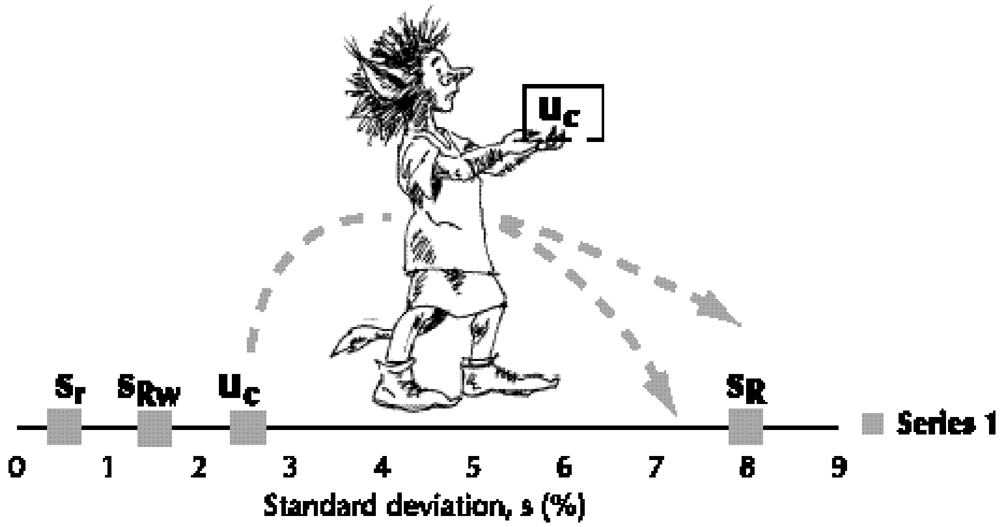
NORDTEST Report
TR 537
Approved 2003-05

HANDBOOK FOR CALCULATION OF MEASUREMENT UNCERTAINTY IN ENVIRONMENTAL LABORATORIES



Bertil Magnusson
Teemu Näykki
Håvard Hovind
Mikael Krysell

Published by Nordtest
Tekniikantie 12
FIN–02150 Espoo
Finland

Phone: + 358 9 455 4600
E-mail: nordtest@nordtest.org

Fax: + 358 9 455 4272
Internet: www.nordtest.org

NT TECHN REPORT 537
Approved 2003-05

Authors:
Bertil Magnusson[1]
Teemu Näykki[2]
Håvard Hovind[3]
Mikael Krysell[4]

NORDTEST project number: 1589-02

Institution:
[1)] SP, Sweden, [2)] SYKE, Finland, [3)] NIVA, Norway, [4)] Eurofins A/S, Denmark

Title (English):

Title (Original): Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories

Abstract:

This handbook is written for environmental testing laboratories in the Nordic countries, in order to give support to the implementation of the concept of measurement uncertainty for their routine measurements. The aim is to provide a practical, understandable and common way of measurement uncertainty calculations, mainly based on already existing quality control and validation data, according to the European accreditation guideline /12/, the Eurolab Technical Report No. 1 /3/ and the ISO/DTS 21748 Guide /8/. Nordtest has supported this project economically in order to promote and enhance harmonisation between laboratories on the Nordic market.

Practical examples, taken directly from the everyday world of environmental laboratories, are presented and explained. However, the approach is very general and should be applicable to most testing laboratories in the chemical field.

The handbook covers all steps in the analytical chain from the arrival of the sample in the laboratory until the data has been reported. It is important to notice that vital parts of the total measurement uncertainty are not included, e.g. sampling, sample transportation and possible gross errors during data storage/retrieval.
The recommendations in this document are primarily for guidance. It is recognised that while the recommendations presented do form a valid approach to the evaluation of measurement uncertainty for many purposes, other suitable approaches may also be adopted – see references in Section 9. Especially the EURACHEM/CITAC-Guide /2/ is useful in cases where sufficient previous data is not available, and therefore the mathematical analytical approach according to GUM /1/ with all different steps is to be used.

Basic knowledge in the use of quality control and statistics is required. In order to make it possible for the reader to follow the calculations, some raw data is given in appendices.

Technical Group: Expert Group Quality and Metrology

ISSN: 0283-7234

Language: English

Pages: 41

Class (UDC): 620.1

Key words: handbook, laboratories, environmental, testing, measurements, uncertainty, chemical

Distributed by:
NORDTEST
Tekniikantie 12
FIN-02150 ESPOO
Finland

Publication code:

Report Internet address:
http://www.nordtest.org/register/techn/tlibrary/tec537.pdf

Handbook
for

Calculation of Measurement Uncertainty

in

Environmental Laboratories

Version 1.2 June 2003

Nordtest project 1589-02

Project participants
Bertil Magnusson, SP, Sweden
Teemu Näykki, SYKE, Finland
Håvard Hovind, NIVA, Norway
Mikael Krysell, Eurofins A/S, Denmark

Drawings by Petter Wang, NIVA, Norway

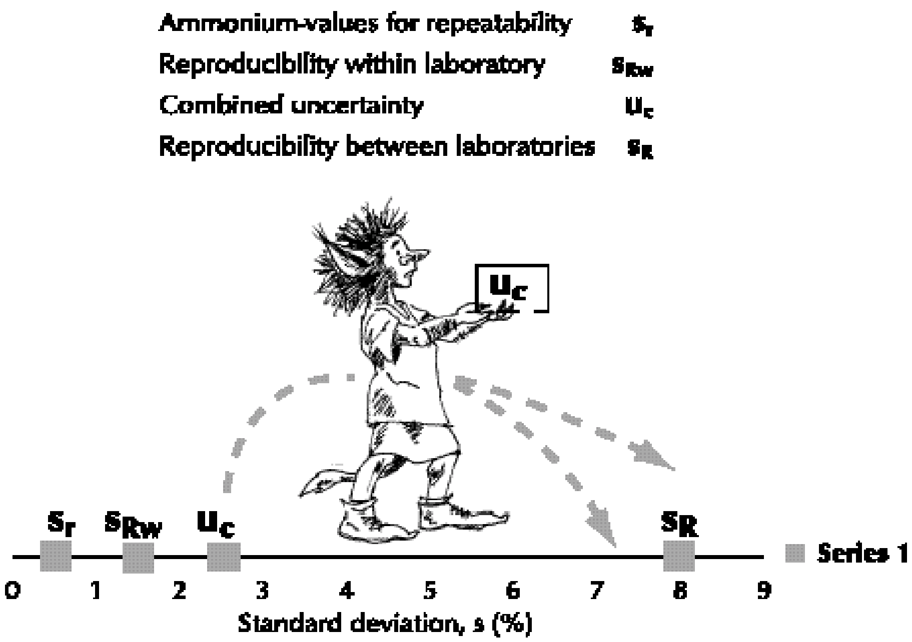
Valuable comments on the contents have been provided by:
Rolf Flykt, Sweden
Irma Mäkinen, Finland
Ulla O. Lund, Denmark
Steve Ellison, UK

Contents:

1 Definitions and abbreviations

s	An estimate of the population standard deviation σ from a limited number (n) of observations (x_i)
$\overline{x}$	Mean value
$u(x)$	Individual standard uncertainty component (GUM, /1/).
u_c	Combined standard uncertainty (GUM, /1/)
U	Expanded combined uncertainty close to 95 % confidence interval
r	Repeatability limit – performance measure for a test method or a defined procedure when the test results are obtained under **repeatability conditions**. **Repeatability conditions:** Conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time. Repeatability (precision under repeatability conditions) is also sometimes called "within run precision" (ISO 3534-1, /6/).
s_r	Repeatability standard deviation of a measurement (can be estimated from a series of duplicate analyses)
R	Reproducibility limit – performance measure for a test method or procedure when the test results are obtained under **reproducibility conditions**. **Reproducibility conditions:** Conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment. Reproducibility (precision under reproducibility conditions) is also sometimes called "between lab precision" (ISO 3534-1, /6/).
s_R	Reproducibility standard deviation of a measurement (can be estimated from validation studies with many participating laboratories or from other interlaboratory comparisons e.g. proficiency testing data) Note: $R = 2.8 \cdot s_R$
R_w	Within-laboratory reproducibility = intermediate measure between r and R, where *operator* and/or *equipment* and/or *time* and/or *calibration* can be varied, but in the same laboratory. An alternative name is intermediate precision
s_{Rw}	Reproducibility within-laboratory standard deviation (can be estimated from standard deviation of a control sample over a certain period of time, preferably one year)

CRM	Certified Reference Material
Certified value	Assigned value given to a CRM, quantified through a certification process (traceable to SI-unit and with a known uncertainty)
Nominal value	Nominal value is the assigned value, e.g. in an interlaboratory comparison where it is the organiser's best representation of the "true value"
u(Cref)	Uncertainty component from the certified or nominal value
bias	Difference between mean measured value from a large series of test results and an accepted reference value (a certified or nominal value). The measure of trueness is normally expressed in term of bias. Bias for a measurement, e.g. for a laboratory or for an analytical method..
u(bias)	Uncertainty component for bias. The *u(bias)*, is always included in the measurement uncertainty calculations
RMS_{bias}	$\sqrt{\frac{\sum (bias_i)^2}{n}}$
Interlaboratory comparison	General term for a collaborative study for either method performance, laboratory performance (proficiency testing) or material certification.



2 Introduction

2.1 Scope and field of application

This handbook is written for environmental testing laboratories in the Nordic countries, in order to give support to the implementation of the concept of measurement uncertainty for their routine measurements. The aim is to provide a practical, understandable and common way of measurement uncertainty calculations, mainly based on already existing quality control and validation data, according to the European accreditation guideline /12/, the Eurolab Technical Report No. 1 /3/ and the ISO/DTS 21748 Guide /8/. Nordtest has supported this project economically in order to promote and enhance harmonisation between laboratories on the Nordic market.

Practical examples, taken directly from the everyday world of environmental laboratories, are presented and explained. However, the approach is very general and should be applicable to most testing laboratories in the chemical field.

The handbook covers all steps in the analytical chain from the arrival of the sample in the laboratory until the data has been reported. It is important to notice that vital parts of the total measurement uncertainty are not included, e.g. sampling, sample transportation and possible gross errors during data storage/retrieval.

The recommendations in this document are primarily for guidance. It is recognised that while the recommendations presented do form a valid approach to the evaluation of measurement uncertainty for many purposes, other suitable approaches may also be adopted – see references in Section 9. Especially the EURACHEM/CITAC-Guide /2/ is useful in cases where sufficient previous data is not available, and therefore the mathematical analytical approach according to GUM /1/ with all different steps is to be used.

Basic knowledge in the use of quality control and statistics is required. In order to make it possible for the reader to follow the calculations, some raw data is given in appendices

2.2 Comment to customers

Previously, laboratories usually reported uncertainty as the standard deviation calculated from data for an internal control sample. The measurement uncertainty also taking into account method and laboratory bias and using a coverage factor of 2, can give uncertainty values which may be a factor of 2 to 5 times higher than previously (Figure 1). However, this does not reflect a change in the performance of the laboratory, just a much better estimation of the real variation between laboratories. In Figure 1, the ammonium results from two laboratories are in good agreement – the difference is about 5 %. You can see this if you look to the right where measurement uncertainty is calculated correctly, but not if you look to the left, where the uncertainty is calculated directly from a control sample and presented as the standard deviation (± 1*s*).

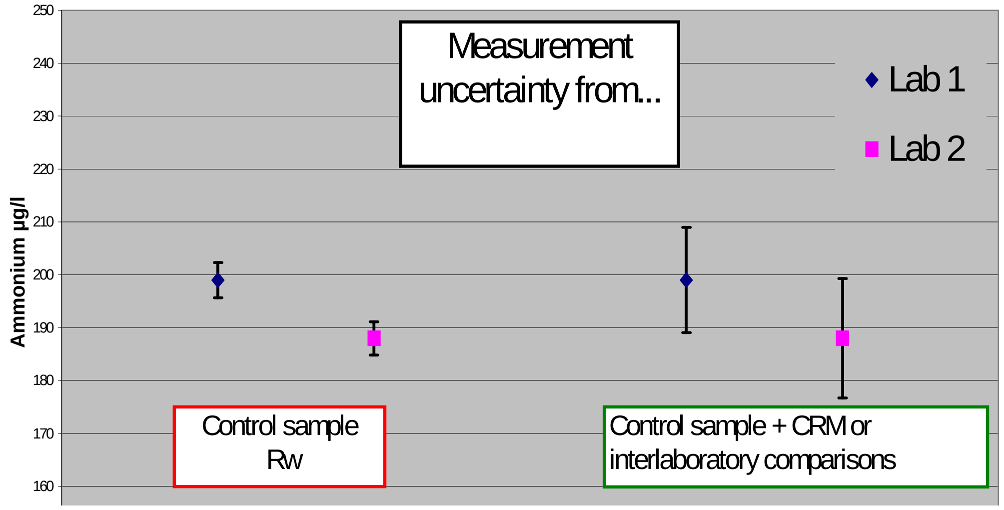


Figure 1. Comparing ammonium results from two laboratories, Lab 1 = 199 µg/L and Lab 2 = 188 µg/L. To the left the error bars are calculated from results on control samples (± 1s) and to the right the error bars are expanded measurement uncertainty.

2.3 About Measurement Uncertainty

What is measurement uncertainty?

- The number after ±
- All measurements are affected by a certain error. The measurement uncertainty tells us what size the measurement error **might** be. Therefore, the measurement uncertainty is an important part of the reported result
- Definition: Measurement uncertainty is ”A parameter associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand” /1, 5/

Who needs measurement uncertainties?

- The customer needs it together with the result to make a correct decision. The uncertainty of the result is important, e.g. when looking at allowable (legal) concentration limits
- The laboratory to know its own quality of measurement and to improve to the required quality

Why should the laboratory give measurement uncertainty?

- As explained above, the customers need it to make correct decisions
- An estimation of the measurement uncertainty is required in ISO 17025 /9/

How is measurement uncertainty obtained?

- The basis for the evaluation is a measurement and statistical approach, where the different uncertainty sources are estimated and combined into a single value
- "*Basis for the estimation of measurement uncertainty is the existing knowledge (no special scientific research should be required from the laboratories). Existing experimental data should be used (quality control charts, validation, interlaboratory comparisons, CRM etc.)*" /12/
- Guidelines are given in GUM /1/, further developed in, e.g., EA guidelines /12/, the Eurachem/Citac guide /2/, in a Eurolab technical report /3/ and in ISO/DTS 21748 /8/

How is the result expressed with measurement uncertainty?

- Measurement uncertainty should normally be expressed as *U*, the combined expanded measurement uncertainty, using a coverage factor k = 2, providing a level of confidence of approximately 95 %
- It is often useful to state how the measurement uncertainty was obtained
- Example, where ± 7 is the measurement uncertainty:

  Ammonium (NH_4-N) = 148 ± 7 µg/L. The measurement uncertainty, 7 µg/L (95 % confidence level, i.e. the coverage factor k=2) is estimated from control samples and from regular interlaboratory comparisons

How should measurement uncertainty be used?

- It can be used as in Figure 1, to decide whether there is a difference between results from different laboratories, or results from the same laboratory at different occasions (time trends etc.)
- It is necessary when comparing results to allowable values, e.g. tolerance limits or allowable (legal) concentrations

3 Calculation of expanded uncertainty, U - overview

A common way of presenting the different contributions to the total measurement uncertainty is to use a so-called fish-bone (or cause-and-effect) diagram. We propose a model (Figure 2), where either the reproducibility within-laboratory (R_w) is combined with estimates of the method and laboratory bias, (error model in Appendix 3) or the reproducibility s_R is used more or less directly, ISO Guide 21748/8/. The alternative way is to construct a detailed fish-bone diagram and calculate/estimate the individual uncertainty contributions. This approach may prove very useful when studying or quantifying individual uncertainty components. It has been shown, though, that in some cases this methodology underestimates the measurement uncertainty /3/, partly because it is hard to include all possible uncertainty contributions in such an approach. By using existing and experimentally determined quality control (QC) and method validation data, the probability of including all uncertainty contributions will be maximised.

Measurement uncertainty model – fish-bone diagram
covering the analytical process from sample arrival to report

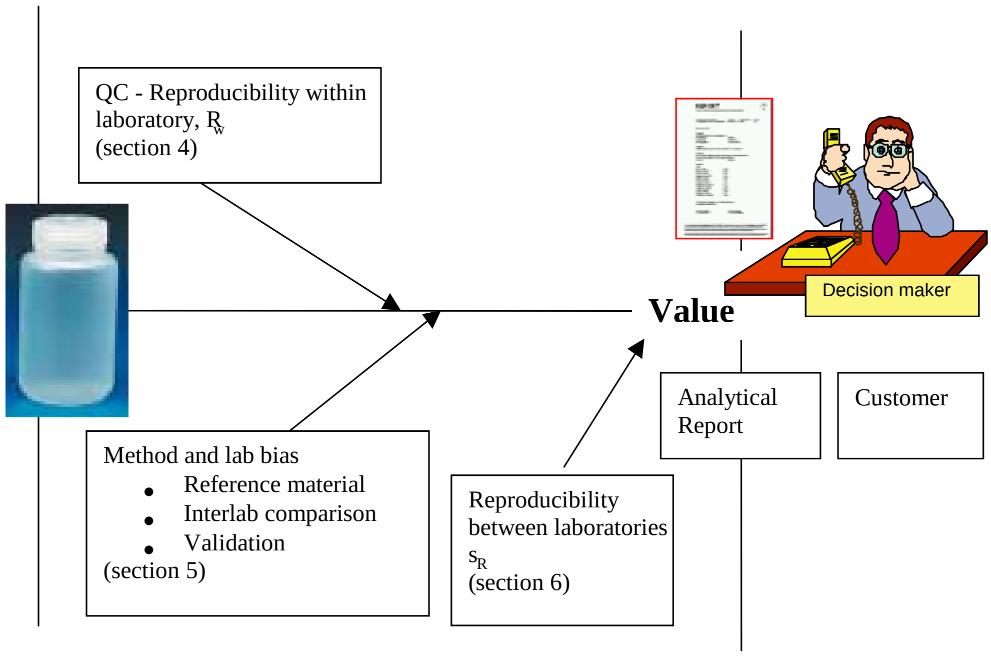


Figure 2. Measurement uncertainty model (fish-bone diagram), where the reproducibility within-laboratory is combined with estimates of the method and laboratory bias. Alternatively, according to ISO guide 21748 /8/, the combined uncertainty u_c can be directly estimated from the reproducibility between laboratories (s_R). This approach is treated in section 6

3.1 Customer needs

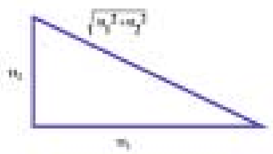
Before calculating or estimating the measurement uncertainty, it is recommended to find out what are the needs of the customers. After that, the main aim of the actual uncertainty calculations will be to find out if the laboratory can fulfil the customer demands with the analytical method in question. However, customers are not used to specifying demands, so in many cases the demands have to be set in dialogue with the customer. In cases where no demands have been established, a guiding principle could be that the calculated expanded uncertainty, U, should be approximately equal to, or less than, 2 times the reproducibility, s_R.

3.2 Flow scheme for uncertainty calculations

The flow scheme presented in this section forms the basis for the method outlined in this handbook. The flow scheme, involving 6 defined steps, should be followed in all cases. The example with NH_4-N in water shows the way forward for calculating the measurement uncertainty using the flow scheme. Explanations of the steps and their components will follow in the succeeding chapters. For each step, there may be one or several options for finding the desired information.

Background for the NH_4-N example – automatic photometric method: The laboratory has participated in 6 interlaboratory comparisons recently. All results have been somewhat higher than the nominal value. The laboratory therefore concludes that there may be a small positive bias. On average, the bias has been +2.2 %. This bias is considered small by the laboratory and is not corrected for in their analytical results, but exists, and is thus another uncertainty component.

For this method, the main sources of uncertainty are contamination and variation in sample handling, both causing random uncertainty components. These uncertainty sources will be included in the calculations below.

Step	Action	Example – Ammonium NH_4-N
1	***Specify Measurand***	Ammonium is measured in water according to EN/ISO 11732 /11/. The customer demand on expanded uncertainty is ± 10 %
2	***Quantify R_w comp.*** ***A control sample*** ***B possible steps not covered by the control sample***	A: Control limits are set to ± 3.34 % (95 % confidence limit) B: The control sample includes all analytical steps.
3	***Quantify bias comp.***	From interlaboratory comparisons over the last 3 years the bias result were 2.4; 2.7; 1.9: 1.4; 1.8: and 2.9. The root mean square (RMS) of the bias is 2.25 %. The uncertainty of the nominal values is $u(Cref)$ = 1.5 %. (see Appendix 4 for explanations)
4	***Convert components to standard uncertainty u(x)***	Confidence intervals and similar distributions can be converted to standard uncertainty /1, 2, 14/. $u(R_w)$ = 3.34/2 = 1.67 % $u(bias) = \sqrt{RMS_{bias}{}^2 + u(Cref)^2}$ $= \sqrt{2.25^2 + 1.5} = 2.71\%$
5	***Calculate combined standard uncertainty, u_c*** 	Standard uncertainties can be summed by taking the square root of the sum of the squares $u_c = \sqrt{u(R_w)^2 + (u(bias))^2} = \sqrt{1.67^2 + 2.71^2} = 3.18$
6	***Calculate expanded uncertainty,*** $U = 2 \cdot u_c$	The reason for calculating the expanded uncertainty is to reach a high enough confidence (app. 95 %) in that the "true value" lies within the interval given by the measurement result ± the uncertainty. $U = 2 \cdot 3.18 = 6.36 \approx 6$ %.

The measurement uncertainty for NH_4-N will thus be reported as ± 6 % at this concentration level.

3.3 *Summary table for uncertainty calculations*

The results of the calculations done in the flow scheme will then be summarised in a summary table.

Ammonium in water by EN/ISO 11732

Measurement uncertainty U (95 % confidence interval) is estimated to ± 6 %. The customer demand is ± 10 %. The calculations are based on control chart limits and interlaboratory comparisons.

		Value	***Relative u(x)***	***Comments***
Reproducibility within-laboratory, R_w				
Control sample $\overline{X}$ = 200 µg/L	*Rw*	Control limits is set to ± 3.34 %	1.67 %	
Other components		--		
Method and laboratory bias				
Reference material	bias	--		
Interlaboratory comparisons	bias	RMS_{bias}= 2.25 % *u(Cref)* = 1.5 %	2.71 %	$u(bias) = \sqrt{RMS_{bias}^2 + u(Cref)^2}$
Recovery test	bias	--		
Reproducibility between laboratories				
Interlaboratory comparisons	*R*	--	8.8 %	Data - see Section 6.2
Standard method	*R*	--		

Combined uncertainty, u_c is calculated from the control sample limits and bias estimation from interlaboratory comparisons. The s_R from interlaboratory comparisons can also be used (see 6.2) if a higher uncertainty estimation is acceptable.

Measurand	***Combined Uncertainty u_c***	***Expanded Uncertainty U***
Ammonium	$\sqrt{1.67^2 + 2.67^2} = 3.18$ %	$3.18 \cdot 2 = 6.4 \approx 6$ %

4 Reproducibility within-laboratory - u(R_w)

In this section the most common ways of estimating the reproducibility within-laboratory component, *u(Rw)*, for the measurement uncertainty calculation are explained:

- Stable control samples covering the whole analytical process. Normally one sample at low concentration level and one at a high concentration level.
- Control samples not covering the whole analytical process. Uncertainties estimated from control samples and from duplicate analyses of real samples with varying concentration levels.
- Unstable control samples.

It is of utmost importance that the estimation must cover all steps in the analytical chain and all types of matrices – worst-case scenario. The control sample data should be run in exactly the same way as the samples e.g. if the mean of duplicate samples is used for ordinary samples, then the mean of duplicate control samples should be used for the calculations.

It is likewise important to cover long-term variations of some systematic uncertainty components **within** the laboratory, e.g. caused by different stock solutions, new batches of critical reagents, recalibrations of equipment, etc. In order to have a representative basis for the uncertainty calculations and to reflect any such variation the number of results should ideally be more than 50 and cover a time period of approximately one year, but the need differs from method to method.

4.1 Customer demands

Some laboratories choose to use the customer demand when setting the limits in their control charts. The actual performance of the method is not interesting, as long as it meets the customer demands on expanded uncertainty. If, for example, the customer asks for data with an (expanded) measurement uncertainty of ± 10 %, then, from our experience, a good starting point is to set the control limits ± 5 %. The $u(R_w)$ used in the calculations will then be 2.5 %.[1] This is just a proposal and the measurement uncertainty calculations will show if these control limits are appropriate.

4.2 Control sample covering the whole analytical process

When a stable control sample is covering 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

[1] Treating the control limits according to GUM /1/ as type B estimate with 95 % confidence limit

samples. If the analyses performed cover a wide range of concentration levels, also control samples of other concentration levels should be used. Example: For NH_4-N two control sample levels (20 µg/L and 250 µg/L) were used during year 2002. The results for the manual analysis method are presented in the table below.

		Value	***Relative u(x)***	***Comments***
Reproducibility within-laboratory, R_w				
Control sample 1 $\overline{X}$ = 20.01 µg/L	s_{Rw}	Standard deviation 0.5 µg/L	2.5 %	From measurements in 2002, n = 75
Control sample 2 $\overline{X}$ = 250.3 µg/L	s_{Rw}	Standard deviation 3.7 µg/L	1.5 %	From measurements in 2002, n = 50
Other components		--		

4.3 Control sample for different matrices and concentration levels

When 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. Example: To estimate the matrix based uncertainties, duplicate analysis of ammonium is performed, and the s_r is estimated from the corresponding R%-chart (Range%-chart /13/), giving the repeatability of analysing natural samples having a normal matrix variation at different concentration levels.

The data set consists of 73 duplicate analyses in the range of 2 µg/L – 16000 µg/L. Most of the results were below 200 µg/L. The data is divided into two parts:

< 15 µg/L and > 15 µg/L

The s_r can be estimated from R%-charts constructed for both concentration ranges. The data is given in Appendix 5. The standard deviation is estimated from the range (see Appendix 8): $s = range / 1.128$.

		Value	***Relative u(x)***	***Comments***
Reproducibility within-laboratory, R_w				
Variation from duplicate analysis 2-15 µg/L: > 15 µg/L:	s_R		5.7 % 3.6 %	n = 43 ($\overline{X}$ = 6.50 µg/L) n = 30 ($\overline{X}$ = 816 µg/L)

At low levels it is often better to use an absolute uncertainty rather than a relative, as relative numbers tend to become extreme at very low concentrations. In this example the absolute *u(r)* becomes 0.37 µg/L for the natural sample (mean concentration 7 µg/L) and 0.5 µg/L for the control sample in Section 4.2 (mean concentration 20 µg/L).

As the estimate from duplicate analysis gives the repeatability component (s_r) only, it should be combined with the control sample results from Section 4.2 to give a better estimate of s_{Rw}. This way, the repeatability component will be included two times, but it is normally small in comparison to the between-days variation.

		Value	***u(x)***	***Comments***
Reproducibility within-laboratory, R_w				
Low level (2-15 µg/L)	s_{Rw}	0.5 µg/L from control sample and 0.37 µg/L from duplicates	0.6 µg/L	Absolute *u(x)* = $\sqrt{0.5^2+0.37^2}$
High level (> 15 µg/L)	s_{Rw}	1.5% from control sample and 3.6% from duplicates	3.9 %	Relative *u(x)* = $\sqrt{1.5^2+3.6^2}$

It can be noticed that the sample matrix has some effect on the variation of the results. The reason for this is not only the matrix, but also the concentration level, as almost all of the duplicate analyses were performed at a concentration level below 10 µg/L. The quantification limit of the measurement was 2 µg/L and the relative standard deviation usually becomes higher near that limit (cf. Figures 4 and 5 in Section 6.3).

4.4 Unstable control samples

If the laboratory does not have access to stable control samples, the reproducibility can be estimated using analysis of natural duplicate samples. The results from the duplicate sample analysis can either be treated in an R-chart, where the difference between the first and second analysis is plotted directly, or as an R %-chart, where the absolute difference between the sample pair is calculated in % of the average value of the sample pair. The latter approach is particularly useful when the concentration varies a lot from time to time.

In this example, duplicate samples for oxygen have been analysed on 50 occasions. The raw data is given in Appendix 6. The concentration variation is limited, so an R-chart approach is chosen. The difference between the first and the second analysis is calculated and plotted in a chart, see Figure 3. In this case, the second result is always subtracted from the first when constructing the R-chart, as it is important to look for systematic differences between the first and the second sample. The standard deviation for the results can be estimated from the average range of the duplicate samples (see Appendix 8), and in this case becomes 0.024. The control limits at ±2s thus lies at ±0.048. The average value of the first determination is 7.53, and the s_r thus equals 100·0.024/7.53 = 0.32 %.

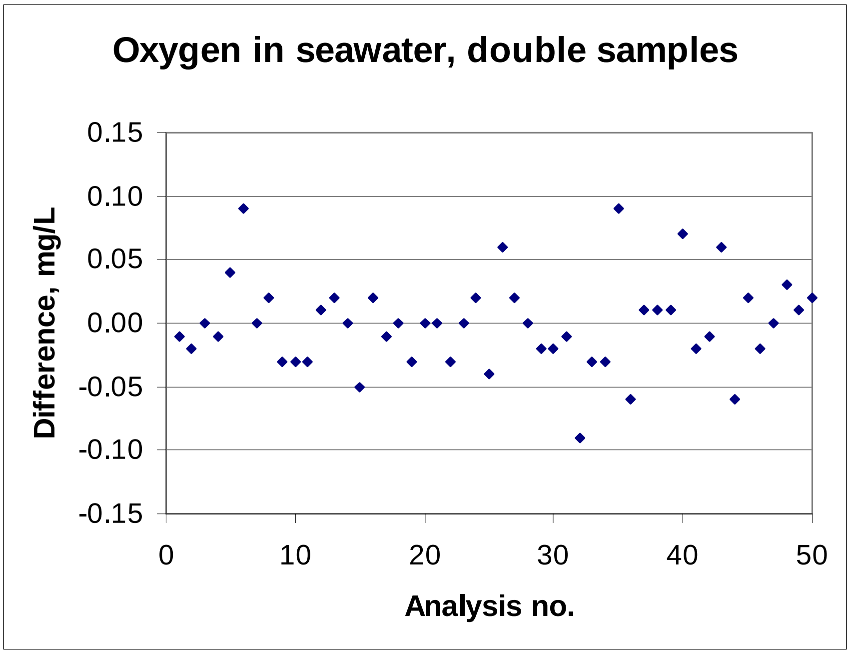
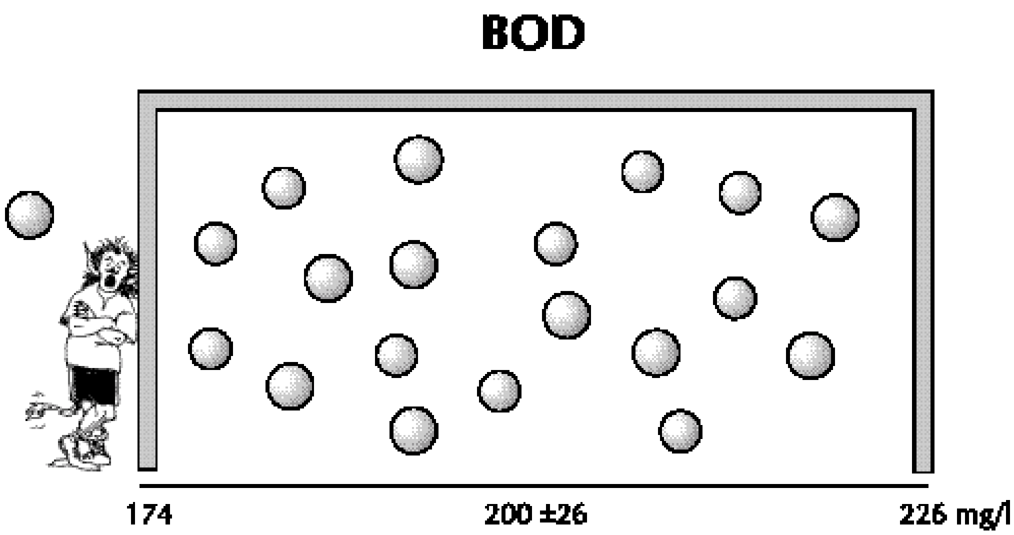


Figure 3. The difference between oxygen duplicate measurements plotted in an R-chart

However, this only gives the within-day variation (repeatability, s_r) for sampling and measurement, and there will also be a "long-term" uncertainty component from the variation in the calibration (here the thiosulphate used for titrating or the calibration of the oxygen probe, depending on method). For this particular analysis, the uncertainty component from the long-term variation in calibration is hard to measure, as no stable reference material or CRM is available for dissolved oxygen. One method would be to calibrate the same thiosulphate solution several times during a few days time, and use the variation between the results. Here we choose to estimate that component by a qualified guess, but laboratories are encouraged to also try the experimental approach.

The total reproducibility within-laboratory for dissolved oxygen thus becomes:

		Value	*Relative u(x)*	*Comments*
Reproducibility within-laboratory, R_w				
Duplicate measurements of natural samples, difference used in r-chart	s_R	s = 0.024 mg/L $\overline{X}$ = 7.53 mg/L	0.32 %	Measurements in 2000-2002, n= 50
Estimated variation from differences in calibration over time		s = 0.5 %	0.5 %	Estimate, based on experience
Combined uncertainty for R_w				
Repeatability + reproducibility in calibration		$\sqrt{0.32^2 + 0.5^2} = 0.59$ %		



5 Method and Laboratory bias – u(bias)

In this chapter the most common ways of estimating the bias components will be outlined, namely the use of CRM, participation in interlaboratory comparisons (proficiency test) and recovery tests. Sources of bias should always be eliminated if possible. According to GUM /1/ a measurement result should always be corrected if the bias is significant and based on reliable data such as a CRM. However, even if the bias is zero, it has to be estimated and treated as an uncertainty component. In many cases the bias can vary depending on changes in matrix. This can be reflected when analysing several matrix CRMs, e.g. the bias could be both positive and negative. Examples are given and explained for the proposed calculations.

For every estimation of the uncertainty from the method and laboratory bias, two components have to be estimated to obtain *u(bias)*:

1) the bias (as % difference from the nominal or certified value)

2) the uncertainty of the nominal/certified value, *u(Cref)* or *u(Crecovery)*

The uncertainty of the bias, *u(bias)* can be estimated by

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

and if only one CRM is used also the s_{bias} have to be included and *u(bias)* can the be estimated /14, 15/[2] by

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

5.1 Certified Reference Material

Regular analysis of a CRM can be used to estimate the bias. The reference material should be analysed in at least 5 different analytical series (e.g. on 5 different days) before the values are used.

In this example the certified value is 11.5 ± 0.5, with a 95 % confidence interval.

Uncertainty component from the uncertainty of the certified value	
Step	**Step**
Convert the confidence interval to *u(Cref)*	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(Cref)*	100·(0.26/11.5) = 2.21%

3	***Quantify Method and laboratory bias***	$\text{bias} = 100 \cdot (11.9 - 11.5)/11.5 = 3.48\,\%$ s_{bias} = 2.2 % (n = 12) *u(Cref)* = 2.21 %
4	***Convert components to standard uncertainty u (x)***	$u(bias) = \sqrt{(bias)^2 + \left(\frac{s_{bias}}{\sqrt{n}}\right)^2 + u(Cref)^2} =$ $\sqrt{(3.48)^2 + \left(\frac{2.2}{\sqrt{12}}\right)^2 + 2.21^2} = 4.2\,\%$

If **several CRM:s** are used, we will get different values for the bias. The uncertainty of the bias estimation will be calculated in the following way (see also section 5.2).

3	***Quantify Method and laboratory bias***	bias CRM1 is 3.48%, s=2,2 (n=12), *u(Cref)*=2.21 % bias CRM2 is -0.9% s=2,0 (n=7)), *u(Cref)*=1.8 % bias CRM3 is 2.4%, s= 2,8 (n=10), *u(Cref)*=1.8 % For the bias the RMS_{bias} = 2.50 mean *u(Cref)*=1,9 %
4	***Convert components to standard uncertainty u (x)***	$u(bias) = \sqrt{RMS_{bias}{}^2 + u(Cref)^2}$ $\sqrt{2.50^2 + 1.9^2} = 3.1\%$

5.2 Interlaboratory comparisons

In this case the results from interlaboratory comparisons are used in the same way as a reference material, i.e. to estimate the bias. 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.

Biases can be both positive and negative. Even if the results appear to give positive biases on certain occasions and negative on others, all bias values can be used to estimate the uncertainty component, RMS_{bias}.

The way forward is very similar to that for reference materials. However, the estimation of the bias from interlaboratory comparisons has more uncertainty to it, and thus usually becomes a bit higher than if CRMs are used. This is partly due to the fact that the certified value of a CRM normally is better defined than a nominal or assigned value in an interlaboratory comparison exercise. In some cases the calculated uncertainty *u(Cref)* from an interlaboratory comparison becomes too high and is not valid for estimating the *u(bias)*.

Uncertainty component from the uncertainty of the nominal value	
Step	**Example**
Find the between laboratory standard deviations, s_R, for the exercises.	The s_R has been on average 9% in the 6 exercises.
Calculate *u(Cref)*	Mean number of participants = 12. $u(Cref) = \frac{s_R}{\sqrt{n}} = \frac{9}{\sqrt{12}} = 2.6\ \%$

The bias has been 2 %, 7 %, -2 %, 3 %, 6 % and 5%, in the 6 interlaboratory comparisons where the laboratory has participated.

3	***Quantify Method and laboratory bias***	RMS_{bias} = 4.6 %, *u(Cref)*= 2.6 %
4	***Convert components to standard uncertainty u (x)***	$u(bias) = \sqrt{RMS_{bias}^2 + u(Cref)^2} =$ $= \sqrt{4.6^2 + 2.6^2} = 5.3\%$

5.3 Recovery

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.

In an experiment the recoveries for an added spike were 95 %, 98 %, 97 %, 96 %, 99 % and 96 % for 6 **different** sample matrices. The average is 96.8 %. The spike of 0.5 mL was added with a micropipette.

Uncertainty component from the definition of 100% recovery, u(Crecovery)	
Step	**Example**
Uncertainty of the 100% recovery. Main components are concentration, *u(conc)* of standard and volume added *u(vol)*	*u(conc)* - Certificate ± 1.2 % (95 % conf. limit) gives = 0.6 % *u(vol)* - This value can normally be found in the manufacturer's specifications, or better use the limits specified in your laboratory. Max bias 1 % (rectangular interval) and repeatability max 0.5 % $u(vol) = \sqrt{\left(\frac{1}{\sqrt{3}}\right)^2 + 0.5^2} = 0.76\,\%$
Calculate *u(Crecovery)*	$\sqrt{u(conc)^2 + u(vol)^2} = \sqrt{0.6^2 + 0.76^2} = 1.0\,\%$

3	***Quantify Method and laboratory bias***	RMS_{bias} = 3.44 % *u(Crecovery)*=1.0 %
4	***Convert components to standard uncertainty u (x)***	$u(bias) = \sqrt{RMS_{bias}{}^2 + u(Cref)^2} =$ $= \sqrt{3.44^2 + 1.0^2} = 3.6\ \%$

6 Reproducibility between laboratories, s_R

If the demand on uncertainty is low, it can be possible to directly use the s_R from interlaboratory comparisons as an approximation of u_c /8/. In such case 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 inhomogenity or matrix variations.

6.1 Data given in standard method

In order to use a figure taken directly from the standard method, the laboratory must prove that they are able to perform in accordance with the standard method /8/, i.e. demonstrating control of bias and verification of the repeatability, s_r. Reproducibility data can either be given as a standard deviation s_R or as reproducibility limit R and then $s_R = R/2.8$

The example below is taken from ISO/DIS 15586 *Water Quality — Determination of trace elements by atomic absorption spectrometry with graphite furnace*. The matrix is wastewater. Combined uncertainty in wastewater, u_c, is taken from the s_R from interlaboratory comparison exercises quoted in the ISO method.

Table 1 ISO/DIS 15586 - Results from the interlaboratory comparison – Cd in water with graphite furnace AAS. The wastewater was digested by the participants.

Cd		**n**	**Outliers**	**Nominal value** µg/L	**Mean** µg/L	**Recovery,** %	s_r %	s_R %
Synthetic	Lower	33	1	0.3	0.303	101	3.5	17.0
Synthetic	Higher	34	2	2.7	2.81	104	1.9	10.7
Fresh water	Lower	31	2		0.572		2.9	14.9
Fresh water	Higher	31	3		3.07		2.1	10.4
Waste water		27	2		1.00		3.1	27.5

Measurand	*Combined Uncertainty u_c*	*Expanded Uncertainty U*
Cd	$u_c = 27.5$ %	$2 \cdot u_c = 55$ % ≈ 50 %

6.2 Data from interlaboratory comparisons

Interlaboratory comparisons are valuable tools in uncertainty evaluation. The reproducibility between laboratories is normally given directly in reports from the exercises as s_R.

These data may well be used by a laboratory (having performed satisfactorily in the comparisons) as the standard uncertainty of the analysed parameter, provided that the comparison covers all relevant uncertainty components and steps (see /9/, section 5.4.6.3). For example, a standard deviation in an interlaboratory comparison, s_R, can be directly used as a combined standard uncertainty, u_c.

Table 2 Summary results (mean values) from 10 interlaboratory comparisons that Lab A has participated in. The reproducibility standard deviation is given in absolute units, s_R and in relative units s_R %.

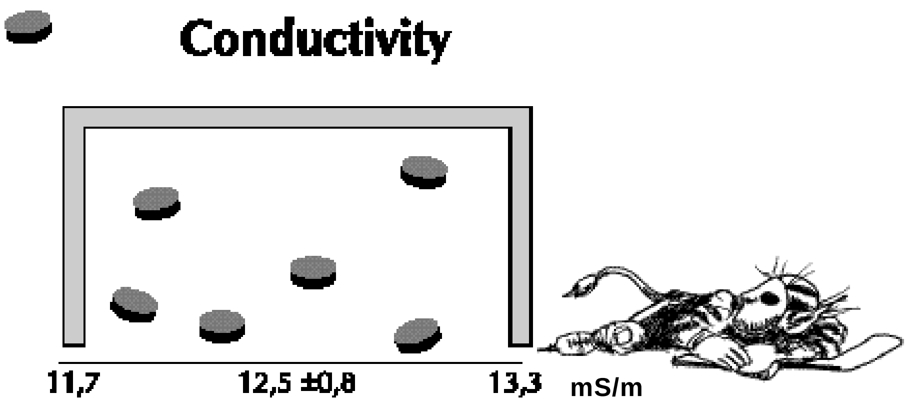
Variable	**Nominal value**	**Lab A % deviation**	s_R *(abs)*	s_R %	**No. of labs**	**Excluded**
pH	7.64	-0.037	0.101		90	5
Conductivity, mS/m	12.5	-2.8	0.40	3.2	86	6
Alkalinity, mmol/L	0.673	2.3	0.026	3.9	60	3
Turbidity, FNU	1.4	-9.1	0.1	14.2	44	3
NH_4-N, µg/L	146	2.2	12.0	8.8	34	5
NO_3-N, µg/L	432	-1.6	16.3	3.7	39	6

In Table 2 we find that for conductivity, for instance, the mean value for the results from 10 interlaboratory comparisons is 12.5 mS/m. The reproducibility relative standard deviation is 0.4 (3.2 %), which is an average (or pooled) standard deviation between the laboratories in the different interlaboratory comparisons and this value can be taken as an estimate of combined uncertainty i.e.

$$u_c\,(conductivity) = s_R = 0.4 \text{ mS/m, thus } U = 2 \cdot 0.4 = 0.8 \text{ mS/m}$$

If we take the ammonium results, we have a mean nominal value of 146 µg/L, and we find that the reproducibility, s_R, is 8.8 %. Thus $U = 2 \cdot 8.8 = 17.6 = 18$ % at this concentration level.

Comment: In Section 3 the expanded uncertainty for ammonium is 5 % using an automated method in one highly qualified laboratory.



7 Examples

In this chapter, practical examples on how measurement uncertainty can be calculated using the method of this handbook are given.

7.1 *Ammonium in water*

Ammonium in water has already been treated in section 3.2 and section 6.2 . The results are summarised in Table 3.

Table 3 Measurement uncertainty of ammonium in water – comparison of different calculations

Uncertainty calculations based on	Relative expanded uncertainty, U	Comment
Control sample + proficiency testing	± 6 %	Uncertainty for one good laboratory- level 200 µg/L.
Interlaboratory comparisons	± 18 %	Uncertainty in general among laboratories – level 150 µg/L

7.2 *BOD in wastewater*

Biological Oxygen Demand, BOD, is a standard parameter in the monitoring of wastewater. This example shows how data from ordinary internal quality control can be used together with CRM results or data from interlaboratory comparison exercises to calculate the within-lab reproducibility and bias components of the measurement uncertainty. The results are summarised in Table 4

Table 4 Measurement uncertainty of BOD in water - comparison of different calculations

Uncertainty calculations based on	Relative expanded uncertainty, U	Comment
Control sample + CRM	± 10 %	
Control sample + interlaboratory comparisons	± 10 %	n = 3, unreliable estimate
Interlaboratory comparisons	± 16 %	Uncertainty in general among laboratories

For BOD at high concentrations, using the dilution analytical method, the major error sources are the actual oxygen measurement and variation in the quality of the seeding solution. These errors will be included in the calculations.

The raw data from the internal quality control, using a CRM, used for the calculations is shown in Appendix 7.

The laboratory has only participated in three interlaboratory comparison exercises the last 2 years (Table 5). At least six would be needed, so here we estimate the bias two different ways – with CRM and with interlaboratory comparisons.

Table 5 BOD - results from interlaboratory comparisons

Exercise	**Nominal value**	**Laboratory result**	**Bias**	s_R	**Number of labs**
	mg/L	mg/L	%	%	
1	154	161	+ 4.5	7.2	23
2	219	210	- 4.1	6.6	25
3	176	180	+2.3	9.8	19
$\overline{X}$			+0.9	7.87[3]	22.3
RMS_{bias}			3.76	-	-

[3] If s_R or the number of participants vary substantially from exercise to exercise, then a pooled standard deviation will be more correct to use. In this case, where the variation in s_R is limited, we simply calculate the mean s_R (the corresponding pooled standard deviation becomes 7.82, an insignificant difference).

Example A: BOD with Internal quality control + a CRM

Step	*Action*	*Example: BOD in wastewater*
1	Specify Measurand	BOD in wastewater, measured with EN1899-1 (method with dilution, seeding and ATU). The demand on uncertainty is ± 20 %.
2	Quantify *u(Rw)* A control sample B possible steps not covered by the control sample	A: The control sample, which is a CRM, gives an s = 2.6 % at a level of 206 mg/L O_2. s = 2.6 % is also when setting the control chart limits. B: The analysis of the control sample includes all analytical steps after sampling
3	Quantify Method and laboratory bias	The CRM is certified to 206 ±5 mg/L O_2. The average result of the control chart is 214.8. Thus, there is a bias of 8.8 mg/L = 4.3 %. The s_{bias} is 2.6 % (n=19) The *u(Cref)* is 5 mg/L / 1.96 = 1.2 %
4	Convert components to standard uncertainty *u(x)*	$u(R_w)$ = 2.6 % $u(bias) = \sqrt{bias^2 + \left(\frac{s_{bias}}{\sqrt{n}}\right)^2 + u(Cref)^2}$ $= \sqrt{4.3^2 + \left(\frac{2.6}{\sqrt{19}}\right)^2 + 1.2^2} = 4.5\ \%$
5	Calculate combined standard uncertainty, u_c	$u_c = \sqrt{2.6^2 + 4.5^2} =\ 5.2\ \%$
6	Calculate expanded uncertainty, $U = 2 \cdot u_c$	$U = 2 \cdot 5.2 = 10.4 \approx 10\ \%$

Example B: BOD with Internal quality control + interlaboratory comparison results

Step	*Action*	*Example: BOD in wastewater*
1	Specify Measurand	BOD in wastewater, measured with EN1899-1 (method with dilution, seeding and ATU). The demand on uncertainty is ± 20 %.
2	Quantify *u(Rw)* A control sample B possible steps not covered by the control sample	A: The control sample, which is a CRM, gives an *s* of 2.6 % at a level of 206 mg/L O_2. s = 2.6 % is also used as s when setting the control chart limits. B: The analysis of the control sample includes all analytical steps after sampling
3	Quantify Method and laboratory bias Data from Table 5	RMS_{bias} = 3.76 $u(Cref) = \frac{s_R}{\sqrt{n}} = \frac{7.9}{\sqrt{22.3}} = 1.67$
4	Convert components to standard uncertainty *u(x)*	$u(R_w)$ = 2.6 % $u(bias) = \sqrt{RMS_{bias}{}^2 + u(Cref)^2} =$ $\sqrt{3.76^2 + 1.67^2} = 4.11\%$
5	Calculate combined standard uncertainty, u_c	$u_c = \sqrt{2.6^2 + 4.11^2} = 4.86$ %
6	Calculate expanded uncertainty, $U = 2 \cdot u_c$	$U = 2 \cdot 4.86 = 9.7 \approx 10$ %

7.3 PCB in sediment

In this example, the $u(R_w)$ is estimated from a quality control sample and the u(bias) is estimated from two different sources: in the first example the use of a CRM and in the second example participation in interlaboratory comparisons. In the summary table both ways of calculating the *u(bias)* will be compared.

For this analysis, the sample-work up is a major error source (both for random and systematic errors), and it is thus crucial that this step is included in the calculations. The number of interlaboratory comparisons is too few to get a good estimate.

Example C: PCB with Internal quality control + a CRM

Step	*Action*	*Example: PCB in sediment*
1	Specify Measurand	Sum of 7 PCB:s in sediment by extraction and GC-MS(SIM). Demand on expanded uncertainty is ± 20 %.
2	Quantify *u(Rw)* A control sample B possible steps not covered by the control sample	A: The control sample, which is a CRM, gives an s_{Rw} = 8 % at a level of 150 µg/kg dry matter. s_{Rw} = 8 % is also used when setting the control chart limits. B: The analysis of the control sample includes all steps except for drying the sample to determine the dry weight. The uncertainty contribution from that step is considered small and is not accounted for.
3	Quantify method and laboratory bias	The CRM is certified to 152 ± 14 µg/kg. The average result of the control chart is 144. Thus, there is a bias = 5.3 %. The s_{bias} = 8 % (n=22) *u(Cref)* 14 µg/kg/1.96, which is 4.7 % relative.
4	Convert components to standard uncertainty *u(x)*	$u(R_w)$ = 8 % $u(bias) = \sqrt{bias^2 + \left(\frac{s_{bias}}{\sqrt{n}}\right)^2 + u(Cref)^2}$ $= \sqrt{5.3^2 + \left(\frac{8}{\sqrt{22}}\right)^2 + 4.7^2} = 7.29$
5	Calculate combined standard uncertainty, u_c	$u_c = \sqrt{8^2 + 7.29^2}$ = 10.8 %
6	Calculate expanded uncertainty, $U = 2 \cdot u_c$	$U = 2 \cdot 10.8 = 21.6 \approx 22\ \%$

Example D: PCB with Internal quality control + interlaboratory comparison

Step	*Action*	*Example: PCB in sediment*
1	Specify Measurand	Sum of 7 PCB:s in sediment by extraction and GC-MS(SIM). Demand on expanded uncertainty is 20 %.
2	Quantify $u(R_w)$ A control sample B possible steps not covered by the control sample	A: The control sample, which is a stable in-house material, gives s_{Rw} = 8 % at a level of 150 µg/kg dry matter. s_{Rw} = 8 % is also used as s when setting the control chart limits. B: The analysis of the control sample includes all steps except for drying the sample to determine the dry weight. The uncertainty contribution from that step is considered small and is not accounted for.
3	Quantify Method and laboratory bias	Participation in 3 interlaboratory comparisons with concentration levels similar to the internal quality control. The bias in the 3 exercises has been –2 %, -12 % and –5 %. RMS_{bias} = 7.6 The s_R in the three exercises has been 12 %, 10 % and 11 %, on average s_R = 11 % (n=14) $u(Cref) = \frac{11}{\sqrt{14}} = 2.9\%$
4	Convert components to standard uncertainty $u(x)$	The $u(R_w)$ is 8 % $u(bias) = \sqrt{RMS_{bias}^2 + u(Cref)^2} =$ $\sqrt{7.6^2 + 2.9^2} = 8.1\%$
5	Calculate combined standard uncertainty, u_c	$u_c = \sqrt{8^2 + 8.1^2} = 11.4$
6	Calculate expanded uncertainty, $U = 2 \cdot u_c$	$U = 2 \cdot 11.4 = 22.8 \approx 23\%$

Summary table for PCB measurement uncertainty calculations

PCB in sediment by extraction and GC-MS (SIM)

Measurement uncertainty U (95 % confidence interval) is estimated to ± 20 % (relative) for 7 PCB:s in sediments at a level of 150 µg/kg dry weight. The customer demand is ± 20 %. The calculations are based on internal quality control using a stable sample, CRM and the participation in a limited amount of interlaboratory comparison exercises.

		Value	*u(x)*	*Comments*
Reproducibility within-laboratory, R_w				
Control sample $\bar{X}$ = 160 µg/kg dry weight	$u(R_w)$	12.8 µg/kg dry weight	8 %	
Other components		too small to be considered		
Method and laboratory bias				
Reference material		Bias: 5.3 % s_{bias} = 8 ; n = 22 $u(Cref)$ = 4,7 %	$u(bias)$ = 5.88	$u(bias)=$ $\sqrt{bias^2 + \left(\frac{s_{bias}}{\sqrt{n}}\right)^2 + u(Cref)^2}$
Interlaboratory comparison n = 3		RMS_{bias} = 7.6 $u(Cref)$ = 2.9 %	$u(bias)$ = 8.1	$u(bias)=$ $\sqrt{RMS_{bias}^2 + u(Cref)^2}$

Combined uncertainty, u_c, is calculated from internal quality control and the maximum bias - interlaboratory comparisons.

Measurand	*Combined Uncertainty u_c*	*Expanded Uncertainty U*
PCB	$u_c = \sqrt{8^2 + 8.1^2} = 11.4$	$U = 2 \cdot u_c = 2 \cdot 11.4 = 22.8 \approx 23\,\%$

Conclusion: In this case the calculation of the *u(bias)* gives similar results regardless of whether CRM or interlaboratory comparison results are used. Sometimes interlaboratory comparisons will give considerably higher values, and it might in such cases be more correct to use the CRM results.

7.4 Concentration ranges

The measurement uncertainty will normally vary with concentration, both in absolute and relative terms. If the concentration range of the reported data is large, it is thus often necessary to take this into account. For lead (Pb) in water, a recovery experiment was carried out a number of times to investigate within-lab reproducibility over the measurable range – the major component of the measurement uncertainty at low levels. The following results were obtained:

Table 6 Within-lab reproducibility and recovery for Pb determined with ICP-MS at different concentration levels.

Addition, µg/L	Pb, % recovery	s, %
0.01	109.7	53.8
0.1	125.2	12.1
0.4	91.8	5
1	98.4	3.0
10	98	1.7
10	100.5	1.3
100	105.5	1.4

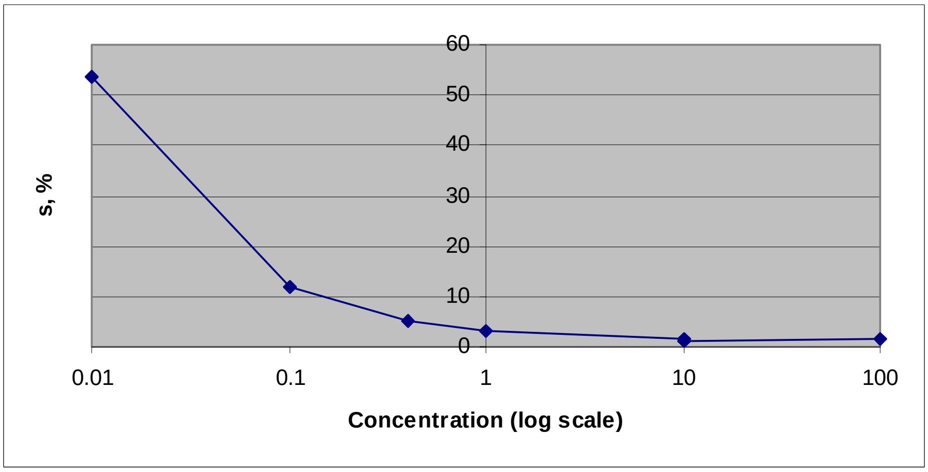


Figure 4 Within-lab reproducibility for Pb over the concentration range

It is clear from the results that the measurement uncertainty, here represented by *s*, is strongly concentration dependent. Two approaches are recommended for using these data:

(1) To divide the measurable range into several parts, and use a fixed relative measurement uncertainty or absolute uncertainty – see Table 7.

Table 7 Within-lab reproducibility for Pb divided into three concentration ranges

Within-lab reproducibility Pb		
Range (µg/L)	*s(rel)*	*s(rel) or (abs)*
< 0.1	50 %	0.01 (µg/L)
0.1 - 10	10 %	10 %
> 10	2 %	2 %

In the second column *s* is relative and given in %. In the third column *s* is also relative but an absolute value is given in the lower range close to the detection limit.

(2) To use an equation that describes how the measurement uncertainty varies with concentration

Plotting *s* % against 1/concentration gives a straight line, and a relatively simple equation. (see Figure 5).

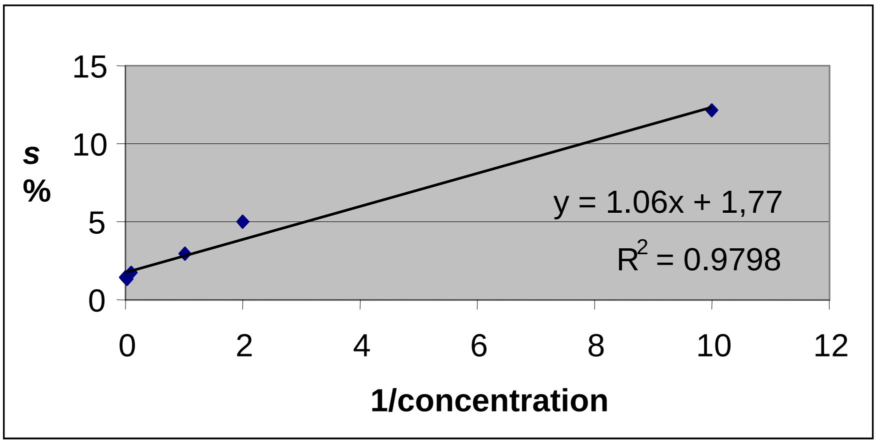


Figure 5: The relationship between within-lab reproducibility and the inverted concentration for Pb in the range 0.1 – 100 µg/L.

The straight-line equation above tells us that the within-lab reproducibility equals 1.06 multiplied with 1/concentration plus 1.77. For example, at a concentration of 2 µg/L the within-lab reproducibility becomes $1.06 \cdot 1/2 + 1.77 = 2.3$ %. When reporting to customers, the laboratory can choose between quoting the formula or calculating the measurement uncertainty for each value, using the formula. For further reading, see for example /2/.

8 Reporting uncertainty

This is an example on what a data report could look like, when measurement uncertainty has been calculated and is reported together with the data. The company and accreditation body logotypes are omitted, and the report does not contain all information normally required for an accredited laboratory. It is recommended to use either relative or absolute values for the benefit of the customer.

Analytical Report

Sample identification: P1 – P4
Samples received: 14 December 2002
Analysis period: 14 –16 December 2002

Results

NH_4-N (µg/L):

Sample	Result	U	Method
P1	103	±6%	23B
P2	122	±6%	23B
P3	12	±10%	23B
P4	14	±10%	23B

TOC (mg/L)

Sample	Result	U	Method
P1	40	±4.0	12-3
P2	35	±3.5	12-3
P3	10	±1.0	12-3
P4	9	±0.9	12-3

Signed: Dr Analyst

The laboratory should also prepare a note explaining how the measurement uncertainty has been calculated for the different parameters. Normally, such an explanatory note should be communicated to regular customers and other customers who ask for information. An example is given below:

Note on measurement uncertainty from Dr Analyst's laboratory

Measurement uncertainty:
U = expanded Measurement Uncertainty, estimated from control sample results, interlaboratory comparison and the analyses of CRMs, using a coverage factor of 2 to reach approximately 95% confidence level.

NH_4-N: *U* is estimated to 6% above 100 µg/L and 10% below 100 µg/L.

TOC: *U* is estimated to 10% over the whole concentration range.

References:

- Guide To The Expression Of Uncertainty In Measurement (GUM)
- Quantifying Uncertainty in Analytical Measurement. EURACHEM/CITAC Guide
- Handbook for calculation of measurement uncertainty in environmental laboratories

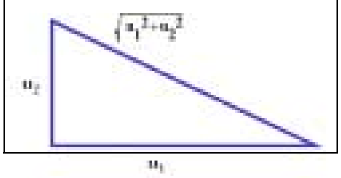
9 References

1. Guide To The Expression Of Uncertainty In Measurement (GUM). BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML. International Organization of Standardization, Geneva Switzerland, 1st Edition 1993, Corrected and reprinted 1995.
2. Quantifying Uncertainty in Analytical Measurement. EURACHEM/CITAC Guide, 2nd Edition, 2000
3. Measurement Uncertainty in Testing, Eurolab Technical Report No. 1/2002
4. Interlaboratory comparison test data, personal communication, H. Hovind, NIVA, Norway.
5. International Vocabulary of Basic and General Terms in Metrology (VIM). ISO, 1993
6. ISO/IEC 3534-1-2, Statistics – Vocabulary and symbols Parts 1-2
7. ISO 5725-1-6:1994, Accuracy (trueness and precision) of measurement methods and results
8. ISO/DTS 21748:2003, Guide to the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation
9. EN ISO/IEC 17025:2000, General Requirements for the Competence of Calibration and Testing Laboratories
10. ISO/TR 13530:1997, Water quality – Guide to analytical quality control for water analysis
11. EN ISO 11732:1997, Water quality -- Determination of ammonium nitrogen by flow analysis (CFA and FIA) and spectrometric detection
12. EA-4/xx: Proposed "EA guideline on The Expression of uncertainty in quantitative testing". Project to be finalised in 2003. (www.european-accreditation.org/)
13. ISO8258, First edition, 1991-12-15, Shewhart Control Charts
14. V. J. Barwick, Ellison L.R., Analyst, 1999, 124, 981-990
15. E. Hund, D.L. Massart and J. Smeyers-Verbeke, Operational definitions of uncertainty. TrAC, 20 (8), 2001

10 Appendices

Appendix 1: Empty flow scheme for calculations

Before starting: Always identify the main error sources, to make sure that they are included in the calculations.

Step	*Action*	*Measurand:*
1	*Specify Measurand*	(measurand) in (matrix) by (method) The customer demand on expanded uncertainty is ± _ %.
2	*Quantify $u(R_w)$* *A control sample* *B possible steps not covered by the control sample*	A: B:
3	*Quantify method and laboratory bias*	
4	*Convert components to standard uncertainty u(x)*	
5	*Calculate combined standard uncertainty, u_c* 	
6	*Calculate expanded uncertainty, $U = 2 \cdot u_c$*	

Appendix 2: Empty summary table

(measurand) in (matrix) by (method)
Measurement uncertainty U (95 % confidence interval) is estimated to ± _ % (relative) for (measurand) in (matrix) at a level of _ (unit). The customer demand is ± _ %. The calculations are based on (control samples/control limits/CRM/interlaboratory comparison/other).

		Value	***Relative u(x)***	***Comments***
Reproducibility within-laboratory, R_w				
Control sample $\overline{X}$ = (conc) (unit)	s_{Rw}			
Other components				
Method and laboratory bias				
Reference material	bias			
Interlaboratory comparison	bias			
Recovery test	bias			
Reproducibility between laboratories				
Interlaboratory comparison	s_R			
Standard method	s_R			

Combined uncertainty, u_c, is calculated from __ and bias from __.

Measurand	***Combined Uncertainty u_c***	***Expanded Uncertainty U***
		$2 \cdot u_c =$

Appendix 3: Error model used in this handbook

This model is a simplification of the model presented in the ISO guide /8/:

$$y = m + (\delta + B) + e$$

y	measurement result of a sample
m	expected value for y
δ	method bias
B	laboratory bias – the uncertainty for these are combined to *u(bias)*
e	random error at within-laboratory reproducibility conditions, R_w

Uncertainty estimation in section 3 to 5

$$u(y)^2 = s_{Rw}^{\ 2} + u(bias)^2$$

$s_{Rw}^{\ 2}$	The estimated variance of e under within-laboratory reproducibility conditions – intermediate precision. In the ISO guide the repeatability, s_r is used as an estimate of e.
$u(bias)^2$	The estimated variance of method bias and laboratory bias.

Uncertainty estimation in section 6

The combined uncertainty *u(y)* or u_c can also be estimated by from reproducibility data.

$u(y)^2 = s_L^{\ 2} + s_r^{\ 2} = s_R^{\ 2}$ - equation A6 ref. /8/

where s_R^2 is the estimated variance under reproducibility conditions and where s_L^2 is either the estimated variance of B if one method is used by all laboratories or an estimated variance of B and δ if several different methods have been used in the collaborative study and s_r^2 is the estimated variance of e.

Comment
For samples that are more inhomogeneous and have big variations in matrix the estimation of the measurement uncertainty of the method can become too low. However we recommend the use of repeatability limit for duplicate analyses $r = 2.8 \cdot s_r$ in order to control sample inhomogenity.

Appendix 4: Uncertainty of bias for NH_4-N in section 3.2

Results for a laboratory from interlaboratory comparisons of NH_4-N in water.

Exercise	**Nominal value x_{ref}**	**Laboratory result x_i**	**Bias**	s_R	**Number of labs**
	mg/L	mg/L	%	%	
1999 1	81	83	2.4	10	31
2	73	75	2.7	7	36
2000 1	264	269	1.9	8	32
2	210	213	1.4	10	35
2001 1	110	112	1.8	7	36
2	140	144	2.9	11	34
$\overline{X}$			**+ 2.18**	8.8	34
RMS			**2.25**	-	-

$$RMS \text{ of the bias } = \sqrt{\frac{\sum bias_i^{\ 2}}{n}} = \sqrt{\frac{2.4^2 + 2.7^2 + ...2.9^2}{n}} = 2.25 \text{ % (rel)}$$

$$u(Cref) = \frac{s_R}{\sqrt{n}} = \frac{8.8}{\sqrt{34}} = 1.5 \text{ % (rel)}$$

A t-test shows that the bias (+2.18 %) is not significant (t = 0.6). However, in order not to complicate the calculations when the bias is small, t-test are normally not performed.

The mean value of s_R is used. If differences in number of laboratories and s_R are very big pooled standard deviations should be used. In this case the pooled standard deviation is 8.9 % for s_R which is the same as the mean value of 8.8 %.

Appendix 5: Raw data for NH_4-N in section 4.3

The estimation of the standard deviation from the range is explained in Appendix 8

concentration < 15 µg/L

Sample	*X1*	*X2*	$\overline{X} = \frac{x_{i1} + x_{i2}}{2}$	$d = x_{i1} - x_{i2}$	$100 \cdot \frac{\lvert d \rvert}{\overline{X}} = r\%$	
1	*7.46*	*7.25*	7.355	0.210	2.855	
2	*9.01*	*9.17*	9.090	-0.160	1.760	
3	*3.6*	*3.1*	3.350	0.500	14.925	
4	*6.48*	*6.48*	6.480	0.000	0.000	
5	*14.49*	*14.12*	14.305	0.370	2.587	
6	*10.84*	*9.89*	10.365	0.950	9.165	
7	*4.61*	*5*	4.805	-0.390	8.117	
8	*2.6*	*2.42*	2.510	0.180	7.171	
9	*2.8*	*2.62*	2.710	0.180	6.642	
10	*5.84*	*6.19*	6.015	-0.350	5.819	
11	*2.12*	*2.5*	2.310	-0.380	16.450	
12	*2.3*	*2.11*	2.205	0.190	8.617	
13	*2.52*	*2.89*	2.705	-0.370	13.678	
14	*3.71*	*3.71*	3.710	0.000	0.000	
15	*7.43*	*7.43*	7.430	0.000	0.000	
16	*8.83*	*8.51*	8.670	0.320	3.691	
17	*9.12*	*8.79*	8.955	0.330	3.685	
18	*8.24*	*7.9*	8.070	0.340	4.213	
19	*2.62*	*2.78*	2.700	-0.160	5.926	
20	*3.33*	*3.33*	3.330	0.000	0.000	
21	*2.69*	*2.69*	2.690	0.000	0.000	
22	*12.09*	*12.09*	12.090	0.000	0.000	
23	*4.24*	*4.24*	4.240	0.000	0.000	
24	*10.49*	*10.64*	10.565	-0.150	1.420	
25	*3.68*	*3.52*	3.600	0.160	4.444	
26	*9.37*	*9.37*	9.370	0.000	0.000	
27	*2.22*	*2.06*	2.140	0.160	7.477	
28	*6.1*	*6.1*	6.100	0.000	0.000	
29	*2.96*	*2.86*	2.910	0.100	3.436	
30	*14.02*	*13.7*	13.860	0.320	2.309	
31	*4.24*	*3.62*	3.930	0.620	15.776	
32	*5.1*	*4.61*	4.855	0.490	10.093	
33	*2.78*	*2.62*	2.700	0.160	5.926	
34	*8.52*	*6.81*	7.665	1.710	22.309	
35	*12.82*	*14.05*	13.435	-1.230	9.155	
36	*3.17*	*2.4*	2.785	0.770	27.648	
37	*11.28*	*11.43*	11.355	-0.150	1.321	
38	*14.31*	*13.82*	14.065	0.490	3.484	
39	*4.01*	*4.48*	4.245	-0.470	11.072	
40	*3.27*	*3.58*	3.425	-0.310	9.051	
41	*9.98*	*10.29*	10.135	-0.310	3.059	
42	*12.56*	*13.66*	13.110	-1.100	8.391	
43	*3.35*	*2.88*	3.115	0.470	15.088	
		Mean:	**6.499**		**6.4363**	= mean range (%)

s(r) % = range(mean)/1.128 = 5.71 %

concentration > 15 µg/L

Sample	*X1*	*X2*	$\overline{X} = \frac{x_{i1} + x_{i2}}{2}$	$d = x_{i1} - x_{i2}$	$100 \cdot \frac{\lvert d \rvert}{\overline{X}} = r\%$	
1	*37.62*	*36.85*	37.235	0.770	2.068	
2	*16.18*	*16.56*	16.370	-0.380	2.321	
3	*28.82*	*28.65*	28.735	0.170	0.592	
4	*4490*	*4413*	4451.500	77.000	1.730	
5	*135.7*	*124.7*	130.200	11.000	8.449	
6	*62.56*	*62.25*	62.405	0.310	0.497	
7	*158.9*	*159.2*	159.050	-0.300	0.189	
8	*16540*	*16080*	16310.000	460.000	2.820	
9	*31.26*	*30.12*	30.690	1.140	3.715	
10	*58.49*	*60.11*	59.300	-1.620	2.732	
11	*740.5*	*796.2*	768.350	-55.700	7.249	
12	*130.3*	*126.9*	128.600	3.400	2.644	
13	*29.35*	*29.19*	29.270	0.160	0.547	
14	*1372*	*1388*	1380.000	-16.000	1.159	
15	*36.55*	*44.74*	40.645	-8.190	20.150	
16	*22.57*	*23.37*	22.970	-0.800	3.483	
17	*34.75*	*33.15*	33.950	1.600	4.713	
18	*92.93*	*94.01*	93.470	-1.080	1.155	
19	*40.6*	*42.23*	41.415	-1.630	3.936	
20	*80.36*	*86.36*	83.360	-6.000	7.198	
21	*15.76*	*18.54*	17.150	-2.780	16.210	
22	*78.22*	*73.76*	75.990	4.460	5.869	
23	*48.89*	*50.91*	49.900	-2.020	4.048	
24	*17.65*	*16.72*	17.185	0.930	5.412	
25	*36.56*	*35.3*	35.930	1.260	3.507	
26	*51.89*	*52.2*	52.045	-0.310	0.596	
27	*197.5*	*206.5*	202.000	-9.000	4.455	
28	*70.32*	*69.22*	69.770	1.100	1.577	
29	*29.99*	*30.62*	30.305	-0.630	2.079	
30	*31.9*	*32.36*	32.130	-0.460	1.432	
		Mean:	**816.331**		**4.0843**	= mean range (%)

s(r) % = range(mean)/1.128 = **3.62** %

Appendix 6: Raw data for oxygen in Section 4.4

Data plotted in Figure 3. "Range" equals the absolute value of the difference between Result 1 and Result 2.

Res. 1 mg/L	Res. 2 mg/L	Range mg/L
8.90	8.91	0.01
8.99	9.01	0.02
8.90	8.90	0.00
9.11	9.12	0.01
8.68	8.64	0.04
8.60	8.51	0.09
8.81	8.81	0.00
8.02	8.00	0.02
7.05	7.08	0.03
6.98	7.01	0.03
7.13	7.16	0.03
6.79	6.78	0.01
6.55	6.53	0.02
4.68	4.68	0.00
5.28	5.33	0.05
7.42	7.40	0.02
7.62	7.63	0.01
5.88	5.88	0.00
6.03	6.06	0.03
6.33	6.33	0.00
5.90	5.90	0.00
6.24	6.27	0.03
6.02	6.02	0.00
9.13	9.11	0.02
9.10	9.14	0.04
8.50	8.44	0.06
8.73	8.71	0.02
8.09	8.09	0.00
7.56	7.58	0.02
6.30	6.32	0.02
6.43	6.44	0.01
7.25	7.34	0.09
7.28	7.31	0.03
8.00	8.03	0.03
8.38	8.29	0.09
9.23	9.29	0.06
9.09	9.08	0.01
9.37	9.36	0.01
9.38	9.37	0.01
9.32	9.25	0.07
8.47	8.49	0.02
8.27	8.28	0.01
8.37	8.31	0.06
8.09	8.15	0.06
8.05	8.03	0.02
7.38	7.40	0.02
7.49	7.49	0.00
4.52	4.49	0.03
4.45	4.44	0.01
4.29	4.27	0.02
	mean range:	0.026
	mean range/1.128:	0.024

Appendix 7: Raw data for BOD in example A and B

Results in mg/L O_2 consumption. The certified value of the CRM is 206 ± 5 mg/L. As the average of two results is always reported for ordinary samples, the s is also calculated from the average of each sample pair in the internal quality control.

Date	Res. 1	Res. 2	Average
12-09-00	218.90	214.77	216.84
01-03-01	206.46	220.83	213.65
13-03-01	221.18	210.18	215.68
02-04-01	215.00	206.50	210.75
14-08-01	194.96	218.03	206.50
05-09-01	218.65	216.55	217.60
19-09-01	223.86	212.19	218.03
16-10-01	215.58	213.01	214.30
07-11-01	196.26	214.93	205.60
28-11-01	210.89	206.89	208.89
11-12-01	228.40	222.73	225.57
13-12-01	206.73	229.03	217.88
15-01-02	207.00	208.47	207.74
22-01-02	224.49	213.66	219.08
30-01-02	201.09	214.07	207.58
11-02-02	218.83	223.13	220.98
06-03-02	216.69	218.22	217.46
18-09-02	206.36	227.96	217.16
02-10-02	215.21	226.18	220.70
		Average:	214.84
		s:	5.58
		s%:	2.60

Appendix 8: Estimation of standard deviation from range

Number of samples	Factor , d_2	**Estimation of standard deviation from range (max-min), /1/ and /13, page 11/.**
n=2	1.128	
n=3	1.693	
n=4	2.059	
n=5	2.326	The standard deviation, s can be estimated from
n=6	2.534	
n=7	2.704	$s = \frac{range}{d_2}$
n=8	2.847	
n=9	2.970	
n=10	3.078	
For comparison		
Rectangular interval	3.464	(Example, see Appendix 5 and 6)
95 % conf. limit.	3.92	

TECHNICAL REPORTS FROM EXPERT GROUP QUALITY AND METROLOGY

351 Guðmundsson, H., Calibration and uncertainty of XRF measurements in the SEM. Espoo 1997. Nordtest, NT Techn Report 351. 42p. NT Project No. 1280-96.

352 Rydler, K.-E. & Nilsson, H., Transparency of national evaluation methods of accreditation in the field of electrical measurements. Espoo 1997. Nordtest, NT Techn Report 352. 61p. NT Project No. 1255-95.

353 Tambo, M. & Søgaard, T., The determination of gas density - part 1: State of the art in the Nordic countries. Espoo 1997. Nordtest, NT Techn Report 353. 19p. NT Project No. 1254-95.

354 Tambo, M. & Søgaard, T., The determination of gas density - part 2: Intercomparison - Calibration of gas density meters with nitrogen. Espoo 1997. Nordtest, NT Techn Report 354. 91 p. NT Project No. 1254-95.

355 Tambo, M. & Søgaard, T., The determination of gas density - part 3: A guideline to determination of gas density. Espoo 1997. Nordtest, NT Techn Report 353. 29p. NT Project No. 1254-95.

357 Källgren, H., Laboratory weighing - Procedures according to OIML international;
Weights of classes E_1, E_2, F_1, F_2, M_1, M_2, M_3 - Part 1: Testing procedures for weights, - Part 2: Pattern evaluation report. Espoo 1997. Nordtest, NT Techn Report 357. 82p. NT Project No. 1249-95.

379 Ravn, T., Pitkäniemi, M., Wallin, H., Bragadóttir, M., Grønningen, D. & Hessel, H., Checking of UV/VIS spectrophotometers. Oslo 1998. Nordic committee on food analysis, NMKL Procedure No.7 (1998). 14 p. NT Project No. 1339-96.

403 Holmgren, M., Observing validation, uncertainty determination and traceability in developing Nordtest test methods. Espoo 1998. Nordtest, NT Techn Report 403. 12 p. NT Project No. 1277-96.

418 Views about ISO/IEC DIS 17025 - General requirements for the competence of testing and calibration laboratories. Espoo 1999. Nordtest, NT Techn Report 418. 87 p. NT Project No. 1378-98.

419 Virtanen, V., Principles for measuring customers satisfaction in testing laboratories. Espoo 1999. Nordtest, NT Techn Report 419. 27 p. NT Project No. 1379-98.

420 Vahlman, T., Tormonen, K., Kinnunen, V., Jormanainen, P. & Tolvanen, M., One-site calibration of the continuous gas emission measurement methods at the power plant. Espoo 1999. Nordtest, NT Techn Report 420. 18 p. NT Project No. 1380-98.

421 Nilsson, A. & Nilsson, G., Ordering and reporting of measurement and testing assignments. Espoo 1999. Nordtest, NT Techn Report 421. 7 p. NT Project No. 1449-99.

430 Rasmussen, S.N., Tools for the test laboratory to implement measurement uncertainty budgets. Espoo 1999. Nordtest, NT Techn Report 430. 73 p. NT Project No. 1411-98.

431 Arnold, M., Roound robin test of olfactometry. Espoo 1999. Nordtest, NT Techn Report 431. 13 p. NT Project No. 1450-99.

429 Welinder, J., Jensen, R., Mattiasson, K. & Taastrup, P., Immunity testing of integrating instruments. Espoo 1999. Nordtest, NT Techn Report 429. 29 p. NT Project No. 1372-97.

443 Guttulsrød, G.F, Nordic interlaboratory comparison measurements 1998. Espoo 1999. Nordtest, NT Techn Report 443. 232 p. (in Dan/Nor/Swed/Engl) NT Project No. 1420-98.

452 Gelvan, S., A model for optimisation including profiency testing in the chemical laboratories. Espoo 2000. Nordtest, NT Techn Report 452. 15 p. NT Project No. 1421-98.

537 Magnusson, B., Näykki, T., Hovind, H. & Krysell, M., Handbook for calculation of measurement uncertainty in environmental laboratories. Espoo 2003. Nordtest, NT Techn Report 537. 41 p. NT Project No. 1589-02.

NORDTEST
TECHNICAL REPORT 537

Nordtest, founded in 1973, is an institution under the Nordic Council of Ministers and acts as a joint Nordic body in the field of conformity assessment. The emphasis is on the development of Nordic test methods and on Nordic co-operation concerning conformity assessment. The main task is to take part in the development of international pre-normative activity. Nordtest is yearly funding projects in its field of activity.

Nordtest endeavours to

- promote viable industrial development and industrial competitiveness, remove technical barriers to trade and promote the concept "Approved Once Accepted Everywhere" in the conformity assessment area
- work for health, safety, environment in methods and standards
- promote Nordic interests in an international context and Nordic participation in European co-operation
- finance joint research in conformity assessment and the development and implementation of test methods
- promote the use of the results of its work in the development of techniques and products, for technology transfer, in setting up standards and rules and in the implementation of these
- co-ordinate and promote Nordic co-operation in conformity assessment
- contribute to the Nordic knowledge market in the field of conformity assessment and to further development of competence among people working in the field

Nordic
COUNCIL OF MINISTERS