summary: This section explains the so-called single-lab validation approach. We will look at the formalization of this approach published by Nordtest. Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories. B. Magnusson, T. Näykki, H. Hovind, M. Krysell. Nordtest technical report 537, ed. 3. Nordtest, 2011. Available on-line from http://www.nordtest.info/index.php/technical-reports/item/handbook-for-calculation-of-measurement-uncertainty-in-environmental-laboratories-nt-tr-537-edition-3.html Therefore in this course this approach is often called "the Nordtest approach". The single-lab validation approach, contrary to the ISO GUM modeling approach, does not go deeply into the measurement procedure and does not attempt to quantify all uncertainty sources individually. Instead uncertainty sources are quantified in large "batches" via components that take a number of uncertainty sources into account. Most of the data that are used come from validation of the analytical procedure. This is the reason for the word "validation" in the name of the approach. This type of approach is also sometimes called the "top-down" approach.

Sections 10.1 to 10.3 present the Nordtest approach step by step and explain the way of obtaining necessary data. Section 10.4 gives a "roadmap" of the Nordtest approach. Section 10.5 presents a practical example of applying the Nordtest approach in the case of determination of acrylamide in snacks by liquid chromatography mass spectrometry (LC-MS).

***

[1] *Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories*. B. Magnusson, T. Näykki, H. Hovind, M. Krysell. Nordtest technical report 537, ed. 3. Nordtest, **2011**. Available on-line from http://www.nordtest.info/index.php/technical-reports/item/handbook-for-calculation-of-measurement-uncertainty-in-environmental-laboratories-nt-tr-537-edition-3.html

***

The slides presented in this section are available from here:

single-lab_validation_nordtest_uncertainty.pdf 33 KB

Uncertainty estmation approach based on validation and Quality Control Data “the Nordtest approach"

Ivo Leito
University of Tartu
Institute of Chemistry

2013

Nordtest Technical Report 537, 3rd ed (2011)
http://www.nordtest.info/

Single-laboratory validation approach

Effects contributing to uncertainty

Random | **Systematic**

The two groups of uncertainty contributions are quantified separately and then combined:

$$u_c = \sqrt{u_1^{\ 2} + u_2^{\ 2}}$$

Uncertainty arising from random effects

Uncertainty accounting for *possible* bias

at „long term“ level!

Single lab validation approach: in practice (1)

The main equation:

$$u_c = \sqrt{u(R_W)^2 + u(bias)^2}$$

Within-laboratory reproducibility

This component accounts for the random effects

Uncertainty of the estimate of the laboratory and the method bias

This component accounts for the systematic effects

This and subsequent equations work with absolute and relative values

Nordtest Technical Report 537, 3rd ed (2011)
http://www.nordtest.info/

Absolute vs relative uncertainties: Rules of Thumb

- **At low concentrations (near detection limit, trace level) use absolute uncertainties**
  - Uncertainty is not much dependent on analyte level
- **At medium and higher concentrations use relative uncertainties**
  - Uncertainty is roughly proportional to analyte level
- **In general: whichever is more constant**

Appendix E.4 from Quantifying Uncertainty in Analytical Measurement, EURACHEM/CITAC Guide, Second Edition (2000)
Available from : http://www.eurachem.org/

Single lab validation approach: in practice

Steps of the process:

1. Specify measurand
2. Quantify R_w component ***u*(*R*w)**
3. Quantify bias component ***u*(*bias*)**
4. Convert components to standard uncertainties ***u*(*x*)**
5. Calculate combined standard uncertainty u_c
6. Calculate expanded uncertainty ***U***

u(R_w)

- $u(R_w)$ is the uncertainty component that takes into account long-term variation of results within lab, that means: within-lab reproducibility (s_{Rw})
- Ideally:
  - The same sample
    - Sample similar to test samples – matrix, concentration, homogeneity
  - The same lab
  - The same procedure
  - Different days (preferably over 1 year)
  - Different persons
  - Different reagent batches
  - …

Include sample preparation!

Repeatability	<	Within-lab reproducibility	<	Combined uncertainty
s_r	<	s_{Rw}	<	u_c

$u(R_w)$

$$u(R_w) = s_{Rw}$$

Ideally: separately for different matrices and different concentration levels!

The control sample analysis has to cover the whole analytical process

u(bias)

- The *possible* **bias** of lab's results from the best estimate of true value is taken into account

Include sample preparation!

- *u*(*bias*) can be found:
  - From **repeated** analysis of the same samples with a reference procedure
  - From **repeated** analysis of certified reference materials (CRMs)
  - From **repeated** interlaboratory comparison measurements
  - From **repeated** spiking experiments

Ideally: several reference materials, several PTs because the bias will in most cases vary with matrix and concentration range

u(bias)

$$u(bias) = \sqrt{RMS_{\text{bias}}{}^2 + u(Cref)^2}$$

This component accounts for the average bias of the laboratory results from the C_{ref}

This component accounts for the average uncertainty of the reference values C_{ref}

u(bias)

- The averaging is done using the **root mean square:**

$$bias_i = Clab_i - Cref_i \qquad RMS_{bias} = \sqrt{\frac{\sum (bias_i)^2}{n}}$$

$$u(Cref_i) = \frac{s_i}{\sqrt{n_i}} \qquad u(Cref) = \sqrt{\frac{\sum u(Cref_i)^2}{n}}$$

- *n*: the number of bias estimates used
  - If *n* is too small then the bias component will include a large share of random effects and may be overestimated

u(bias): only one CRM

- If only one single CRM is used:

$$u(bias) = \sqrt{RMS_{\text{bias}}{}^2 + s_{\text{bias}}^2 / n + u(Cref)^2}$$

Uncertainty due to possible bias

Evaluation of uncertainty due to bias, ideally:

- **Separately for different sample matrices**
- **Separately for different concentration levels**

Roadmap:

Possible bias

$u(Cref_i)$ from certificates

$$u(Cref_i) = \frac{s_i}{\sqrt{n_i}} \rightarrow u(Cref) = \sqrt{\frac{\sum u(Cref_i)^2}{n}}$$

$$bias_i = Clab_i - Cref_i$$

$$RMS_{bias} = \sqrt{\frac{\sum (bias_i)^2}{n}} \rightarrow u(bias) = \sqrt{RMS_{bias}^{\ 2} + u(Cref)^2}$$

Uncertainty due to random effects

$$u(R_w) = s_{RW}$$

Combined standard uncertainty

$$u_c = \sqrt{u(R_w)^2 + u(bias)^2}$$

13

10. The single-lab validation approach

10.1. Principles

In the Nordtest approach the uncertainty is regarded as being due to two components:

1. The **within-lab reproducibility** (intermediate precision) component. This uncertainty component takes into account all uncertainty sources that are **random in the long term** (i.e. several months, preferably one year). So, quite some uncertainty sources that are systematic within a day will become random in the long term. A typical example is titration if new titrant is prepared weekly. Within a given week the titrant concentration is a systematic effect, but in the long term it becomes random, because many batches of titrant will be involved. A similarly typical example is calibration graph if it is prepared daily: every day the possible bias in calibration is a systematic effect, but in the long term it becomes random.
2. The **bias component**. This component takes into account the systematic effects that cause long-term bias (but not those that just cause bias within a given day). The **long-term bias** can be regarded as sum of procedure bias (bias inherent in the nature of the procedure) and laboratory bias (bias caused by the way how the procedure is implemented in the laboratory).

Introduction to uncertainty estmation based on validation and quality control data (the Nordtest approach)

http://www.uttv.ee/naita?id=17909

https://www.youtube.com/watch?v=9oOX4CUsWjI

The main equation of the Nordtest approach is here:

$$u_c = \sqrt{u(R_w)^2 + u(bias)^2} \quad (10.1)$$

Here $u\,(R_W)$ stands for the within-lab reproducibility component of uncertainty and u (*bias*) stands for the uncertainty component taking into account possible bias. The resulting measurement uncertainty u_c is not directly related to any specific result, because it is calculated using data from the past measurements. Therefore it can be said that the uncertainty obtained with the Nordtest approach characterizes the analysis procedure rather than a concrete result. If the uncertainty of a concrete result is needed then it is *assigned* to the result.

Because of this it is necessary to decide whether to express the uncertainty in absolute terms (i.e. in the units of the measured quantity) or in the relative terms (i.e. as a ratio of uncertainty to the value of the measured quantity or a percentage of the value of the measured quantity). The rules of thumb:

- **At low concentrations (near detection limit, trace level) use absolute uncertainties**
  Uncertainty is not much dependent on analyte level
- **At medium and higher concentrations use relative uncertainties**
  Uncertainty is roughly proportional to analyte level
- **In general: use whichever is more constant with changing concentration.**

The main equation of the Nordtest approach. Absolute and relative quantities

http://www.uttv.ee/naita?id=17911

https://www.youtube.com/watch?v=MH8CixjySjI

Overview of the practical implementation of the Nordtest approach

http://www.uttv.ee/naita?id=17912

https://www.youtube.com/watch?v=hPrncfXr7Ok

The main steps of the process of measurement uncertainty evaluation with the Nordtest approach:

1. Specify measurand
2. Quantify R_W component $\boldsymbol{u(R_W)}$
3. Quantify bias component $\boldsymbol{u(bias)}$
4. Convert components to standard uncertainties $\boldsymbol{u(x)}$
5. Calculate combined standard uncertainty $\boldsymbol{u_c}$
6. Calculate expanded uncertainty $\boldsymbol{U}$

***

[1] A typical example is titration if new titrant is prepared weekly. Within a given week the titrant concentration is a systematic effect, but in the long term it becomes random, because many batches of titrant will be involved. A similarly typical example is calibration graph if it is prepared daily: every day the possible bias in calibration is a systematic effect, but in the long term it becomes random.

10.2. Uncertainty component accounting for random effects

Estimating the within-lab reproducibility component of uncertainty

http://www.uttv.ee/naita?id=17913

https://www.youtube.com/watch?v=9-803-VViKo

The within-lab reproducibility (intermediate precision) component takes into account all uncertainty sources that are random in the long term (i.e. several months, preferably one year). So, quite some uncertainty sources that are systematic within a day will become random in the long term. The simplest way to find $u\ (R_W)$ is from a number of repeated measurements of a control sample, organized, e.g. as a control chart. Alternatively, the pooled standard deviation approach as explained in section 6 can be used. If this is done then the $u\ (R_W)$ can be found based on several different control samples, so that it will be an average value of all of them.

The number of values used for evaluation of $u\ (R_W)$ must be sufficiently large. An initial estimate of $u\ (R_W)$ can be obtained with 10-15 values but thereafter more data should be collected. Even more importantly, the time period during which the data are collected, must be sufficiently long (at least several months, preferably around a year) so that all the sources of variability in the procedure are taken into account. So, 10 values collected over a five-month time period is a better option than 20 values collected during 1.5 months.

Depending on situation $u\ (R_W)$ can be used as absolute or relative value.

It is important to stress that $u\ (R_W)$ should be estimated separately for different matrixes and different concentration levels.

10.3. Uncertainty component accounting for systematic effects

Estimating the uncertainty component due to possible bias

http://www.uttv.ee/naita?id=17910

https://www.youtube.com/watch?v=hLGZsW_o81o

The component *u* (*bias*) takes into account *possible* bias of the measurement procedure.

Reliably determining the bias of the procedure is not easy for the following reasons:

1. If the random effects are strong then this needs a very large number of measurements. When a limited number of measurements are made then the bias estimate will always contain a contribution from random effects, which will make the bias estimate artificially higher. Even more so – the procedure can actually have no bias at all.
2. Even with formally similar matrixes the bias can differ by magnitude and even by sign (e.g. when determining pesticides in different varieties of apples, determining drug residues in blood plasma from different patients, etc). This means that having determined the bias in one variety of apples this bias is not automatically applicable for another variety and we can speak about uncerteinty in applying the bias to another variety.
3. Bias is always determined against a reference value, which also has an uncertainty.

This is why we speak about *possible* bias and the uncertainty component *u* (*bias*) quantifies our limited knowledge about bias.

Bias refers to difference between our measured value and a reference value. Therefore, for finding *u* (*bias*) we need a sample or a material with a reference value. In broad terms there are four different possibilities how the *u* (*bias*) can be determined:

Possibility of determining *u* (*bias*)	Pros	Cons
Analysing the sample with a reference analysis procedure	Bias can be determined very reliably, because the determined bias corresponds to the bias with exactly the same sample matrix as is in the real sample.	It is usually very difficult to find a suitable reference procedure. Therefore this possibility is not often used at routine labs.
Certified reference material (CRM) In simplified terms certified reference material is a material, in which the content of the analyte (or analytes) is reliably known (the material has a certificate). If the certified reference material's matrix is similar to real samples (i.e. it is not a pure compound or a solution but is e.g. milk powder, soil or blood plasma) then we call it certified matrix reference material.	Bias can be determined quite reliably, because the reference values of CRM-s are generally quite reliable.	Availability of certified reference materials is limited and their matrixes are often better homogenised than in the case of real samples, leading to somewhat optimistic bias estimates.
Using samples of interlaboratory comparisons as reference samples	Is often usable at routine labs, because labs usually participate in	The reference values are mostly derived from participant results and have therefore low reliability (high

	interlaboratory comparisons with the procedures that	uncertainty). This leads to overestimated *u* (*bias*).
Using spiking studies	Can be done at the laboratory and can be done with the real samples, thereby exactly matching the matrix.	The main problem in bias determination by spiking is dispersing the analyte in the sample in the same way as the native analyte in the sample. In the case of inhomogenous matrixes this can be very difficult.

In the case of all four possibilities it is critical to include also sample preparation in bias determination.

In the case of all four possibilities it is necessary to make an as large as possible number of replicate measurements in order to separate the bias from random effects as efficiently as possible. Bias generally changes from matrix to matrix and usually is different at different concentration levels. So, it is important to use several CRMs, several interlaboratory comparisons, etc.

The lower is the reliability of the reference value the higher is the *u* (*bias*) estimate.

The *u* (*bias*) component is found according to the following equation:

$$u(bias) = \sqrt{RMS_{\text{bias}}{}^{2} + u(Cref)^{2}} \quad (10.2)$$

RMS_{bias} is the average (root mean square) bias and is found as follows:

$$RMS_{bias} = \sqrt{\frac{\sum (bias_i)^2}{n}} \quad (10.3)$$

Where n is the number of bias determinations carried out and each $bias_i$ is a result of an individual bias determination and is found as follows:

$$bias_i = Clab_i - Cref_i \quad (10.4)$$

Where $Clab_i$ is a mean of the results of analyte determination in the reference sample (e.g. in the CRM) obtained by the laboratory and $Cref_i$ is the reference value of the reference sample. It is important that $Clab_i$ corresponds to a number of replicates.

u (*Cref*) is the average standard uncertainty of the reference values of the reference samples and is found as follows:

$$u(Cref) = \sqrt{\frac{\sum u(Cref_i)^2}{n}} \quad (10.5)$$

Here u ($Cref_i$) is the standard uncertainty of the i-th reference value. In the case of CRM analysis, spiking or analysis with a reference procedure the u ($Cref_i$) can usually be reasonably found. However, in the case of interlaboratory comparisons where the consensus value of the participants is used as the reference value a reliable uncertainty of the reference value cannot be found. The best estimate in that case would be the standard deviation of the average value after elimination of outliers:

$$u(Cref_i) = \frac{s_i}{\sqrt{n_i}} \quad (10.6)$$

Here s_i is the standard deviation of the participants in the i-th intercomparison after elimination of outlayers and n_i is the number of participants (again after eliminating the outliers) in the i-th intercomparison.

In the special case if a number of bias determinations were carried out using one single CRM the equation 10.2 changes into the following form:

$$u(bias) = \sqrt{RMS_{\text{bias}}{}^2 + s_{\text{bias}}^2 / n + u(Cref)^2} \quad (10.7)$$

Where s_{bias} is the standard deviation of the bias estimates obtained and n is the number of bias estimates obtained.

Depending on situation the $u(bias)$ can be used as absolute or as relative value.

Finally, it is important to stress that the bias uncertainty component should be estimated separately for different matrixes and different concentration levels.

***

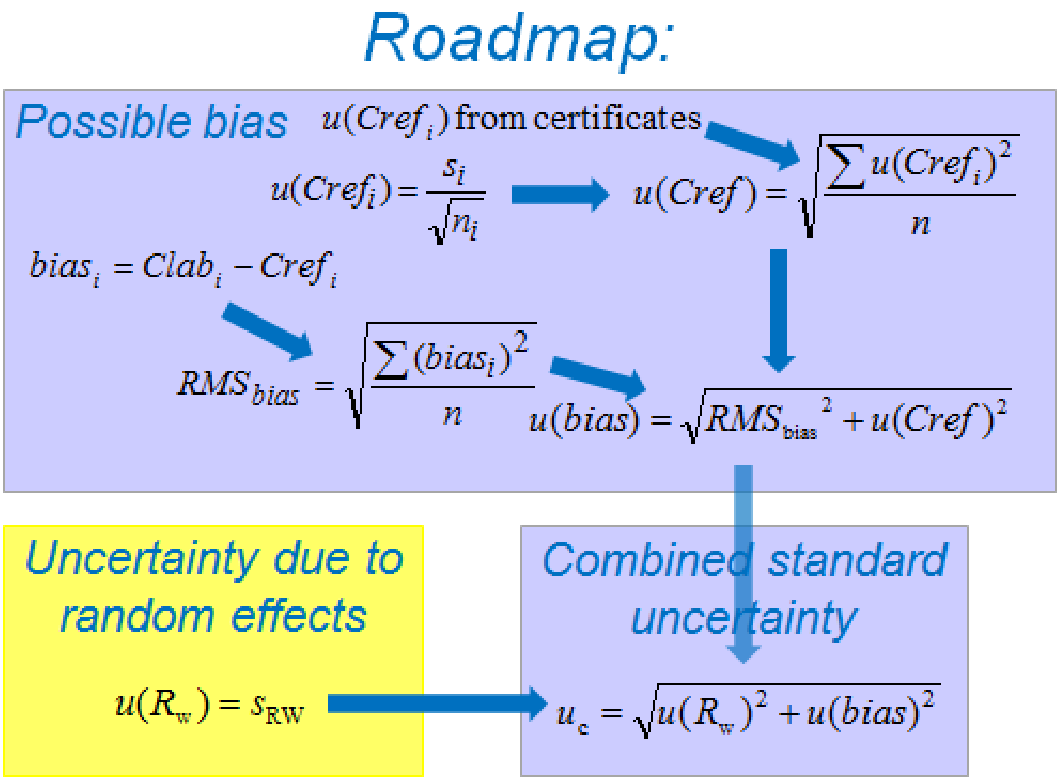
[1] In simplified terms certified reference material is a material, in which the content of the analyte (or analytes) is reliably known (the material has a certificate). If the certified reference material's matrix is similar to real samples (i.e. it is not a pure compound or a solution but is e.g. milk powder, soil or blood plasma) then we call it certified matrix fererence material.

10.4. Roadmap

Roadmap of uncertainty estimation using the Nordtest approach

http://www.uttv.ee/naita?id=17914

https://www.youtube.com/watch?v=KNjUAiq5mEQ



Scheme 10.1. Roadmap of measurement uncertainty estimation using the single-lab validation approach (the Nordtest approach).

10.5. Determination of acrylamide in snacks by LC-MS

Nordtest approach in practice: Determination of acrylamide in snacks by LC-MS Please note that in the tables of the slide "Measurements with the CRMs" the measurement units must be µg/kg, not mg/l.

http://www.uttv.ee/naita?id=18163

https://www.youtube.com/watch?v=P94xWjC6Og0

Some comments on this example:

1. We use relative uncertainties in this example. The reason is that the concentration in these samples is quite high and in LC-MS analysis variability of the results is often roughly proportional to the values of the results.

2. In this example we assume that the matrixes of the used CRMs – potato chips and crisp bread – are sufficiently similar. This means that both reproducibility and bias obtained with these matrixes are similar. In such case the calculated uncertainty is applicable to both of these matrixes. If afterwards we will analyse some matrix that is different from potato chips and crisp bread then we cannot apply to that result the uncertainty estimate that we obtained here. Assessing whether a new matrix is sufficiently similar to the one used for bias evaluation is usually based on experience.

3. The uncertainty estimate $u_c$rel that we have obtained is a relative uncertainty, so it is assigned to a result C_A by the following formula:

$$u_c(C_A) = C_A \frac{u_c \mathrm{rel}}{100\,\%} \quad (10.8)$$

In principle it can be assigned to a result of any magnitude, however, it would not be correct to apply this uncertainty to analysis results that are very different from the one presented here. The difference should not be higher than 3-4 times.

When we use the Nordtest approach for uncertainty estimation then in general we can assume that the number of degrees of freedom is sufficiently large so that the k = 2 uncertainty can be assumed to have roughly 95% coverage probability.

***

[1] Please note that in the tables of the slide "Measurements with the CRMs" the measurement units must be µg/kg, not mg/l.

The slides of this example and the calculation files – both the initial and the solved file – are available from here:

- nordtest_uncertainty_example_acrylamide_lc-ms.pdf 153 KB
- nordtest_uncertainty_example_acrylamide_lc-ms_initial.xls 32 KB
- nordtest_uncertainty_example_acrylamide_lc-ms_solved.xls 33 KB

Single-lab validation approach in practice:

Determination of acrylamide in crisp bread by LC-MS

Ivo Leito
University of Tartu
Institute of Chemistry

2013

1

Single-lab validation approach in practice: Determination of acrylamide in crisp bread by LC-MS

- Concentration level 998 µg/kg
- Laboratory has analysed two certified reference materials (CRMs) with similar matrixes
  - Potato chips and crisp bread
  - The crisp bread CRM is also used as a control sample

2

Certified reference material (CRM)

- The **crisp bread** CRM has the following acrylamide content:

$$C_{\text{acrylamide}} = (1179 \pm 68)\ \mu\text{g/kg} \quad (k = 2,\ \text{norm.})$$

- The **potato chips** CRM has the following acrylamide content:

$$C_{\text{acrylamide}} = (860 \pm 42)\ \mu\text{g/kg} \quad (k = 2,\ \text{norm.})$$

3

Measurements with the CRMs

Crisp bread

Days	*C* (µg/kg)
5.01.2008	1172
6.03.2008	1186
3.04.2008	1153
8.01.2009	1151
18.03.2009	1181
3.04.2009	1147
11.04.2009	1097
16.04.2009	1102
25.04.2009	1162
3.08.2009	1138
28.08.2009	1122
27.11.2009	1191

Mean:	1150 µg/kg
Std Dev:	31 µg/kg

Potato chips

Days	*C* (µg/kg)
3.04.2008	845
3.04.2008	832
3.04.2008	802
27.04.2008	829
27.04.2008	851
27.04.2008	834

Mean:	832 µg/kg
Std Dev:	17 µg/kg

4

Roadmap:

Possible bias

$u(Cref_i)$ from certificates

$$u(Cref_i) = \frac{s_i}{\sqrt{n_i}} \rightarrow u(Cref) = \sqrt{\frac{\sum u(Cref_i)^2}{n}}$$

$$bias_i = Clab_i - Cref_i$$

$$RMS_{bias} = \sqrt{\frac{\sum (bias_i)^2}{n}} \rightarrow u(bias) = \sqrt{RMS_{bias}^{\ 2} + u(Cref)^2}$$

Uncertainty due to random effects

$$u(R_w) = s_{RW}$$

Combined standard uncertainty

$$u_c = \sqrt{u(R_w)^2 + u(bias)^2}$$

5

Finding $u(R_w)$

$$u(R_w) = s_{RW} = 31\ \mu\text{g/kg}$$

$$u(R_w)\_rel = s_{RW}\_rel = 31/1150 \cdot 100 = 2.70\ \%$$

6

Finding u(bias)

Ref value µg/kg	U (k=2) µg/kg	u_c µg/kg	Lab result µg/kg	u_c_rel %	$bias_i$ µg/kg	bias_rel %
1179	68	34	1150	2.88	-29	-2.45
860	42	21	832	2.44	-28	-3.24

u(Cref)_rel=	2.67 %
RMS_{bias}_rel=	2.87 %
u(bias)_rel =	3.92 %

Result:

u_c_rel =	4.8 %
u_c =	48 µg/kg
U_rel (k=2) =	9.5 %
U (k=2) =	95 µg/kg

- Acrylamide content in the sample

$C_{acrylamide}$ = (998 ± 95) µg/kg   (k = 2, norm.)

11. Comparison of the approaches

Brief summary: This section summarized the main properties of the uncertainty estimation approaches, their advantages and drawbacks.

Whenever possible, one of the so-called single-lab approaches should be used. The interlaboratory approaches are only suitable for getting very crude uncertainty estimates. We recommend using them only in case when the laboratory actually does not have the measurement in place yet and wants to know, approximately what uncertainty can be obtained.

If the laboratory has competence and time to carry out investigations of the analytical procedures the the modeling approach is often suitable. If the laboratory has limited time, but has validation and quality control data then the single-lab validation approach is the most suitable.

Comparison of measurement uncertainty estimation approaches

http://www.uttv.ee/naita?id=17917

https://www.youtube.com/watch?v=TlpJ1c-9Rx8

The following table summarizes the pros and cons of the approaches:

Modelling

- Advanced laboratories
  - Extra work usually required
  - Deep knowledge required
- Danger to underestimate uncertainty
- Promotes **thinking**, high value in **teaching**

Single-lab validation

- Routine laboratories
  - Lots of data needed
  - **Less extra work** required
- Realistic uncertainty estimates
- Teaching value is lower than with modelling

Interlaboratory validation

- Minimal work or knowledge required
- s_R value has to be known
- Crude uncertainty estimates

PT approach

- Minimal work or knowledge required
- Very crude uncertainty estimates
- Should be used only as first approximation

Brief summary: This section explains that measurement uncertainty estimates are indispensable if we want to compare two measurement results.

Comparing measurement results using measurement uncertainty estimates

http://www.uttv.ee/naita?id=18095

https://www.youtube.com/watch?v=I6nYn6Pe7f0

12. Comparing measurement results

13. Additional materials and case studies

Brief summary: This section collects some additional materials, examples and case studies on specific chemical analysis techniques.

13.1. Different analytical techniques

A large number of measurement uncertainty estimation examples (example uncertainty budgets) is available from the following address: http://www.ut.ee/katsekoda/GUM_examples/

13.2. Dissolved oxygen by Winkler method

Dissolved oxygen (DO) is one of the most important dissolved gases in water. Sufficient concentration of DO is critical for the survival of most aquatic plants and animals [2] as well as in waste water treatment. DO concentration is a key parameter characterizing natural and wastewaters and for assessing the state of environment in general. Besides dissolved CO_2, DO concentration is an important parameter shaping our climate. It is increasingly evident that the concentration of DO in oceans is decreasing [3 - 6].

Accurate measurements of DO concentration are very important for studying these processes, understanding their role and predicting climate changes.

The Winkler titration method is considered the most accurate method for DO concentration measurement. Careful analysis of uncertainty sources relevant to the Winkler method was carried out and the results are presented as a **„Report on improved high-accuracy Winkler method for determination of dissolved oxygen concentration“**.

In this report it is described how the Winkler method was optimized for minimizing all uncertainty sources as far as practical. The most important improvements were: gravimetric measurement of all solutions, pre-titration to minimize the effect of iodine volatilization, accurate amperometric end point detection and careful accounting for dissolved oxygen in the reagents. As a result, the developed method is possibly the most accurate method of determination of dissolved oxygen available. Depending on measurement conditions and on the dissolved oxygen concentration the combined standard uncertainties of the method are in the range of 0.012 – 0.018 mg dm^{-3} corresponding to the k = 2 expanded uncertainty in the range of 0.023 – 0.035 mg dm^{-3} (0.27 – 0.38%, relative). This development enables more accurate calibration of electrochemical and optical dissolved oxygen sensors for routine analysis than has been possible before. Most of this report is based on the article I. Helm, L. Jalukse, I. Leito, Anal. Chim. Acta. 741 (2012) 21-31 (ref [1]).

Preparation of this report was supported by the European Metrology Research Programme (EMRP), project ENV05 "Metrology for ocean salinity and acidity".

EMRP
European Metrology Research Programme
Programme of EURAMET

The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union

1. I. Helm, L. Jalukse, I. Leito, Anal. Chim. Acta. 741 (2012) 21-31.
2. R.F. Keeling, H.E. Garcia, PNAS, 99(12) (2002) 7848 – 7853.
3. G. Shaffer, S. M. Olsen, J. O. P. Pedersen, Nat. Geosci. 2 (2009) 105-109.
4. R. F. Keeling, A. Körtzinger, N. Gruber, Annu. Rev. Mar. Sci. 2 (2010) 199–229.
5. D. Gilbert, N.N. Rabalais, R.J. Diaz, J. Zhang, Biogeosci. Discuss. 6 (2009) 9127–9160.
6. J.P.A. Hobbs, C. A. McDonald, Journal Fish Biol. 77 (2010) 1219 – 1229.

g_winkler_report_280613.pdf 1.95 MB

13.3. Coulometric KF titration

Active ingredients in pharmaceuticals, carbon-fiber composites, polymers, novel cellulose-based active paper, food powders, biomass – all of these and many other solid materials are highly affected by moisture when processing into various products. Errors and inconsistencies in moisture measurement and control in industrial processes lead to decreased process speed/throughput and increased wastage, shortened durability of biomaterials, increased energy consumption in drying and increased fine particle emissions in biomass combustion.

Coulometric Karl Fischer method is currently the most accurate method of moisture measurement. A **Survey of the factors determining the uncertainty of coulometric Karl Fischer titration method** has been carried out.

This survey gives an overview of the factors that determine the uncertainty of coulometric Karl Fischer (cKF) method for water determination. Distinction is made between uncertainty sources originating from the cKF method itself and uncertainty sources due to sample handling. The "compound" uncertainty sources – repeatability, reproducibility and bias that actually incorporate the contributions from these two classes of uncertainty sources – are also briefly discussed.

Based on the literature data the most influential uncertainty sources of coulometric KF titration method are possible chemical interferences, instrument instability and the accuracy of the end point determination. The uncertainty sources due to sample handling are more problematic with solid samples than with liquid samples. The most important sample preparation related uncertainty sources are change of water content in the sample before the measurement and incomplete transfer of water from the sample to the reaction vessel.

The general conclusion is that although the uncertainty sources of the cKF method are in general rather well known the discussion is almost always either qualitative or limited to the compound uncertainty sources and there is very limited quantitative information available on the contributions of the actual uncertainty sources.

Preparation of this report was supported by the European Metrology Research Programme (EMRP), project SIB64 METefnet "Metrology for Moisture in Materials".

EMRP
European Metrology Research Programme
Programme of EURAMET

The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union

coulometric_kf_titration_measurement_uncertainty_sources_survey.pdf 227 KB

14. Tests and Exercises

This section contains a compilation of all the tests and exercises of this course.

- The concept of measurement uncertainty (MU) - Self-test 1
- The origin of measurement uncertainty - Self-test 2
- The Normal distribution - Self-test 3.1
- Mean, standard deviation and standard uncertainty - Self-test 3.2
- Standard deviation of the mean - Self-test 3.4
- Rectangular and triangular distribution - Self-test 3.5
- The Student distribution - Self-test 3.6
- Quantifying uncertainty components - Self-test 4.1
- Calculating the combined standard uncertainty - Self-test 4.2
- Looking at the obtained uncertainty - Self-test 4.3
- Expanded uncertainty - Self-test 4.4
- Presenting measurement results - Self-test 4.5
- Sources of uncertainty - Self-test 5.3
- Treatment of random and systematic effects - Self-test 5.4
- Random and systematic effects revisited - Self-test 6
- Precision, trueness, accuracy - Self-test 7
- Overview of measurement uncertainty estimation approaches - Self-test 8
- Model equation - Self-test 9.2
- Standard uncertainties of the input quantities - Self-test 9.5
- The ISO GUM Modeling approach - Self-test 9 A and Self-test 9 B
- Uncertainty component accounting for random effects - Self-test 10.2
- Uncertainty component accounting for systematic effects - Self-test 10.3
- Determination of acrylamide in snacks by LC-MS - Self-test 10.5 A and Self-test 10.5 B
- Comparison of the approaches - Self-test 11

Frequently asked questions

How many decimal places should we leave after comma when presenting results?

The number of decimals after the comma depends on the order of magnitude of the result and can be very different. It is more appropriate to ask, how many significant digits should be in the uncertainty estimate. This is explained in the video in section 4.5. The number of decimals according to that video is OK for the results, unless there are specific instructions given how many decimals after the point should be presented. When presenting result together with its uncertainty then the number of decimals in the result and in uncertainty must be the same.

If we need to find standard deviation of those wihtin-lab reproducibility measurments, then we need certainly use the pooled one? We can not take the simpliest standard deviation, which is calculated by standard deviationformula?

The within-lab reproducibility standard deviation sRW characterises how well can the measurement procedure reproduce the same results on different days with the same sample. If the sample is not the same (as in this self-test) then if you just calculate the standard deviation of the results then the obtianed standard deviation includes both the reproducibility of the procedure and also the difference between the samples. The difference between the samples is in the case of this self-test much larger than the within-lab reproducibility. So, if you simply calculate the standard deviation over all the results then you will not obtain within-lab reproducibility but rather the variability of analyte concentrations in samples, whith a (small) within-lab reproducibility component added.

In estimation of uncertainty via the modelling approach: When when we can use the Kragten approach and when we just use the combination of uncertainties?

In principle, you can always use the Kragten approach. However, if the relative uncertainties of the input quantities are large, and especially if such a quantity happens to be in the denominator, then the uncertainty found with the Kragten approach can differ from that found using equation 4.11. This is because the Kragten approach is an approximative approach.

Exactly what is human factor? I thought that it may be for example person’s psychological conditions and personal experience and so on? This will definitely influence measurement, but is this taken into account then?

The "human factor" is not a strict term. It collectively refers to different sources of uncertainty that are due to the person performing the analysis. These uncertainty sources can either cause random variation of the results or systematic shift (bias). In the table below are some examples, what uncertainty sources can be caused by the "human factor". In correct measurement uncertainty estimation the “human factor” will be automatically taken into account if the respective uncertainty sources are taken into account.

Uncertainty source	Type	Taken into account by
Variability of filling a volumetric flask to the mark, variability of filling the pipette to the mark	Random	Repeatability of filling the flask/pipetting
Systematically titrating until indicator is very strongly coloured	Systematic (causes systematically higher titration results)	Uncertainty of the titration end-point determination

Systematically grinding the sample for shorter time than should be done, leading to less dispersed sample and lowered recovery	Systematic	Uncertainty due to sample preparation (uncertainty due to recovery)

Can we report as V = (10.006 ± 0.016) mL at 95 % CL at coverage factor of 2?

We use in this course the conventional rounding rules for uncertainty. Therefore uncertainty ±0.0154 ml is rounded to ±0.015 ml. Sometimes it is recommended to round uncertainties only upwards (leading in this case to ±0.016 ml). However, in the graded test quizes please use the conventional rounding rules.

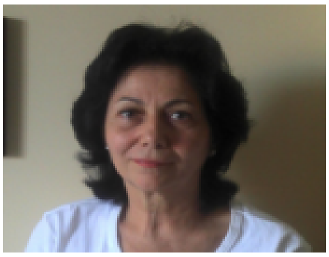
What our participants say?

Vasileios Ragias,
Greece
April 2015

"I have five years experience in accreditation and Q.A. in my lab, and though participated in many seminars, it is the first time I fully understood uncertainty. The course was over my expectations and the Tutors overdid themselves for our benefit. I am so pleased that I already started spreading the word to my colleagues!"

"I would like to thank the organisers for a very well organised and executed programme for Measurement Uncertainty in Chemical Analysis. I will definitely use this in trying to elaborate the concept to laboratories here in Trinidad and Tobago. As the Manager of the Trinidad and Tobago Laboratory Accreditation Service, this has truly enhanced my knowledge and application of the concept to real situations."

Karlene Lewis
Trinidad and Tobago
April 2015

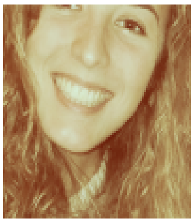


Valya Ruseva
Bulgaria
May 2015

"I would like to congratulate the Course team for a job well done and for success of this course. Indeed it was the best course I have participated and it was a privilege to learn such pragmatic approaches to measurement uncertainty estimation. Thank you very much for sharing knowledge and for so positive readiness to teach and help! "

"I am not surprised at all it was elected the "e-course of the year" in Estonia. The division of the course into sections, video lectures, reading materials and self-tests were excellent. The tutors did not answer quizzes for students. They helped students resolve problems on their own and issues of violation of rules guiding students contribution to learning forum discussions were never overlooked but immediately dealt with tactfully. This made both tutors and students attach seriousness to the course. I have learnt many new things in this course and the organisers deserve lots of commendations. Please accept my heartfelt gratitude for a course well planned and congratulations! "

Emmanuel Ofoe Ofosu
Ghana
April 2015



Mafalda Flores, Portugal April 2014

"This course was a very pleasant surprise! Honestly, when I enrolled in the course did not have high expectations because I have complicated schedules, but since they allow students to manage their own time in the way that suits you helps. I think much of the success of this course is due to this and it's a really a big "Pro" of the course. With this course I deepened my knowledge in an area where I have expertise but where I lacked certain important points: the course was a response to these failures.

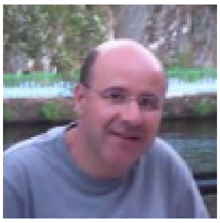
No doubt that all this information will be very useful to me in the short and long term and I thank you for your work and commitment and the whole team. It is true that knowledge is sharing, and that feeling was quite notorious even among students enrolled who were helping each other.

In short, I think the University of Tartu did an excellent job and I would be happy to attend more courses of this kind."

"First of all I would like to thank the UT team for the past six weeks, I had a blast!

I didn't know anything about measurement uncertainty when I started this course, but now such figures suddenly make sense to me. Interesting, useful and fun (believe it or not)!"

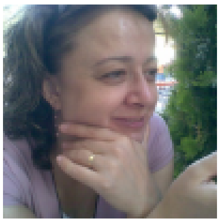
Susanne Jacobsson, Sweden April 2014



Luis Cabrita, Portugal April 2014

"I was neither an absolute beginner nor an expert in measurement uncertainty, but after this course I feel I improved a lot. The course introduced me to concepts I was not very familiar with, and clarified some questions I had in my mind for some time but had no one to ask to. The materials explained the ideas in a few simple words, and the support from teachers throughout the course was very good. It was a pleasure participating in the course and I will be looking forward to hearing of new courses (validation, etc) from the University of Tartu team.
Thank you and Congratulations!"

"It was a good pleasure for me to attend such a well organized and perfectly presented course. I have read/listened a lot on the topic before, but this was the best presentation without any frustrations and mass of calculations/formulae etc. I would like to thank to team members, University of Tartu and all the participants with their valuable questions and discussions."



Kevser Topal, Turkey April 2014

The full participant feedback from Spring 2014 is available from this link:

u_mooc_participants_feedback_2014.pdf 467 KB