Universität Stuttgart

Measurement uncertainty revisited

Alternative approaches to uncertainty evaluation

based on EUROLAB Technical Report No. 1/2007

Dr.-Ing. Michael Koch
Institute for Sanitary Engineering, Water Quality and Solid Waste Management
Universität Stuttgart
Dep. Hydrochemistry
Bandtaele 2
70569 Stuttgart GERMANY
Tel.: +49 711 685 65444 / Fax: +49 711 685 67809
e-mail: Michael.Koch@iswa.uni-stuttgart.de

Universität Stuttgart

GUM – Guide to the expression of uncertainty in measurement

- acknowledged as the master document of measurement uncertainty
- Main GUM principles:
  - uncertainty evaluation is comprehensive, accounting for all relevant sources of measurement error
  - uncertainties arising from random and systematic effects are treated alike, i.e. are expressed and combined as variances of associated probability distributions
  - statistical evaluation of measurements (Type A) and alternative techniques, based on other data / information (Type B), are recognised and utilised as equally valid tools
  - uncertainties of final results are expressed as standard deviations (standard uncertainty) or by multiples of standard deviations (expanded uncertainty) with a specified numerical factor (coverage factor).

Why is the GUM often criticised as inapplicable?

- the GUM almost exclusively treats a single approach for uncertainty evaluation: the "**modelling approach**", based on a comprehensive mathematical model of the measurement procedure, where every uncertainty contribution is associated with a dedicated input quantity, the uncertainty contributions are evaluated individually and combined as variances.
- This is often (mis)conceived as being "the GUM approach" for uncertainty evaluation

Other approaches

- the GUM principles admit a variety of approaches, but this fact was buried under a plethora of papers and lectures celebrating the "modelling approach" as a new paradigm in measurement quality assurance.
- Alternative "empirical approaches" have only recently received greater attention.
- Data utilised in these approaches are typically precision and bias data obtained from within-laboratory validation studies, quality control, interlaboratory method validation studies, or proficiency tests

Are those alternative approaches GUM-conform?

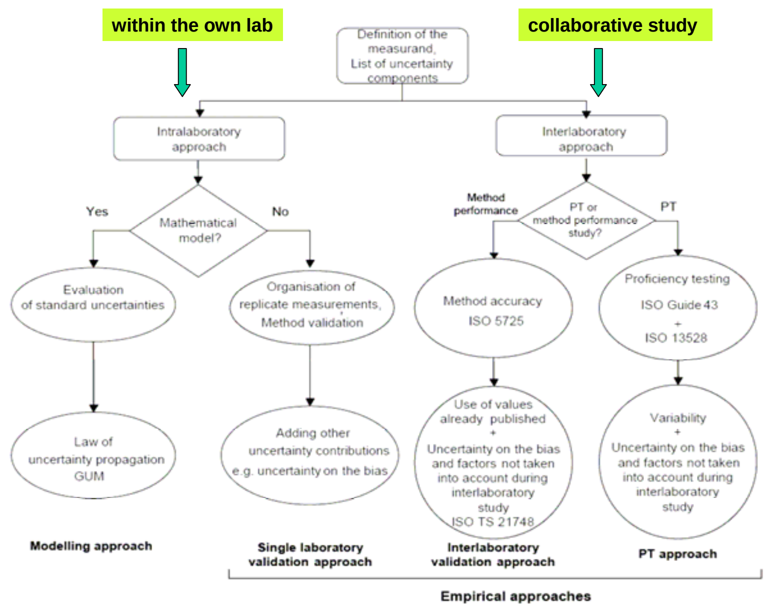
- Yes, if the GUM principles are observed
  - a clear definition of the measurand, i.e. the quantity to be measured
  - a comprehensive specification of the measurement procedure and the test items, and
  - a comprehensive analysis of the effects impacting the measurement results.

Empirical approaches

- use of reproducibility standard deviation from an interlaboratory method validation study
- use of within-laboratory data (data from method validation studies and quality control carried out in the lab)
- use of laboratory performance data from PT

Uncertainty evaluation is a difficult task, prone to mistakes

- Measurement uncertainty is often significantly underestimated
  - In the modelling approach e.g. major uncertainty contributions may be lacking, input uncertainties may be misestimated, and correlations may be overlooked
  - In the empirical approach, significant effects which have not been included in the experimental design for the method performance investigation, e.g. variations of test items or test conditions, will be missing in a (collaborative or within-laboratory) reproducibility standard deviation



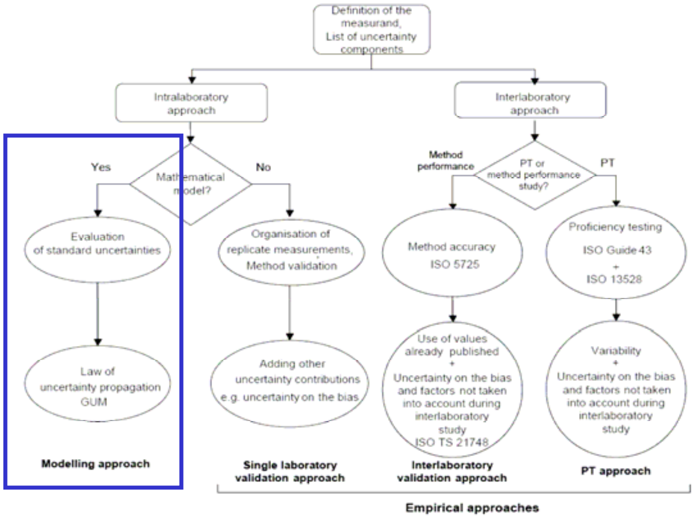
Common points between the different approaches

- always important
  - Define clearly, with no ambiguity the measurand or the characteristic to be measured, analysed or tested
  - Analyse the measuring or testing process carefully in order to identify the major components of uncertainty and to examine if they are taken on board in the application of the law of propagation of uncertainty or if they are active during the repetition of observations organised to evaluate repeatability and reproducibility or if they are included in collaborative studies.
  - It is also important to admit that in some situations, it is not possible to identify the individual components of the uncertainty. The symptom of this can be seen when the uncertainty evaluated by applying the modelling approach leads to a smaller uncertainty than the variation observed in laboratory intercomparisons

Sampling

- Where sampling activities are performed, it is also important to define the measurand clearly
  - do we seek information related to the test item transmitted to the laboratory for analysis or
  - do we need information concerning the batch (the sampling target)
- It is obvious that the uncertainty will be different in both cases

The modelling approach



Universität Stuttgart

The modelling approach

- based on a model formulated to account for the interrelation of all the influence quantities that significantly affect the measurand
- corrections are assumed to be included in the model to account for all recognised, significant systematic effects
- the application of the law of propagation of uncertainty enables evaluation of the combined uncertainty on the result
- the approach depends on partial derivatives for each influence quantity, so depends on an equation for the measured result

Koch, M.: Measurement uncertainty revisited – SADCMET PT Workshop 2007 Dar es Salaam

The modelling approach

- typical output of the modelling approach is an "uncertainty budget"
- for each input quantity x_i
  - the standard uncertainty $u(x_i)$ is determined
  - and the sensitivity coefficient $c_i = \partial y/\partial x_i$
  - resulting in the uncertainty contribution $u_i(y) = c_i \times u(x_i)$
- Unless correlation among input quantities has to be taken into account, the standard uncertainty u(y) is given by the root sum of squares of the uncertainty contributions u_i

$$u(y) = \sqrt{\sum u_i^2(y)}$$

Universität Stuttgart

The modelling approach

- By default in an uncertainty budget absolute uncertainties are used. Conversion to relative uncertainties is always possible but requires due care (other sensitivity coefficients)
- As an obvious benefit, an uncertainty budget provides information about the relative magnitude of the various uncertainty contributions. This information is particularly useful when planning improvements of the measurement procedure.

Koch, M.: Measurement uncertainty revisited – SADCMET PT Workshop 2007 Dar es Salaam

The modelling approach
Example: PT reference values for As

- Description of the measurand:
  We want to know the concentration of As in the final PT sample

Universität Stuttgart

The modelling approach
Example: PT reference values for As

- Description of the procedure:
  - A stock solution is prepared by dissolving a As_2O_3 (with a certain purity; difference weighing on an analytical balance) in a certain amount of analytical grade water (difference weighing on a toploader balance)
  - This stock solution is diluted by weighing a certain amount of the stock solution (difference weighing on a toploader balance) and filling up to a certain amount (also difference weighing on a toploader balance)

Koch, M.: Measurement uncertainty revisited – SADCMET PT Workshop 2007 Dar es Salaam

The modelling approach
Example: PT reference values for As

- Description of the procedure:
  - A certain amount of this dilute solution is weighed (difference weighing on a toploader balance) and diluted to the final amount (difference weighing on a bigger balance)
  - The density is gravimetrically measured with a pycnometer to calculate the concentration

Universität Stuttgart

The modelling approach
Example: PT reference values for As

- The input quantities can be derived from the mathematical model
- For all weighings of material with a density significantly different from the calibration mass pieces, a buoyancy correction has to applied (in our case all weighings of aqueous solutions)

Koch, M.: Measurement uncertainty revisited – SADCMET PT Workshop 2007 Dar es Salaam

The modelling approach
Example: PT reference values for As

- The mathematical model

$$c_{lot} = \frac{m_{As_2O_3} \cdot P \cdot F_{As/As_2O_3}}{m_{ss\_t} \cdot K} \cdot \frac{m_{ss} \cdot K}{m_{dil\_t} \cdot K} \cdot \frac{m_{dil} \cdot K}{m_{lot} \cdot K} \cdot \rho_{lot}$$

$$= \frac{m_{As_2O_3} \cdot F_{As/As_2O_3} \cdot P \cdot m_{ss} \cdot \rho_{lot} \cdot m_{dil}}{m_{ss\_t} \cdot m_{lot} \cdot K \cdot m_{dil\_t}}$$

m_{As2O3} = mass of arsenic oxide in stock solution in g
P = purity
$F_{As/As2O3}$ = quotient of molecular masses
m_{ss_t} = total mass of stock solution in g
K = buoyancy correction factor
m_{ss} = mass of stock solution in the diluted solution in g
m_{dil_t} = total mass of diluted solution
m_{dil} = mass of diluted solution in the final lot
m_{lot} = total mass of the lot in g
ρ_{lot} = density of the lot in g/l

The modelling approach
Example: PT reference values for As

- Identifying the sources of uncertainties
  - for all weighings
    - precision of the weighing
    - trueness of the balance (linearity)
    - uncertainty of the buoyancy correction factor
  - the purity of the chemical
  - the molecular masses of As and O
  - density measurement

The modelling approach
Example: PT reference values for As

- All uncertainty contributions have to be quantified as standard uncertainty $u(x_i)$ of the input quantity x_i
  - with type A estimation (statistical information)
  - or type B estimation (all other informations)

The modelling approach
Example: PT reference values for As

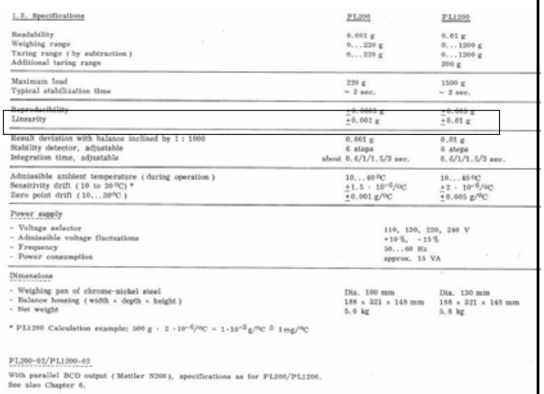
- quantifying the precision of weighings
  - modelling experiments using approximately the same masses as during sample preparation
  - 20 difference weighings → standard deviation = standard uncertainty

40g + 2g

tare	total	difference
40,0029	42,0029	2,0000
40,0027	42,0027	2,0000
40,0026	42,0028	2,0002
40,0026	42,0028	2,0002
40,0027	42,0027	2,0000
40,0026	42,0027	2,0001
40,0026	42,0026	2,0000
40,0025	42,0026	2,0001
40,0025	42,0026	2,0001
40,0025	42,0026	2,0001
40,0024	42,0026	2,0002
40,0024	42,0026	2,0002
40,0024	42,0026	2,0002
40,0024	42,0026	2,0002
40,0024	42,0026	2,0002
40,0024	42,0026	2,0002
40,0024	42,0026	2,0002
40,0024	42,0025	2,0001
40,0024	42,0025	2,0001
40,0024	42,0025	2,0001
	mean	2,0001
	std	7,86398E-05
	rstd	0,004%

The modelling approach
Example: PT reference values for As

- quantifying the trueness of weighings
  - the manufacturer allows for a certain tolerance in the linearity of the balance
  - this tolerance is taken as rectangular distribution
  - $\rightarrow s = a/\sqrt{3}$



1.2. Specifications	PL200	PL1200
Readability	0.001 g	0.01 g
Weighing range	0...220 g	0...1200 g
Taring range (by subtraction)	0...220 g	0...1200 g
Additional taring range		300 g
Maximum load	220 g	1500 g
Typical stabilization time	~ 2 sec.	~ 2 sec.
Reproducibility	±0.0005 g	±0.005 g
Linearity	±0.001 g	±0.01 g
Result deviation with balance inclined by 1 : 1000	0.001 g	0.01 g
Stability detector, adjustable	6 steps	6 steps
Integration time, adjustable	about 0.6/1/1.5/3 sec.	0.6/1/1.5/3 sec.
Admissible ambient temperature (during operation)	10...40 °C	10...40 °C
Sensitivity drift (10 to 30 °C) *	±1.5 · 10⁻⁶/°C	±2 · 10⁻⁶/°C
Zero point drift (10...30 °C)	±0.001 g/°C	±0.005 g/°C
Power supply		
- Voltage selector	110, 130, 220, 240 V	
- Admissible voltage fluctuations	+10 %, -15 %	
- Frequency	50...60 Hz	
- Power consumption	approx. 15 VA	
Dimensions		
- Weighing pan of chrome-nickel steel	Dia. 100 mm	Dia. 130 mm
- Balance housing (width × depth × height)	188 × 321 × 148 mm	188 × 321 × 148 mm
- Net weight	5.6 kg	5.8 kg

* PL1200 Calculation example: 500 g · 2 · 10^{-6}/°C = 1 · 10^{-3} g/°C ≙ 1 mg/°C

PL200-02/PL1200-02

With parallel BCD output (Mettler N200), specifications as for PL200/PL1200. See also Chapter 6.

The modelling approach
Example: PT reference values for As

- uncertainty of the balance
  - since precision and trueness are additive the sensitivity coefficients $c_i = \partial y/\partial x_i = 1$
  - with that $u_{balance} = \sqrt{u^2_{precision} + 2 \cdot u^2_{trueness}}$

Uncertainty for	200g + 500g					
parameter	*specification*	*probability distribution*	*divisor*	*standard uncertainty*	*sensitivity coefficient*	*uncertainty contribution*
precision	0,065211881	normal	1	0,06521188	1	0,065211881
trueness (lin)	0,01	rectangular	√3	0,0057735	1	0,005773503
trueness (lin)	0,01	rectangular	√3	0,0057735	1	0,005773503
					u_c	**0,065721047**

The modelling approach
Example: PT reference values for As

- buoyancy correction

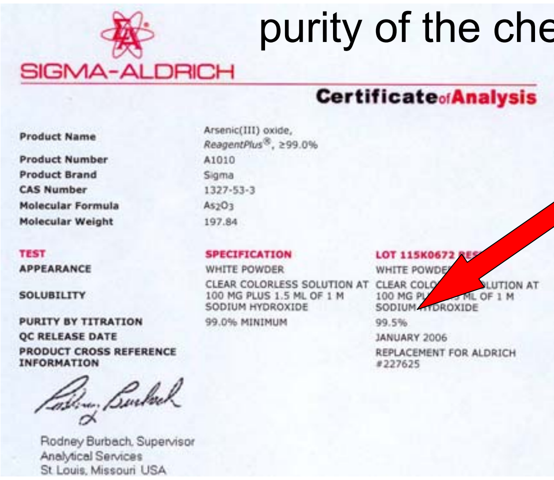
$$K_i = \frac{1 - \frac{\rho_{air}}{\rho_{cal}}}{1 - \frac{\rho_{air}}{\rho_i}}$$

with ρ_{air} = 1,1788 g/l (average air density) and
ρ_{cal} = 8000 g/l (approximate density of the metallic calibration mass pieces) and
ρ_i = 1001 g/l (approximate density of an aqueous solution
we get K = 1.00103

the uncertainty can be estimated from possible variations in the lab environment
from O. Rienitz (PTB) PhD Thesis: u_K = 0.00011

The modelling approach
Example: PT reference values for As

purity of the chemical


SIGMA-ALDRICH
Certificate of Analysis

Product Name	Arsenic(III) oxide, ReagentPlus®, ≥99.0%	
Product Number	A1010	
Product Brand	Sigma	
CAS Number	1327-53-3	
Molecular Formula	As_2O_3	
Molecular Weight	197.84	
TEST	SPECIFICATION	LOT 115K0672 RES
APPEARANCE	WHITE POWDER	WHITE POWDE
SOLUBILITY	CLEAR COLORLESS SOLUTION AT 100 MG PLUS 1.5 ML OF 1 M SODIUM HYDROXIDE	CLEAR COLO LUTION AT 100 MG PL ML OF 1 M SODIUM DROXIDE
PURITY BY TITRATION	99.0% MINIMUM	99.5%
QC RELEASE DATE		JANUARY 2006
PRODUCT CROSS REFERENCE INFORMATION		REPLACEMENT FOR ALDRICH #227625

Rodney Burbach, Supervisor
Analytical Services
St Louis, Missouri USA

99.5%

Uncertainty?
It is assumed that the manufacturer can distinguish between 99.5% and 99.6% if they report 99.5%
Therefore rectangular distribution ±0.1%

$$u_P = \frac{0.001}{\sqrt{3}} = 0.00057$$

The modelling approach
Example: PT reference values for As

- Molecular masses of As and O
  - taken from an IUPAC publication
  - uncertainty is neglected

The modelling approach
Example: PT reference values for As

- Density measurement – procedure
  - Bring the sample and a bottle of analytical grade water to the same temperature
  - Weigh the empty pycnometer
  - Fill the pycnometer with sample and weigh it
  - Fill the pycnometer with water and weigh it
- Calculation

$$\frac{\rho_{sample}}{\rho_{water}} = \frac{m_{pycn+sample} - m_{pycn}}{m_{pycn+water} - m_{pycn}} \;\Rightarrow\; \rho_{sample} = \frac{m_{pycn+sample} - m_{pycn}}{m_{pycn+water} - m_{pycn}} \cdot \rho_{water}$$

- and with buoyancy correction

$$\rho_{sample} = \frac{m_{pycn+sample} - m_{pycn}}{m_{pycn+water} - m_{pycn}} \cdot \rho_{water} \cdot \left(1 - \frac{\rho_{air}}{\rho_{water}}\right) + \rho_{air}$$

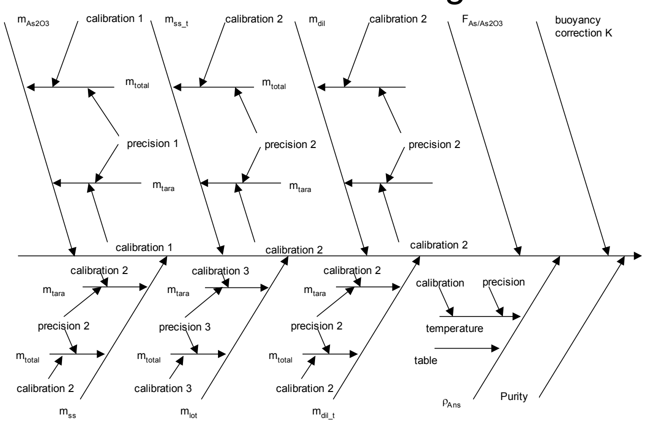
ρ_{water} taken from a PTB table for the measured temperature

The modelling approach
Example: PT reference values for As

- Density measurement uncertainty
  - own uncertainty budget
  - uncertainty sources:
    - balance – as shown above
    - table – uncertainty neglected
    - temperature measurement – uncertainty of the thermometer taken from the calibration certificate
    - density of the air – from normal variations in the lab

The modelling approach
Example: PT reference values for As

- all that uncertainty contributions can be illustrated in a fishbone diagram

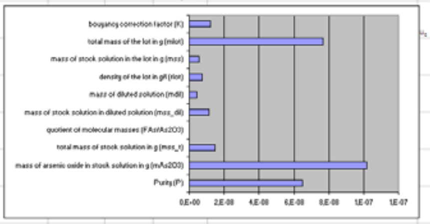


The modelling approach
Example: PT reference values for As

- For each input quantity we calculate in a spreadsheet (as shown by Angelique in 2005)
  - its standard uncertainty $u(x_i)$
  - its sensitivity coefficient $c_i = \partial y / \partial x_i$
  - its uncertainty contribution $u_i(y) = c_i \times u(x_i)$

$$c_{lot} = \frac{m_{As_2O_3} \cdot F_{As/As_2O_3} \cdot P \cdot m_{ss} \cdot \rho_{lot} \cdot m_{ss\_dil}}{m_{ss\_t} \cdot m_{lot} \cdot K \cdot m_{dil}}$$

	parameter	estimated value	specification	probability distribution	divisor	standard uncertainty (u_i)	sensitivity coefficient (c_i)	sensitivity coefficient (c_i)	uncertainty contribution ($c_i u_i$)	abs	
3	Purity (P)	99,50%	0,10%	Rechteck	$\sqrt{3}$	0,00057735	$\frac{m_{As_2O_3} \cdot F_{As/As_2O_3} \cdot m_{ss} \cdot \rho_{lot} \cdot m_{ss\_dil}}{m_{ss\_t} \cdot m_{lot} \cdot K \cdot m_{dil}}$	0,000112682	6,5057E-08	6,5057E-08	from MERCK certificate
4	mass of arsenic oxide in stock solution in g (m_{As2O3})	0,1053	1602MP 40-0,2			0,000167986	$\frac{F_{As/As_2O_3} \cdot P \cdot m_{ss} \cdot \rho_{lot} \cdot m_{ss\_dil}}{m_{ss\_t} \cdot m_{lot} \cdot K \cdot m_{dil}}$	0,000605065	1,01642E-07	1,01642E-07	uncertainty of balance 1
5	total mass of stock solution in g (m_{ss_t})	499,41	PL1200 200-500			0,065721047	$-\frac{m_{As_2O_3} \cdot F_{As/As_2O_3} \cdot P \cdot m_{ss} \cdot \rho_{lot} \cdot m_{ss\_dil}}{m_{ss\_t}^2 \cdot m_{lot} \cdot K \cdot m_{dil}}$	-2,24502E-07	-1,47545E-08	1,47545E-08	uncertainty of balance 2
6	quotient of molecular masses ($F_{As/As2O3}$)	0,75739018				0	$\frac{m_{As_2O_3} \cdot P \cdot m_{ss} \cdot \rho_{lot} \cdot m_{ss\_dil}}{m_{ss\_t} \cdot m_{lot} \cdot K \cdot m_{dil}}$	0,000148033	0	0	from IUPAC data; uncertainty negligible
7	mass of stock solution in diluted solution (m_{ss_dil})	100,00	PL1200 200-200			0,010160985	$\frac{m_{As_2O_3} \cdot F_{As/As_2O_3} \cdot P \cdot m_{ss} \cdot \rho_{lot}}{m_{ss\_t} \cdot m_{lot} \cdot K \cdot m_{dil}}$	1,12119E-06	1,13923E-08	1,13923E-08	uncertainty of balance 2
8	mass of diluted solution (m_{dil})	996,45	PL1200 200-1000			0,039310953	$-\frac{m_{As_2O_3} \cdot F_{As/As_2O_3} \cdot P \cdot m_{ss} \cdot \rho_{lot} \cdot m_{ss\_dil}}{m_{ss\_t} \cdot m_{lot} \cdot K \cdot m_{dil}^2}$	-1,12510E-07	-4,42319E-09	4,42319E-09	uncertainty of balance 2
9	density of the lot in g/l (ρ_{lot})	997,907535				0,06596093	$\frac{m_{As_2O_3} \cdot F_{As/As_2O_3} \cdot P \cdot m_{ss} \cdot m_{ss\_dil}}{m_{ss\_t} \cdot m_{lot} \cdot K \cdot m_{dil}}$	1,12354E-07	7,41095E-09	7,41095E-09	see separate calculation
10	mass of stock solution in the lot in g (m_{ss})	200	PL1200 200-200			0,010160985	$\frac{m_{As_2O_3} \cdot F_{As/As_2O_3} \cdot P \cdot \rho_{lot} \cdot m_{ss\_dil}}{m_{ss\_t} \cdot m_{lot} \cdot K \cdot m_{dil}}$	5,60593E-07	5,69617E-09	5,69617E-09	uncertainty of balance 2
11	total mass of the lot in g (m_{lot})	49900	BFW300			34,09481105	$-\frac{m_{As_2O_3} \cdot F_{As/As_2O_3} \cdot P \cdot m_{ss} \cdot \rho_{lot} \cdot m_{ss\_dil}}{m_{ss\_t} \cdot m_{lot}^2 \cdot K \cdot m_{dil}}$	-2,24686E-09	-7,66064E-08	7,66064E-08	uncertainty of balance 3
12	bouyancy correction factor (K)	1,00103149				0,00011	$-\frac{m_{As_2O_3} \cdot F_{As/As_2O_3} \cdot P \cdot m_{ss} \cdot \rho_{lot} \cdot m_{ss\_dil}}{m_{ss\_t} \cdot m_{lot} \cdot K^2 \cdot m_{dil}}$	-0,000112003	-1,23203E-08	1,23203E-08	from PTB informations
13	result (g/l)	0,00011212						u_c	1,45046E-07		
14	**result in mg/l**	**0,11211854**									
15	**standard uncertainty in mg/l**	**0,00014506**									
16	**rel. uncertainty**	**0,13%**									
17	**exp. uncertainty**	**0,00029009**									
18	**exp. rel. uncertainty**	**0,26%**									



The modelling approach
Example: PT reference values for As

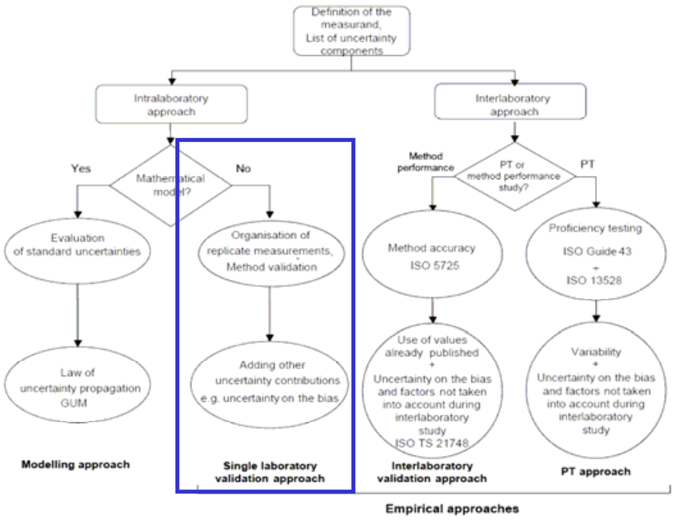
- The big advantage of the modelling approach:
  - the biggest contribution can be identified
  - in this case the weighing of the chemical

The modelling approach

- Scope of uncertainty data
  - An uncertainty budget refers to a specified measurement.
  - But the algorithm behind the uncertainty budget applies to all measurements made using the same measurement system and procedure on comparable test items.
  - For any new measurement, the (combined) standard uncertainty u(y) is obtained by plugging the input data x_i and $u(x_i)$ for this measurement into the algorithm, which then will return y and u(y).
  - Of course, if the input data are close to those for a previous measurement, the standard uncertainty u(y) will be about the same as obtained before

Universität Stuttgart

The single laboratory validation approach



Universität Stuttgart

The single laboratory validation approach

- Basic principle
  - *Measurement accuracy = precision + trueness*
  - *Measurement uncertainty =*
    within-lab reproducibility + uncertainty on the bias
- Measurement uncertainty is estimated as a root sum of squares of a standard deviation s characterising the (im)precision of the measurement and an estimate b accounting for measurement bias, which gives the standard uncertainty u according to the schematic equation

$$u = \sqrt{s^2 + b^2}$$

The single laboratory validation approach

- Bias correction
  - measurement bias is investigated, and corrective actions are taken to remove/reduce such bias to the greatest possible extent.
  - The bias-related uncertainty estimate accounts for the potential bias left after correction.
  - In practice, however, it happens quite often that significant bias is found, but the data are not sufficient for deriving a sound correction.
  - For example, it may be doubtful whether a single-level correction, based on measurements of a single standard, is applicable to the entire measuring range.
  - Then additional measurements, e.g. including another standard, should be made in order to characterise the bias to an appropriate degree. If this is not possible or not practical, a pragmatic alternative is to increase the uncertainty to account for the observed bias instead of attempting any correction

Universität Stuttgart

The single laboratory validation approach

- Data on precision
  - The precision of a measurement procedure is investigated during method validation, monitored in quality control, and quantified by standard deviations obtained from replicate measurements on appropriate test items.
  - Depending on the conditions two different standard deviations can be obtained
    - s_{rw} the within-laboratory repeatability standard deviation, obtained under repeatability conditions: same operator, same equipment, short-time repetition.
    - s_{Rw} the within-laboratory reproducibility standard deviation, obtained under within-laboratory reproducibility conditions (often called “intermediate conditions”): different operators (if applicable), different equipment (if applicable), long-time repetition.

Koch, M.: Measurement uncertainty revisited – SADCMET PT Workshop 2007 Dar es Salaam

The single laboratory validation approach

- Data on precision
  - For the purpose of estimating measurement uncertainty, the **within-laboratory reproducibility standard deviation s_{Rw}** will be used.
  - The repeatability standard deviation s_{rw} is **not** normally a suitable uncertainty estimate, since it excludes major uncertainty contributions.

The single laboratory validation approach

- Data on bias
  - It is understood that measurement bias is eliminated to the greatest possible extent.
  - Residual bias is investigated during method validation, monitored in quality control, and quantified by deviations of measurement results on appropriate test items from corresponding reference values.
  - Most often reference materials are used for this purpose, but alternatively a reference measurement procedure may be used.

The single laboratory validation approach

- The bias contribution to measurement uncertainty is obtained from the mean deviation, the uncertainty of the reference value, and the (im)precision of the mean value of the replicate measurements made in the bias investigation:

$$b = \sqrt{\Delta^2 + u_{ref}^2 + \frac{s^2}{n}}$$

The single laboratory validation approach

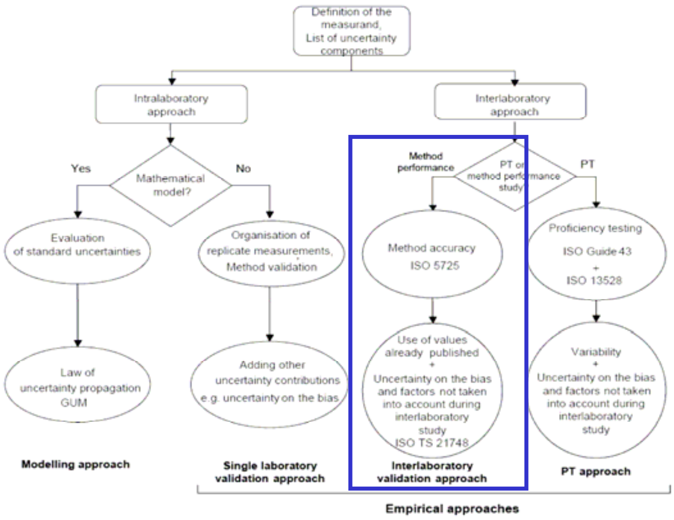
- Often different data on bias, obtained from different measurement series, will be available.
- Then these data should be compared and combined into a joint estimate for the uncertainty on bias, preferably as a function of the measurand level.
- In absence of within-laboratory bias investigations the PT approach (see later) may be used. In this case bias estimates are obtained from PT data (deviation of the laboratory´s result from the assigned value) while the within-laboratory reproducibility standard deviation is used as precision estimate.
- If bias estimates are not available at all, a pragmatic approach would be to expand the within-laboratory standard deviation using a rule-of-thumb factor. For the chemical field, e.g., average proportions between various within-laboratory and interlaboratory precision data were published.
- Considering that a factor of two is quite commonly observed in such studies, $u \approx 2\, s_{Rw}$ could be used as a preliminary estimate of measurement uncertainty in absence of bias data.

The single laboratory validation approach

- Scope of uncertainty data
  - provided that the measurements are under statistical control, uncertainty estimates obtained using this approach are applicable for all measurements within the scope of the measurement procedure.
  - The application range of the uncertainty estimates is determined by the range covered in the validation study and the on-going quality control.
  - Therefore these investigations should include appropriate within-scope variations, e.g. different levels of the measurand and different types of test items

Universität Stuttgart

The interlaboratory validation approach



Koch, M.: Measurement uncertainty revisited – SADCMET PT Workshop 2007 Dar es Salaam

The interlaboratory validation approach

- For standard test procedures, trueness and precision are usually determined by an interlaboratory comparison (see ISO 5725-2).
- The main performance characteristics obtained in such studies are
  - s_r the repeatability standard deviation
  - s_R the interlaboratory reproducibility standard deviation
- For the purpose of estimating measurement uncertainty, the reproducibility standard deviation s_R will be used.
- The repeatability standard deviation s_r is **not** normally a suitable uncertainty estimate, since it excludes major uncertainty contributions.

Universität Stuttgart

The interlaboratory validation approach

- Bias
  - When suitable reference test objects are available, the interlaboratory validation study may also include an investigation of bias.
  - However, since the (interlaboratory) reproducibility standard deviation already comprises systematic effects due to different ways of operation in the laboratories involved (laboratory bias), such study will only address method bias.
  - Most often method bias is not significant or not relevant and is not specified as a separate performance characteristic.

The interlaboratory validation approach

- Estimation of uncertainty
  - the default uncertainty estimate from an interlaboratory validation study is, as a standard uncertainty u:

$$u = s_R$$

Universität Stuttgart

The interlaboratory validation approach

- According to *ISO/TS 21748 Guide to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation* this estimation may be applied if the laboratory can prove
  - that the tests are carried out in conformity with the standard, and in particular
  - that the measuring conditions and test items are consistent with those in the interlaboratory comparison, and
  - that for its implementation of the test procedure, trueness and precision are compatible with the inter-laboratory comparison data

Koch, M.: Measurement uncertainty revisited – SADCMET PT Workshop 2007 Dar es Salaam

The interlaboratory validation approach

- Scope of uncertainty data
  - Provided that the measurements are under statistical control, the reproducibility standard deviation s_R is applicable for all measurements within the scope of the standard procedure.
  - For out-of scope applications, i.e. if the test conditions or the test objects substantially deviate from those in the interlaboratory validation study, the effect of these deviations has to be estimated and combined with the reproducibility standard deviation.
  - For this purpose the following schematic equation applies:

$$u = \sqrt{s_R^2 + \sum u_{other}^2}$$

Approach using PT data

- The use of PT data for estimating measurement uncertainty is still under debate and authoritative references are few
- But if a laboratory has successfully participated in an inter-laboratory proficiency test, it may also utilise the results for estimating the measurement uncertainty for the measurement procedure used

Approach using PT data

- PT data normally deliver
  - a reproducibility standard deviation s_R
  - the laboratory's deviation Δ from the assigned value
  - an uncertainty estimate u_{ass} for the assigned value should also be available

Universität Stuttgart

Approach using PT data

- Similar to the single laboratory validation approach the uncertainty could be estimated according to $u^2 = s^2 + b^2$, where
  - precision s could be derived from within-laboratory standard deviation (e.g. control charts)
  - and bias from the deviation Δ in the PT according to the formula

$$b = \sqrt{\Delta^2 + u_{ass}^2 + \frac{s^2}{n}}$$

Koch, M.: Measurement uncertainty revisited – SADCMET PT Workshop 2007 Dar es Salaam

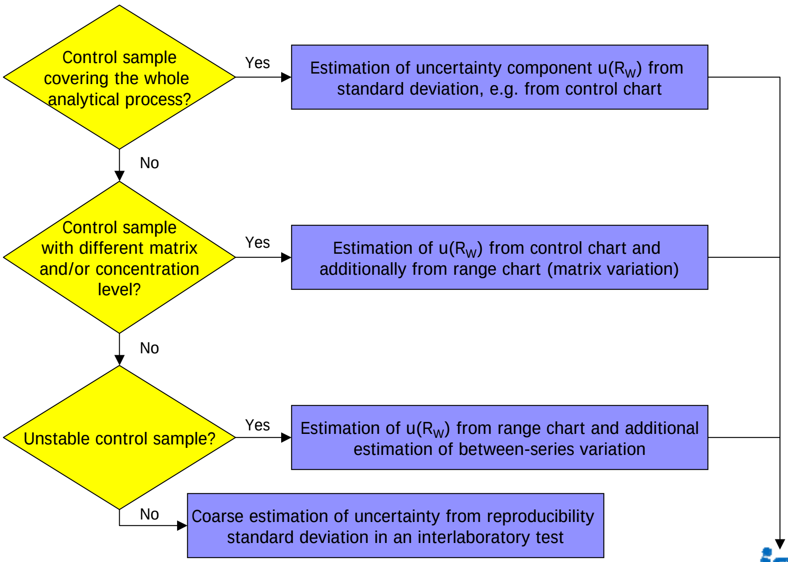
Approach using PT data

- Correction for bias
  - The bias estimate from PT studies should not normally be used for any correction of the results.
  - If the observed bias is regarded as unacceptable the laboratory has to take action and resolve this issue.

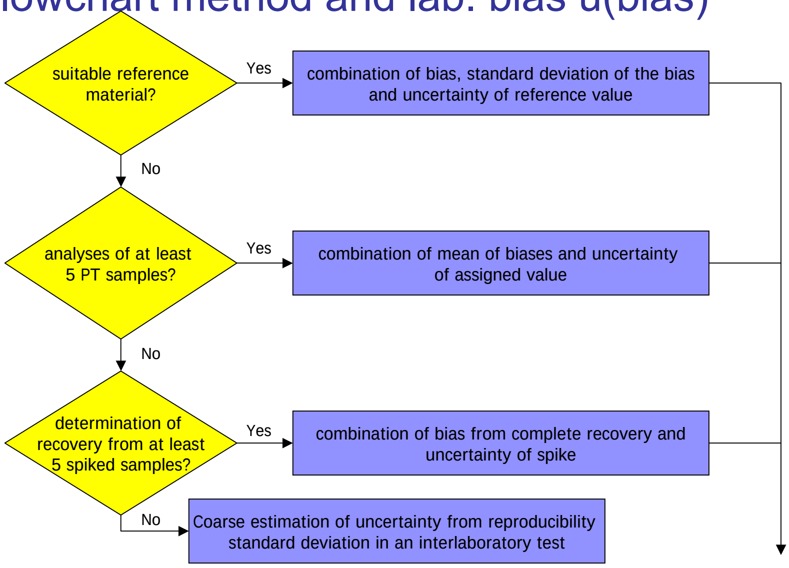
NORDTEST approach

- An approach using a combination of single laboratory validation, interlaboatory validation and PT data is described in the NORDTEST „Handbook for calculation of measurement uncertainty in environmental laboratories“ and in a German Guideline for estimating measurement uncertainty based on validation data

Flowchart reproducibility $u(R_W)$



Flowchart method and lab. bias u(bias)



Reproducibility within-laboratory

- quantification of random variations has to be done under the same conditions as in routine analysis
- i.e.:
  - neither under repeatability conditions
  - nor under reproducibility conditions
  - but under between-series conditions
- this is called here „reproducibility within-laboratory“

Reproducibility within the laboratory R_w - method 1

Control sample covering the whole analytical process

- if
  - the control sample covers the whole analytical process and
  - has a matrix similar to the samples,
- the within-laboratory reproducibility at that concentration level can simply be estimated from the analyses of the control sample
- If the analyses performed cover a wide range of concentration levels, also control samples of other concentration levels should be used.

		value	rel. Uncertainty	Comments
Reproducibility within the lab R_w				
control sample 1 $\overline{X}$ = 20.01 µg/l	s_{Rw}	standard deviation 0.5 µg/l	2.5 %	from 75 measure-ments in 2002
control sample 1 $\overline{X}$ = 250.3 µg/l	s_{Rw}	standard deviation 3.7 µg/l	1.5 %	from 50 measure-ments in 2002
other components		---		

Reproducibility within the laboratory R_w – method 2

Control samples for different matrices and concentrations

- if
  - a synthetic control solution is used for quality control, and
  - the matrix type of the control sample is **not** similar to the natural samples
- we have to take into consideration uncertainties arising from different matrices
- These can be estimated from the repeatability with different matrices (range control chart)

		value	u(x)	Comments
Reproducibility within the lab R_w				
low level (2-15 µg/l)	s_{Rw}	0.5 µg/l from the mean control chart 0.37 µg/l from the range control chart	0.6 µg/l	Absolute: $u(x)=\sqrt{0.5^2+0.37^2}$
high level (>15 µg/l)	s_{Rw}	1.5 % from the mean control chart 3.6 % from the range control chart	3.9 %	Relative: $u(x)=\sqrt{1.5\%^2+3.6\%^2}$

Note: The repeatability component is included two times!!

Reproducibility within the laboratory R_w – method 3

Unstable control samples

- if
  - the laboratory does not have access to stable control samples (e.g. measurement of dissolved oxygen)
- it is possible only to estimate uncertainty components from repeatability via the range control chart
- the „long-term“ uncertainty component (from batch to batch) has to be estimated e.g. by a qualified guess

		value	u(x)	Comments
Reproducibility within the laboratory R_w				
Duplicate measurements of natural samples	s_r	s = 0.024 mg/l mean: 7.53 mg/l	0.32 %	from 50 measurements
Estimated variation from differences in calibration over time		s = 0.5 %	0.5 %	based on experience
Combined uncertainty for R_w Repeatability + Reproducibility in calibration		$\sqrt{0.32\%^2+0.5\%^2}=0.59\%$		

Method and Laboratory bias

- can be estimated from
  - the analysis of certified reference materials
  - the participation in proficiency tests
  - from recovery experiments
- Sources of bias should always be eliminated if possible
- According to GUM a measurement result should always be corrected if the bias is significant and based on reliable data such as analysis of a CRM.
- In many cases the bias can vary depending on changes in matrix. This can be reflected when analysing several matrix CRMs

Method and Laboratory bias u(bias)
Components of uncertainty

- the bias (as % difference from the nominal or certified value)
- the uncertainty of the bias determination
- the uncertainty of the nominal/certified value $u(C_{ref})$

Method and Laboratory bias u(bias) - method 1a
Use of one certified reference material

- The reference material should be analysed in at least 5 different analytical series
- Example: Certified value: 11.5 ± 0.5 (95% confidence interval)

Uncertainty component from the uncertainty of the certified value	
Convert the confidence interval	The confidence interval is ± 0.5. Divide this by 1.96 to convert it to standard uncertainty: 0.5/1.96=0.26
Convert to relative uncertainty $u(C_{ref})$	100·(0.26/11.5)=2.21%

Method and Laboratory bias u(bias) - method 1a
Use of one certified reference material

- Quantify the bias
  - the CRM was analysed 12 times. The mean is 11.9 with a standard deviation of 2.2%
  - This results in:

$$bias = 100 \cdot (11.9 - 11.5)/11.5 = 3.48\% \quad and$$

$$s_{bias} = 2.2\% \quad with \quad n = 12$$

- Therefore the standard uncertainty is:

$$u(bias) = \sqrt{(bias)^2 + \left(\frac{s_{bias}}{\sqrt{n}}\right)^2 + u(C_{ref})^2} =$$

$$\sqrt{(3.48\%)^2 + \left(\frac{2.2\%}{\sqrt{12}}\right)^2 + 2.21\%^2} = 4.2\%$$

Method and Laboratory bias u(bias) - method 1b
Use of several certified reference material

- Quantification of the bias
  - bias CRM1 is 3.48%, s=2.2% (n=12), $u(C_{ref})$=2.21%
  - bias CRM2 is –0.9%, s=2.0% (n=7), $u(C_{ref})$=1.8%
  - bias CRM3 is 2.4%, s=2.8% (n=10), $u(C_{ref})$=1.8%
  - RMS_{bias} then is:

$$RMS_{bias} = \sqrt{\frac{\sum (bias_i)^2}{n}} = \sqrt{\frac{3.48\%^2 + (-0.9\%)^2 + 2.4\%^2}{3}} = 2.5\%$$

  - and the mean uncertainty of the certified value $u(C_{ref})$: 1.9%
- This results in the total standard uncertainty of the bias:

$$u(bias) = \sqrt{RMS_{bias}^2 + u(C_{ref})^2} = \sqrt{2.5\%^2 + 1.9\%^2} = 3.1\%$$

Method and Laboratory bias u(bias) – method 2
Use of PT results

- In order to have a reasonably clear picture of the bias from interlaboratory comparison results, a laboratory should participate at least 6 times within a reasonable time interval

Uncertainty component from the uncertainty of the nominal value	
between laboratory standard deviations s_R	s_R has been on average 9% in the 6 exercises
Convert to relative uncertainty $u(C_{ref})$	Mean number of participants= 12 $u(C_{ref}) = \frac{s_R}{\sqrt{n}} = \frac{9\%}{\sqrt{12}} = 2.6\%$

Or: $u(C_{ref}) = 1.25 \cdot \frac{s_R}{\sqrt{n}}$ for a robust mean to be in accordance with ISO 13528

Universität Stuttgart

Method and Laboratory bias u(bias) – method 2
Use of PT results

- Quantification of the bias
  - In the 6 participations the biases have been: 2%, 7%, -2%, 3%, 6% and 5%
  - Therefore RMS_{bias} is:

$$RMS_{bias} = \sqrt{\frac{\sum (bias_i)^2}{n}} = \sqrt{\frac{2\%^2 + 7\%^2 + (-2\%)^2 + 3\%^2 + 6\%^2 + 5\%^2}{6}} = 4.6\%$$

- and the total standard uncertainty of the bias:

$$u(bias) = \sqrt{RMS_{bias}^2 + u(C_{ref})^2} = \sqrt{4.6\%^2 + 2.6\%^2} = 5.3\%$$

iswa

Universität Stuttgart

Method and Laboratory bias u(bias) – method 3
From Recovery Tests

- Recovery tests, for example the recovery of a standard addition to a sample in the validation process, can be used to estimate the systematic error. In this way, validation data can provide a valuable input to the estimation of the uncertainty.
- Example: In an experiment the recoveries for an added spike were 95 %, 98 %, 97 %, 96 %, 99 % and 96 % for 6 different sample matrices. The spike of 0.5 mL was added with a micropipette.

uncertainty component from spiking	
uncertainty of the concentration of the spike u(conc)	from the certificate: 95% confidence intervall = ± 1.2 % u(conc) = 0.6 %
uncertainty of the added volume u(vol)	from the manufacturer of the micro pipette: max. bias: 1% (rectangular interval), repeatability: max. 0.5% (standard dev.) $u(vol) = \sqrt{\left(\frac{1\%}{\sqrt{3}}\right)^2 + 0.5\%^2} = 0.76\%$
uncertainty of the spike $u(c_{recovery})$	$\sqrt{u(conc)^2 + u(vol)^2} = \sqrt{0.6\%^2 + 0.76\%^2} = 1.0\%$

iswa

Method and Laboratory bias u(bias) – method 3
From Recovery Tests

- Quantification of the bias:
- RMS_{bias}:

$$RMS_{bias} = \sqrt{\frac{5\%^2 + 2\%^2 + 3\%^2 + 4\%^2 + 1\%^2 + 4\%^2}{6}} = 3.44\%$$

- Therefore the total standard uncertainty of the bias is:

$$u(bias) = \sqrt{RMS_{bias}^2 + u(C_{recovery})^2} = \sqrt{3.44\%^2 + 1.0\%^2} = 3.6\%$$

Combination of the uncertainties
(Reproducibility within-laboratory and bias)

- Reproducibility R_w (from control samples and other estimations)
- bias u(bias) (from CRM, PT or recovery tests)
- Combination:

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

Calculation of the expanded uncertainty

- for the conversion to an approx. 95% confidence level

$$U = 2 \cdot u_c$$

Coarse estimation by direct use of reproducibility standard deviations

- If the demand on uncertainty is low: $u_c = s_R$
- The expanded uncertainty becomes: $U = 2 \cdot S_R$
- This may be an overestimate depending on the quality of the laboratory – worst-case scenario
- It may also be an underestimate due to sample inhomogeneity or matrix variations

Reproducibility standard deviation from a standard

- The laboratory must first prove that it is able to perform in accordance with the standard method
  - „no“ significant bias
  - verification of the repeatability
- The expanded uncertainty then is:

$$U = 2 \cdot s_R$$

Universität Stuttgart

Reproducibility standard deviation from a standard Example – Mercury according to EN 1483

Tabelle 2: Verfahrenskenndaten (reproducibility variation coefficient → VC_R)

	Alle Laboratorien										
	Probe	l	n	NAP %	Wahrer Wert µg/l	$\bar{\bar{x}}$ µg/l	σ_R µg/l	VC_R %	σ_r µg/l	VC_r %	Wiederfindungsrate %
drinking water	A	21	62	9	0,819	0,831	0,2500	30,1	0,1310	15,8	101,5
surface water	B	20	59	13	1,474	1,459	0,3918	26,9	0,1855	12,7	99,0
waste water	C	21	68	0	5,732	5,799	1,3745	23,7	0,5746	9,9	101,2

- Expanded uncertainty for drinking water:
  $U = 2 \cdot VC_R \approx 60\ \%$

Koch, M.: Measurement uncertainty revisited – SADCMET PT Workshop 2007 Dar es Salaam

Universität Stuttgart

Reproducibility standard deviation from a PT

- The laboratory must have been successfully participating in the PT
- If the comparison covers all relevant uncertainty components and steps (matrix?)
- The expanded uncertainty then also is:

$$U = 2 \cdot s_R$$

iswa

Universität Stuttgart

Reproducibility standard deviation from a PT Example – Mercury in a Univ. Stuttgart PT

reproducibility variation coefficient

Niveau	Vorgabe [µg/l]	rob. Standarda… chung [µg/l]	…Standardab… hung [%]	Ausschlussgren… ben [µg/l]	Ausschlussgren… nten [µg/l]	Ausschlussgren… ben [%]	Ausschlussgrenze unten [%]	Anzahl Werte	außerhalb unten	außerhalb oben	außerhalb [%]
1	0,584	0,1334	22,86	0,889	0,341	52,25	-41,60	37	3	1	10,8
2	1,248	0,2256	18,09	1,748	0,830	40,07	-33,46	39	3	1	10,3
3	1,982	0,3502	17,67	2,756	1,333	39,06	-32,75	39	1	0	2,6
4	3,238	0,4726	14,60	4,263	2,352	31,65	-27,36	41	2	2	9,8
5	3,822	0,4550	11,90	4,793	2,960	25,40	-22,55	38	0	1	2,6
6	4,355	0,7704	17,69	6,057	2,927	39,10	-32,78	40	1	0	2,5
7	5,421	0,7712	14,23	7,090	3,973	30,78	-26,71	41	1	1	4,9
8	6,360	0,7361	11,57	7,928	4,963	24,65	-21,96	38	5	1	15,8
9	6,553	0,9177	14,00	8,536	4,829	30,25	-26,31	39	2	0	5,1
10	7,361	0,9965	13,54	9,508	5,486	29,16	-25,48	40	1	3	10,0
11	8,063	1,0672	13,24	10,357	6,051	28,46	-24,94	38	5	2	18,4
12	9,359	0,9854	10,53	11,444	7,481	22,29	-20,06	40	2	2	10,0
							Summe	470	26	14	8,5

- $u_c = s_R \approx 20\%$
- $U \approx 40\%$

iswa

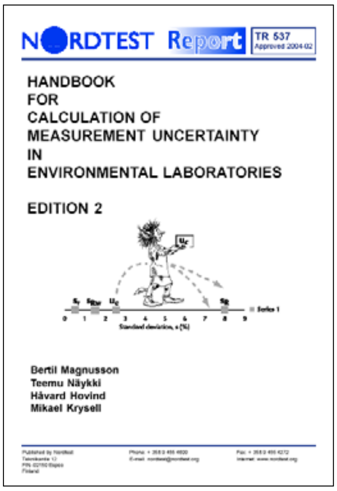
Universität Stuttgart

Conclusion

- The method described is an easy way to estimate measurement uncertainty from data that are already available in many cases
- It is a holistic approach, you cannot forget an important uncertainty source
- It does not give you information about the source of the uncertainty

Universität Stuttgart

Where to get the NORDTEST-Handbook?


NORDTEST Report TR 537 Approved 2004-02

HANDBOOK
FOR
CALCULATION OF
MEASUREMENT UNCERTAINTY
IN
ENVIRONMENTAL LABORATORIES

EDITION 2

Bertil Magnusson
Teemu Näykki
Håvard Hovind
Mikael Krysell

- The Handbook is available from http://www.nordicinnovation.net/nordtest.cfm as technical report No. 537 and on the workshop CD

Verification of measurement uncertainty estimates

- From where do we know that our estimate is realistic?
- We have to check that
- But how?

Verification of measurement uncertainty estimates

- Checks using within-laboratory precision
  - Compare the estimated standard uncertainty with the standard deviation of a series of results on an appropriate test item over a period of time
  - The standard uncertainty for a routine test method should never be smaller than the long-term precision for the same method and test material;
  - if the standard uncertainty is significantly smaller than the observed within-laboratory standard deviation, the uncertainty estimate should be reviewed immediately.

Verification of measurement uncertainty estimates

- Checks based on certified reference materials (CRM) or suitable test materials
  - Measure a suitable test material or CRM of known assigned value x_{ref} with small uncertainty.
  - Check the difference d between observed value x and reference value x_{ref} against the expanded uncertainty U(x).
  - If the difference d is equal to or greater than the expanded uncertainty U(d), it should be concluded that the uncertainty fails to account for the observed bias on the material.
  - The uncertainty estimate should be reviewed and appropriate steps taken to identify the source of the bias.

Universität Stuttgart

Verification of measurement uncertainty estimates

- Checks based on reference methods
  - Reference methods provide independent reference values.
  - A single such value can be used to check an uncertainty estimate in the same way as using a single CRM value

Koch, M.: Measurement uncertainty revisited – SADCMET PT Workshop 2007 Dar es Salaam

Verification of measurement uncertainty estimates

- Checking an uncertainty estimate against proficiency test results
  - The assessment of the uncertainty estimates is performed using the zeta score

$$\zeta = \frac{x - x_a}{\sqrt{u(x)^2 + u(x_a)^2}}$$

Verification of measurement uncertainty estimates

- Interpretation of ζ-scores
  - Uncertainty overestimated
    $|\zeta|$ always significantly < 2
    - The estimated uncertainty is clearly higher than the laboratory performance suggests.
    - This could be acceptable, especially if the reported uncertainty is lower than or equal to the target value of uncertainty (that is, within the customer's requirements).
    - However if there is a need for lower uncertainty, a new estimate has to be made.

Verification of measurement uncertainty estimates

- Interpretation of ζ-scores
  - Correct
    most values of |ζ| in the range 0 to 2
    - Here one could think that all is clear-cut, but we have to bear in mind that there are many sources that are not always tested in a PT scheme, including sampling, analyte stability, sample inhomogeneity in real samples, and other concentration levels

Universität Stuttgart

Verification of measurement uncertainty estimates

- Interpretation of ζ-scores
  - Uncertainty underestimated
    |ζ| frequently over 2
    - The estimated uncertainty is clearly lower than the laboratory is performing.
    - The uncertainty estimate should be revised to obtain a more realistic estimate

Verification of measurement uncertainty estimates

- Checks based on comparison of results with other laboratories
  - The same principles used for checks based on proficiency testing can be used for comparison with other laboratories after collaborative measurement of several test items.

Universität Stuttgart

Verification of measurement uncertainty estimates

- Comparison with other uncertainty estimates
  - When checking whether two uncertainty estimates agree or disagree, one should keep in mind that the precision of uncertainty estimates is often very limited.
  - For example, for an empirical standard deviation determined from 10 repeated measurements, the coefficient of variation is 24 %, and F-tests on two such standard deviations would not be considered significant with standard deviations differing by less than a factor of about 1.8.
  - It would therefore be unreasonable to expect different uncertainty estimates to agree very closely.

Koch, M.: Measurement uncertainty revisited – SADCMET PT Workshop 2007 Dar es Salaam

Examples and literature

- The EUROLAB technical report contains 10 detailed examples
- The report also contains a list of 27 relevant standards, guidelines, books and internet websites